A Nanosyntax Analysis of Idioms   Will Nediger

Idioms have presented a long-standing puzzle for syntactic theory: they behave in some ways like individual lexical items (in that their meaning is not compositionally predictable), and in some ways like phrases (in that they show evidence of having internal syntactic structure). In lexicalist theories, in which the syntax is fed by atomic lexical items, this behavior is paradoxical – it is difficult to explain how lexical items could have internal syntactic structure, or how entire phrases could be interpreted non-compositionally. I present an analysis of idioms which seeks to solve this puzzle within the framework of Nanosyntax (Starke 2009, 2011).

Nanosyntax is a non-lexicalist theory, in which the syntactic derivation is not fed by the lexicon. Syntactic structure is built by Merge of individual features, as opposed to lexical items. Lexical items then spell out subtrees of the syntactic structure. A lexical item consists of a phonological string, semantic features, and a subtree; the lexical item can spell out that subtree, or any tree containing that subtree. Spell-Out is cyclic, so after every merger operation, the entire tree is spelled out. Each successful spellout overrides previous spellouts, and a tree is spelled out with as few lexical items as possible.

I argue that the paradox of idioms receives a natural explanation in a Nanosyntax framework. Since all lexical items spell out syntactically complex structures in Nanosyntax, there is no conflict between lexicality and syntactic complexity. I argue that idioms are stored in the lexicon, which accounts for the fact that their meaning is not compositionally predictable, and that they therefore spell out subtrees, giving them internal syntactic structure.

However, not all idioms are created equal. Consider the well-known contrast between the behavior of idioms like (1), which cannot undergo syntactic operations, and those like (2), which can:

(1)   a. He kicked the bucket (idiomatic and literal readings)
        b. The bucket was kicked (only literal reading)
(2)   a. She spilled the beans (idiomatic and literal readings)
        b. The beans were spilled (idiomatic and literal readings)

If the idiom *spill the beans* were to be stored as a single lexical item, the architecture of Nanosyntax would predict (2b) to be ungrammatical unless *the beans were spilled* was also stored in the lexicon, because passivization would break up the subtree spelled out by *spill the beans*, so that idiom would not be able to spell out the post-passivization structure. Given that it is clearly unparsimonious to store every transformed variant of *spill the beans* in the lexicon, I argue that *spill the beans*-type idioms are lexically stored as multiple chunks.

In particular, I argue that idioms are lexically decomposed to the extent that they can be semantically decomposed. As is well known, the ability of idioms to undergo syntactic operations correlates very strongly with the extent to which their meaning can be distributed among their parts (Nunberg et al. 1994). Thus, I propose the following principle:

(3)   *Atomicity of lexical items*
        No lexical item (consisting of a sound/meaning pair and the subtree which it can spell out)
        can be decomposed into two or more possible lexical items.

Thus, for example, *spill the beans* is not a single lexical item because it can be decomposed into two idiom chunks, *spill* (meaning "divulge") and *the beans* (meaning "a secret"). In contrast, *kick the bucket* is a single lexical item because it cannot be decomposed, since idiomatic meanings cannot be individually

identified for *kick* and *the bucket*. (3) ensures that the ability of idioms like (2) to undergo syntactic transformations follows from their ability to be semantically decomposed.

In order to explain why idiom chunks have co-occurrence restrictions (e.g. why *divulge the beans* is ungrammatical), I follow Nunberg et al. (1994) in arguing that co-occurrence restrictions on idiom chunks are semantic in nature. As Nunerg et al. argue, the idiom *spill the beans* consists of a literal spilling-the-beans meaning which is metaphorically mapped to a divulging-the-secret meaning. If *spill* and *the beans* do not co-occur, the literal meaning is unavailable. Under this approach, the existence of idiom families, such as that in (4), is expected – some idiom chunks should be able to co-occur with multiple lexical items, as long as those items all satisfy the semantic co-occurrence restriction.

(4)   get off one's ass, get off one's tush, get off one's rear, get off one's butt

(Nunberg et al. 1994:504)

I also contrast my approach with approaches to idioms in a similar non-lexicalist framework, Distributed Morphology (Halle and Marantz 1993). The key architectural difference between the two frameworks is that DM does not allow spellout of non-terminals, while Nanosyntax does (since features themselves are terminal nodes). Following Caha (2008), I argue that DM's assumption that only terminals can be spelled out forces it to rely on morphological processes such as Morphological Merger (Marantz 1989), which are not necessary in Nanosyntax.

According to the standard DM treatment of idioms, the Encyclopedia states that, for example, *kick* takes on the meaning of "die" in the context of *bucket*, and *bucket* correspondingly takes on no meaning at all in the context of *kick*. I argue that this treatment is at odds with our intuitions about the distributivity of meaning (that is, we have an intuition that the idiom as a whole means "die," not just *kick*) and that it does not provide a natural semantic explanation of the difference between idioms like (1) and those like (2). I also adopt the argument of Williams (2007) that there exist verbal idioms in which the idiomatic meaning cannot be assigned to the verbal root, because the verb receives its literal meaning in the context of the idiom. Williams gives the example in (5):

(5)   a. [[ ]$_V$ this!]$_{VP}$: "as far as V-ing goes, screw you!"
         b. Format this!

Following Williams, I argue that it is not clear in DM how the idiomatic meaning of (5) should be distributed, and that the Nanosyntax approach is therefore preferable.

References
Caha, P. 2008. An argument for the spell-out of non-terminals. Talk given at Workshop on Theoretical Morphology 4. Leipzig: June 20. ⋯ Halle, M. and Marantz, A. 1993. Distributed Morphology and the Pieces of Inflection. In K. Hale and S. Keyser (eds.), *The View from Building 20*, 111-176. Cambridge, MA: MIT Press. ⋯ Marantz, A. 1989. Clitics and phrase structure. In M. Baltin and A. Kroch (eds.), *Alternative Conceptions of Phrase Structure*, 99-116. Chicago: University of Chicago Press. ⋯ Nunberg, G., Sag, I. and Wasow, T. 1994. Idioms. *Language* 70(3): 491-538. ⋯ Starke, M. 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36(1): 1-6. ⋯ Starke, M. 2011. Towards an elegant solution to language variation: Variation reduces to the size of lexically stored trees. Unpublished Ms., Barcelona. ⋯ Williams, E. 2007. Dumping Lexicalism. In G. Ramchand and C. Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces*, 353-382. Oxford: Oxford University Press.