

THEMATIC COMPOSITION AND IDIOM VARIATION

ATTILA CSERÉP

University of Debrecen

Abstract: *The Corpus of Contemporary American English (COCA) has been studied to retrieve variant forms of semantically decomposable idioms that have no thematic composition for the purpose of determining whether thematic composition is a necessary criterion for idiom variation as claimed by Horn (2003). The syntactic variants searched for include passive, raising, tough-movement, relative clauses and wh-questions. Horn's (2003) hypothesis is not fully confirmed, as some variation has been found.*

Keywords: *decomposability, idioms, syntactic variation, thematic composition*

1. Introduction

Idiom variation studies have produced a number of theories concerning the properties that determine the flexibility of idioms. One of the hypotheses proposed by Horn (2003:248-249) uses the property of thematic composition, as well as transparency, as a predictor of idiom variability. As he explains, “[a]n expression has thematic composition if the thematic structure of the verb in its literal sense and that of the verb in its idiomatic sense are identical” (Horn 2003:249). Thematic structure is defined as “the set of semantic roles that a verb assigns to its NP arguments” (Horn 2003:249). An expression that has thematic composition is *break the ice*, where both in the literal and idiomatic senses there is an idea of the removal, or breaking (down), of a rigid entity (Horn 2003:249). *Break the ice* can therefore undergo certain syntactic transformations such as passivization. In contrast, *grasp the nettle* has no thematic composition and cannot therefore vary syntactically. Horn (2003:249) suggests that grasping in its literal sense involves taking hold of an object, but “confronting something does not involve the same sort of action”. Both types of idioms share the property of decomposability, (metaphorical) semantic composition in Horn's (2003:247) terminology, which enables speakers to assign idiomatic senses to the individual idiom components after the metaphorical meaning is known. Thus, *break* means ‘break down’ and *ice* means ‘a fragile/rigid barrier to social interaction’ (Horn 2003:247). The purpose of this paper is to find corpus evidence that confirms or refutes Horn's (2003:248-249) claim that thematic composition rather than semantic composition is a necessary condition for syntactic variation.

2. Idiom Variation and Thematic Composition

2.1.Semantic composition and thematic composition

Horn (2003) distinguishes three major types of idioms based on the properties of semantic composition and thematic composition (see Figure 1). Some idioms lack metaphorical semantic composition; in other words, the idiom components do not carry meaning, such as *kick the bucket* or *shoot the bull*. They are claimed to be fixed (Horn 2003:247). There are expressions that have both semantic and thematic composition, and a subdivision is made within this type on the basis of transparency.

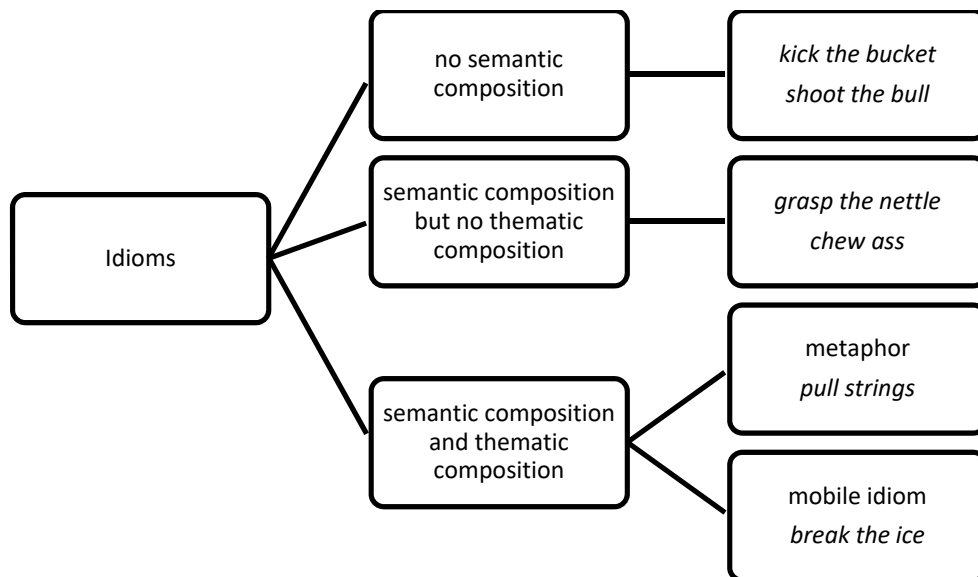


Figure 1. Types of idioms distinguished by Horn (2003)

Transparent expressions are called “metaphors”, and they can undergo several syntactic operations such as passivization, raising, *tough*-movement, relativization and *wh*-question (Horn 2003:261-262). All these syntactic changes affect the surface position of the idiomatic noun phrase by moving it before the verb, as can be seen in the examples below taken from Horn (2003:247, 261-262).

