Two types of multi-lexeme derivatives in Spanish

In this paper we compare morphological properties of lexical blends and compounds in Spanish and find that blends differ from productive NN compounds, but are similar to a different type of compounds that behave more “wordlike” and always respect lexical integrity. We conclude that blends reside in morphology while productive compounds in Spanish are syntactic.

Lexical blends are derivatives formed by intentional truncation and concatenation of two words: usually a beginning of one source word (sw1) is attached to the end of another source word (sw2) as in \(\text{spoon}_{sw1} + \text{fork}_{sw2} = \text{spork}\). Blends resemble compounds because they involve combination of two lexemes and because the semantic relationships between the lexemes are analogous to those found in compounds (e.g., both can have semantic heads: a \(\text{motel = motor + hotel}\) is a type of a hotel). However, the grammatical status of blends is controversial. Is blending a word-formation process resulting in structured units that behave similarly to compounds? Or are they created by an extra-linguistic mechanism that is different from the typical syntactic and/or morphological operations? We address these questions by comparing morphological properties of Spanish NN blends and compounds.

The morphology of Spanish provides us with an opportunity to compare blends and compounds along several dimensions. First, Spanish endocentric NN compounds are known to be mostly left-headed. So, we can compare whether NN blends have a similar distribution with respect to semantic heads. (This is an interesting comparison because the literature on blends mostly comes from languages that have right-headed blends and compounds.) Second, Spanish has gender, a morphological property that is expected to be controlled by the syntactic head of a construction. This allows us to test whether the gender of blends and compounds matches the gender of their semantic head. Following Bisetto and Scalise 2005, we assume that semantic and syntactic heads typically (but not necessarily) coincide with one another. Our null hypothesis was that blends and compounds are produced by the same grammatical mechanism which combines two inflected words in Spanish and, hence, both types of derivatives should have similar morphological properties.

To examine properties of blends, we collected data from Casado Velarde 1985, Piñeros 2004, and Rodríguez González 1989, which resulted in a database of 370 blends (some examples are given below).

\begin{enumerate}
  \item \(\text{burricicleta = burro ‘donkey’ + bicicleta ‘bicycle’ = ‘small town people’s mode of transportation,’ EXOCENTRIC}\)
  \item \(\text{amigovio = amigo ‘friend’ + novio ‘boyfriend’ = ‘boyfriend’, RIGHT-HEADED}\)
  \item \(\text{italián = italiano, m. ‘Italian’ + español, m. ‘Spanish’ = ‘a mix of Italian and Spanish,’ COORDINATE}\)
\end{enumerate}

Because the sources did not always allow us to determine semantic headedness or gender of a blend, and because many blends we collected were neologisms, we designed a survey to determine how speakers of Spanish interpret these blends. For this purpose we selected a sample of 40 blends from our database, focusing on those for which sw1 and sw2 disagreed in gender. Survey participants were native speakers of Spanish recruited over the web. After being familiarized with the notion “blend,” participants were presented with examples of blends like this: “\(\text{burricicleta is a blend of the words burro and bicicleta.}\)” We asked “\(\text{Is burricicleta a type of (i) burro, (ii) bicicleta, (iii) neither, (iv) both?}\), “\(\text{What do you think burricicleta means?}\), “\(\text{Which article would you use with burricicleta (i) la (ii) el?}\), and “\(\text{Have you ever heard of this word before?}\)”. Data from 64 people who answered questions about 20 blends each, was included in the analysis. Results: While speakers were mostly not familiar with the presented blends, they had consistent intuitions about their headedness and
gender. In 780/1023 (76%) responses blends were treated as endocentric, with 60% being right-headed and 40% left-headed. The rate of interspeaker agreement on an individual blend’s headedness ranged between 60% and 87% for all but a handful of blends. A mixed logistic regression (with subject as a random effect) showed that from the two predictors of blend gender, gender of sw2 and gender of the head, the first was by far the strongest (log odds estimate = -3.12, st.error = 0.3, p < 0.01), although both were significant with no interaction between them. Looking at the left-headed blends for which the predictors disagree, we see that these blends match the gender of sw2 rather than the gender of the head (sw1) by the ratio of 2:1, while right-headed blends match the gender of sw2 (which is also the head) by a ratio of 9:1. Thus, the gender of sw2 overrides the gender of the semantic head, but the latter still appears to have a weak effect on blend gender.

The results above are inconsistent with patterning of productive Spanish NN compounds. Our quantitative data on compounds comes from 247 NN compounds from Moyna’s (2000) dissertation. This data includes both the productive NN compounds and the more rare, unproductive ones, many of which are archaic or specialized terms, but still semantically transparent to the speakers. The productive compounds typically consist of two inflected nouns spelled as two words, each of which maintains its own stress (e.g., piedra imán ‘stone, f.’ + ‘magnet, m.’, ‘mineral magnet, f’). The unproductive ones tend to be spelled as one word, have a single main stress, and often have linking vowels (Guevara 2012, Moyna 2004) (cf. 2). None of the compounds spelled as two words in our data were right-headed, while 39% of compounds spelled as one word were right-headed. The gender of two-word compounds always agrees with the gender of the first constituent, even if the compound is exocentric (e.g., bobo escama ‘fool, m. + scale, f.’, type of fish, m.). As a rule of thumb, such compounds are also pluralized by head marking (Guevara 2012). On the other hand, the gender of the one-word compounds tends to agree with the second constituent (see examples in 2), and these compounds are pluralized at the right edge of the word (e.g., tel-aráña-s ‘cloth-spider-pl’ = ‘spiderwebs’).

(2) a. aicola: ajo ‘garlic, m’ + cola ‘glue, f’ = ‘glue made with garlic, f’

b. bocacaz: boca ‘mouth, f’ + caz ‘irrigation canal, m’ = ‘opening in the dam or river, m’

To sum up, blends have different properties from productive compounds which appear to be syntactically left headed (the head determines gender, semantic headedness also appears on the left, compound-internal headmarking of plural is possible). On the other hand, blends behave similarly to the lexicalized compounds in several ways: both form a single phonological word with one main stress, both can be semantically right- and left-headed, for both gender is determined by the right-most N regardless of semantic headedness, and both are pluralized at the right edge. Therefore, we conclude that while productive NN compounds are formed in syntax, blends are formed in a different component of grammar, one where special prominence is given to the right-most stem. This difference cannot be attributed to the difference in the kinds of units that are being combined since both blending and compounding involve combinations of inflected words. A possible alternative would be to say that because blending involves truncation of sw1 (including crucially the inflection), this hampers the ability of sw1 to act as a syntactic head. However, the unproductive compounds that form a single phonological word behave like blends and are not truncated (some of them preserve the inflection on sw1 as in example 2b). Additionally, it’s not clear why semantic headedness should be affected by truncation. Further work is in progress to better understand the morphological and syntactic differences between two-word vs. one-word compounds in Spanish and to quantify their productivity.