- (1) After a few beers, the ice was broken.
- (2) Strings seem to be pulled every time he applies for a promotion.
- (3) The bandwagon was easy to jump on.
- (4) Bill pulled the same strings that Joe pulled to get the promotion.
- (5) Which bandwagon will Fred jump on this week?

Sentence (1) illustrates the passive, example (2) shows raising, (3) is a case of *tough*-movement, while (4) and (5) illustrate relativization and *wh*-question, respectively. Note that raising necessarily involves passivization, since the object noun phrase has to be moved to subject position during passivization before this subject can be raised.

In addition to metaphors, there are semantically and thematically compositional idioms that are more restricted. They are labelled “mobile idioms” and can undergo only three of the aforementioned five operations according to Horn (2003:262-263): passive, raising and *tough*-movement. As a result of these transformations, the idiomatic noun, which is an internal

argument of the verb, ends up in an external argument position as subject. Since these three types of syntactic operations are possible with the most restricted thematically compositional idioms, the same transformations are used as tests to distinguish thematically compositional expressions from those that have semantic composition but lack thematic composition (Horn 2003:247-248). These idioms are claimed to be recalcitrant to the three syntactic operations. To illustrate his point, Horn (2003) gives invented example sentences based on native speaker intuition. His claim is examined here using naturally occurring language data in a large corpus. If Horn (2003) is right, no instance of passive, raising and *tough*-movement is expected to occur with idioms of the latter type.

Horn's (2003:248-249) suggestion concerning the significant role of thematic composition is meant to be an improvement on previous claims that attribute a key role to the property of (metaphorical) semantic composition in determining the flexibility of idiomatic expressions. However, thematic composition is supposed to explain only some types of variation selected by Horn (2003) as test cases, whereas semantic composition has been proposed to account for more variation types, including the premodification or postmodification of the idiomatic noun.

2.2. The corpus and the idioms

Idioms occur relatively infrequently; therefore, a large corpus is necessary to study variation. The Corpus of Contemporary American English (COCA) has been selected, because it contains about 520 million words (Davies 2008–). It is a modern, balanced corpus of American English that contains five genres: spoken, fiction, popular magazines, news and academic English.

The idioms that have semantic composition but no thematic composition have been taken from Horn's (2003) discussion. All the V (Det) N expressions are included, while two prepositional idioms (*give the lie to*, *make a great show of*) are ignored. Table 1 shows the expressions. Horn's (2003:251, 271) discussion in the main body of the article gives *blow one's cool* but his appendix lists the same phrase with the verb *lose*. Both versions have been included. Idioms having no semantic composition are ignored, as the central issue addressed here is Horn's (2003:248-249) claim that semantic composition is not sufficient for syntactic variation and thematic composition is a better predictor of flexibility.

blow one's cool	keep one's cool
catch hell	kiss ass
chew ass	lose one's cool
drop a bomb	lose one's mind
eat one's words	make a face
grasp the nettle	raise hell
hit the hay	screw the pooch

Table 1. List of V (Det) N idiomatic expressions

Queries were made online using the query syntax supplied with the corpus. The texts were searched for the noun in the singular or plural form within the maximum span of 9 words to the left or right of the verbal constituent. Separate searches were performed to collect solid or hyphenated compounds. Since *make a face* had a very high frequency, only a sample of its occurrences was examined. It was decided to study all instances of the form *made* in combination with *face/faces*, because two of the three syntactic tests proposed by Horn (2003:247-248) – passive and raising – require the past participle form of the verb. Searching for *make*, *makes* or *making* to retrieve passive or raising would have been futile. The expression *grasp the nettle* is a British English idiom; therefore, in addition to COCA, the 100-million-word British corpus BNC was also checked through the same web interface (Davies 2004–). If the British and American varieties of *grasp the nettle* are considered to be separate expressions, and *blow one's cool* and *lose one's cool* are also regarded as different, then a total of fourteen idioms have been studied.

The decision was made to extract all occurrences of the idioms in preference to retrieving only forms which exhibit the syntactic operations in question. The methodology adopted here enables us to see the total occurrence of an idiom and what proportion of its occurrences instantiate the given variation. The search query included both the idiomatic verb and noun, as a result of which pronominalization could not be systematically retrieved. Some instances were found in the expanded context where the idiomatic noun was replaced by a pronoun, but they were ignored.

2.3. Results and discussion

The number of variant tokens and the total number of occurrences for each idiom are provided in Table 2. The low frequency of *blow one's cool* cannot help us assess its flexibility. For this reason, it will be ignored in the discussion below. *Screw the pooch* is also rather infrequent, which means that the data for this idiom has to be interpreted with some caution.

Most idioms occur predominantly in their unvaried form. The passive can be found with four of the idioms but with one of them only one instance is attested. A single occurrence is not usually regarded as firm evidence for the acceptability of a variant. Passive forms of *make a face* are not common, but *drop a bomb* and, in the British corpus, *grasp the nettle* are more common in the passive. Only finite passives are shown in Table 2, but one nonfinite passive also occurs (*and there were no faces being made*), and two instances of *get/want one's N Ved* are attested with *chew ass* and *kiss ass* each, as in (6).

(6) ...and now he was going to get his ass chewed.

Idiom	Total	Passive		Raising		Tough-movement		Relative clause		Wh-question	
			%		%		%		%		%
blow one's cool	8	0	0	0	0	0	0	0	0	0	0
catch hell	95	0	0	0	0	0	0	2	2	0	0
chew ass	21	0	0	0	0	0	0	0	0	0	0
drop a bomb	149	3	2	0	0	0	0	8	5	0	0
eat one's words	87	0	0	0	0	0	0	0	0	0	0

grasp the nettle AmE	16	0	0	0	0	0	0	2	13	0	0
grasp the nettle BrE	54	6	11	0	0	0	0	2	4	0	0
hit the hay	36	0	0	0	0	0	0	0	0	0	0
keep one's cool	223	0	0	0	0	0	0	0	0	0	0
kiss ass	133	0	0	0	0	0	0	1	0.8	1	0.8
lose one's cool	218	0	0	0	0	0	0	0	0	0	0
lose one's mind	1022	0	0	0	0	0	0	0	0	0	0
make a face	812	3	0.4	0	0	0	0	24	3	0	0
raise hell	409	1	0.2	0	0	0	0	3	0.7	2	0.5
screw the pooch	13	0	0	0	0	0	0	0	0	0	0

Table 2 Total occurrences of idioms and variants

Raising does not occur. The only example that comes close to it is (7). As the finite verb is missing, it has not been classified as raising.

(7) He puts his hands out in front of him like a nervous salesman fearing the kibosh, calming things, the pooch about to be screwed.

Horn's (2003) third syntactic test, *tough*-movement, is not attested. While the absence of raising and *tough*-movement supports Horn's (2003:248-249) claim that thematic composition is a prerequisite to these transformations, passive forms can be found. The original claim is not fully confirmed.

Horn (2003:261-263) views relative clauses and *wh*-questions as grammatical only with the most flexible idiom type: metaphors. He does not use these operations as tests for thematically noncompositional expressions, but it is logical to assume that if relative clauses and *wh*-questions are unacceptable with non-transparent thematically compositional idioms (mobile idioms), they are also unacceptable with idioms that lack thematic composition. Furthermore, Horn's (2003:261-263) treatment of syntactic variation in metaphors and mobile idioms, discussed in Section 2.1, suggests that if an expression permits these operations, it should also permit passive, raising and *tough*-movement, but not vice versa. Neither claim is supported by corpus evidence. Relativization and *wh*-questions are found with some of the idioms in Table 2, though not with all. If we ignore single instances of a variant, relative clauses occur with six of the thirteen expressions, while *wh*-question is attested with one idiom. Some examples are given below.

(8) ...and the hell he would catch from Master Trevant for letting Saturn escape.

(9) This is the nettle that the president doesn't want to grasp...

(10) Whose ass do you have to kiss to get a deal like that?

Relative clauses also occur with *catch hell*, *grasp the nettle* (AmE) and *kiss ass*, though only one example is found of the latter, but the same idioms are not attested in passive, raising or *tough*-movement, which are all claimed to be more common. The corpus data reveal that with the exception of *grasp the nettle* in the British corpus, relativization is actually more common than the passive. Horn's (2003:261-262) relative clauses all contain the relative pronoun *that*, but the zero pronoun is much more common in the corpus. It outnumbers *that* by 36 to 5.

Ignoring low-frequency *blow one's cool* and instances where a single occurrence of a variant is found, of the thirteen expressions six exhibit some type of variation thought to be ungrammatical by Horn (2003). It could be argued that while thematic composition is not a necessary condition for syntactic variation of the type studied by Horn (2003), it is a slightly better predictor than semantic composition, because it accounts for the data of the remaining seven idioms in Table 2. However, in the pertaining literature, semantic composition is not claimed to be a sufficient criterion for flexibility. Nunberg (1978:127) states that semantic composition does not necessarily ensure the passivizability of an idiom. Nor is it claimed to be the only factor affecting variability. Transparency and discourse function are also proposed to be significant (Gibbs and Nayak 1989:125; Nunberg et al. 1994:509). Therefore, the absence of variant forms in several rows of Table 2 is not a decisive argument against the role of semantic composition. Furthermore, the above discussion assumed that Horn's (2003) categorization of the expressions in Table 2 as semantically compositional (but thematically non-compositional) is correct. It is with this background assumption that the corpus evidence has been examined. A closer look on the expressions casts some doubt on this assumption. *Chew ass* and *kiss ass* lack semantic composition, despite Horn's (2003:248) meaning paraphrases: [administer/deliver] [a reprimand] for *chew ass* and [curry] [favour] for *kiss ass*. Semantic composition does not simply reside in attaching the verbal idiom meaning to the verb and the nominal chunk of the meaning to the noun. It should be based on metaphorical correspondences. Chewing ass as a painful physical act is metaphorically mapped onto a psychologically an emotionally painful experience, but ass does not correspond to reprimand. Similarly, kissing ass can be regarded as a hyperbolic act of currying favour with a person, perhaps as a subtype of the act of ingratiating oneself, but ass is not mapped onto favour. It metonymically stands for the person. The frozenness of these idioms can be satisfactorily explained by their lack of semantic composition. Any reference to thematic composition is unnecessary. Given its lack of semantic composition, what is remarkable is that *kiss ass* is not completely frozen. Altogether six semantically compositional idioms occur in some variant form and six idioms are frozen, if we ignore *blow one's cool* (due to its very low frequency), *chew ass* and *kiss ass* (due to their lack of semantic composition) and single occurrences of a variant. The variation that has been found for a given idiom is undoubtedly not very common, most tokens occur without any syntactic operations.

Horn (2003:267) himself notes that sentences like (11) and (12) seem to be counterexamples to his hypothesis. The idioms in these examples lack thematic composition but exhibit *tough*-movement.

(11) Humble pie is never easy to eat.

(12) One's words are never easy to eat.

However, he dismisses (11) and (12) as conclusive evidence, as neither *eat humble pie* nor *eat one's words* undergo other syntactic operations he uses as diagnostic tests (Horn 2003:267-268). Interestingly, perhaps because of these counterexamples, he later modifies his claim: "all expressions that lack the property of thematic composition display, at most, severely

limited mobility” (Horn 2003:271). This is more in line with the corpus evidence, but note that it is not *tough*-movement that tends to occur with the idioms in Table 2. To what extent the variation is limited can only be judged if data about the same syntactic operations of thematically compositional idioms is examined, but this is outside the scope of the paper.

3. Conclusion

This corpus study of the syntactic variation of semantically compositional but thematically noncompositional idioms has been conducted to find natural language evidence for Horn’s (2003:248-249) hypothesis that thematic composition is a necessary condition for certain types of syntactic variation in idioms. Five types of variation have been examined: passive, raising, *tough*-movement, relative clauses and *wh*-questions. A total of fourteen expressions have been subjected to analysis, of which one (*blow one’s cool*) is so infrequent that the corpus data cannot be regarded as informative, and two idioms (*chew ass*, *kiss ass*) seem to be semantically noncompositional. Of the remaining twelve idioms, six exhibit some sort of variation, though the proportion of variant tokens is often very low. The evidence suggests that Horn’s (2003) hypothesis of the significant role of thematic composition is not fully confirmed. Future research should include a larger sample of idioms to explore the issue more deeply.

References

- Davies, Mark. (2004–) *BYU-BNC. (Based on the British National Corpus from Oxford University Press)* [Online]. Available <http://corpus.byu.edu/bnc/> [Last accessed 2016, July 28].
- Davies, Mark. (2008–) *The Corpus of Contemporary American English: 520 million words, 1990-present* [Online]. Available <http://corpus.byu.edu/coca/> [Last accessed 2016, July 31].
- Gibbs, Raymond W., Jr. and Nandini P. Nayak. 1989. “Psycholinguistic studies on the syntactic behavior of idioms” in *Cognitive Psychology* 21(1):100–138.
- Horn, George M. 2003. “Idioms, metaphors and syntactic mobility” in *Journal of Linguistics* 39(2):245–273.
- Nunberg, Geoffrey D. 1978. *The Pragmatics of Reference*. Bloomington: Indiana University Linguistics Club.
- Nunberg, Geoffrey, Ivan A. Sag and Thomas Wasow. 1994. “Idioms” in *Language* 70(3):491–538.

Notes on the author

Attila Cserép is Senior Lecturer at the Department of English Linguistics, Institute of English and American Studies, University of Debrecen. His primary research interest is phraseology, especially the syntax and semantics of idioms. His current research focus lies in discovering what properties of an idiom affect its variability and how idioms are motivated.