On the phonological status of Spanish compound words

Rajiv Rao

Abstract
Primarily inspired by Bustos de Gisbert (1986), Hualde (2006/2007), and Moyna (2011), specifically by their comments on stress deletion in the left members of some Spanish compounds, the current investigation fills a gap in this field by conducting an acoustic analysis of fundamental frequency evidence of stress in Spanish compounds and phonologically classifying them through the Autosegmental-Metrical model. Using a data elicitation task of 30 noun+noun compounds demonstrating syntactic, semantic, orthographic, and phonological variation, eight speakers of Mexican Spanish recorded two iterations of each compound, which is embedded at the beginning or at the end of a declarative carrier phrase. The acoustic analysis reveals that, as expected, right members categorically exhibit F0 evidence of stress. However, only the 15 compounds written as two orthographic units show strong evidence of left stress. This is especially so in utterance-initial position. The number of unstressed syllables between left and right stresses determines the tonal sequences produced in left members. In compounds spelled as two orthographic units, the phonological targets of each member resemble those of two simple content words in broad focus declaratives. Evidence of left stress in orthographically united compounds occurs in less than 20% of cases, and these are viewed as carrying postlexical secondary stress. The tonal sequences of this group are more wide-ranging than those of the first group. The results have implications for language processing and raise questions for the study of compound stress in Spanish.

1. Introduction
In the most general of terms, compounding in Spanish and other Romance languages has been described as a creative process whereby two or more words or words and stems form a longer word (Alemany Bolufer 1920: 152–4, Montes Giraldo 1968: 23–4, Manteca 1987: 333–5, Lang 1990: 65–72, among many others). However, seminal
large-scale studies on this topic for Spanish, such as Bustos de Gisbert (1986) and Moyna (2011), demonstrate that a precise definition of compounds based on morphological, (morpho)syntactic, semantic, orthographic, and prosodic criteria is difficult to achieve. While the former two authors take into consideration all these areas in their research on Spanish compounds, detailed experimental characterizations of the prosody of compounds and variables that may affect it remain relatively unexplored. Thus, Moyna (2011: 287) concludes that ‘[t]he area of compound prosody is begging for fresh insights that will help clarify, confirm, or correct many previous assertions [...]’.

In the last few decades, research on Spanish prosody, especially on the acoustic correlates of stress and the intonation of various utterance types, has grown enormously thanks to frameworks of intonational phonology such as the Autosegmental–Metrical (hereafter, AM) model (Pierrehumbert 1980, Beckman & Pierrehumbert 1986, Pierrehumbert & Beckman 1988, Ladd 1996, 2008). Most such studies use laboratory approaches in order to control the effects of independent variables and reduce potential confounds. To date, Hualde (2006/2007) contains the only detailed study of the stress properties of Spanish compounds. It provides semantic and syntactic evidence to support his claim that certain classes of compounds delete their left member’s stress, making them just one prosodic word, while others maintain both stresses, and are thus two prosodic words. However, Hualde (2006/2007) does not give much detailed acoustic evidence to support his claims, nor does he focus on tonal differences between the compounds considered. Also, since Moyna’s (2011) large-scale study was published after Hualde (2006/2007), we are now aware of further variables mentioned in the former that could have been explored in the latter.

This paper is largely motivated by the work outlined to this point. It begins answering Moyna’s (2011) call for work on the prosody of compounds, and is intended to do so in a way that complements and expands upon Hualde (2006/2007). It bridges the gap between previous research couched in models of intonational phonology and our current knowledge of compounds. This paper details an acoustic analysis of Mexican Spanish speakers’ production of 30 noun + noun (hereafter, NN) compounds that result in another noun (that is, [NN]_N). This particular class of compound is of interest because, according to Bustos de Gisbert (1986: 182), its characterization has raised debate. In order to shed further light on this debate, the paper aims to illustrate evidence for one or two stresses in each of the compounds considered, to propose phonological transcriptions of the compounds based on intonational movement associated with each member, and to discuss the effects of relevant morphological, (morpho)syntactic, semantic, orthographic, and phonological characteristics available in previous literature through the lens of the experimental results provided here.

2. Conceptual background

This section provides an operational definition of compound as it will be used in this paper. It also covers the acoustic correlates of stress in Spanish, and reviews concepts relevant to Spanish intonational phonology and to stress in Spanish compounds.
2.1 Characteristics of compounds

One major distinction to make in the description of the properties of compounds is that of functional (hereafter, F) versus lexical (hereafter, L) categories. Moyna (2011: 23–4) formulates a hypothesis in which all syntactic heads can be defined through the features [±L] and [±F]. Anything classified as [+L], which is a requirement of members forming a compound, either possesses meaning itself (for example, traditional categories such as noun, verb, adjective or adverb) or is needed by another item in order to properly convey meaning (for example, categories such as word class marker, numeral, weak quantifier or modal verb). On the other hand, compounding is not carried out using [−L] items (for example, categories such as case and verbal inflection marker, determiner or strong quantifier). In terms of the end product of combining two [+L] items, Moyna (2011: 24) states that it should be a lexeme (that is, part of a main lexical class) as opposed to a function word such as a pronoun or a preposition.

From a syntactic point of view, Moyna (2011: 28–30) also states that the connection between members of compounds can typically be classified as apposition (for example, escritor–director ‘writer–director’; escritor ‘writer’ + director ‘director’), complementation (for example, lavaplatos ‘dishwasher’; lava ‘wash’ + platos ‘dishes’), coordination (for example, ajiaceite ‘garlic and olive oil sauce’; ajo ‘garlic’ + aceite ‘oil’) and modification (for example, hombre lobo ‘wolf man’; hombre ‘man’ + lobo ‘wolf’). These categories are applied within two general syntactic classes of compounds: HIERARCHICAL, containing a head and a modifier of that head (for example, hombre lobo ‘wolf man’) and CONCATENATIVE, with a flat representation of constituents (for example, escritor–director ‘writer–director’). In hierarchical cases, the head can come before or after its modifier, thus creating a further classification scheme of HEAD–INITIAL (for example, hombre lobo ‘wolf man’) versus HEAD–FINAL (for example, ciencia–ficción ‘science fiction’; ciencia ‘science’ + ficción ‘fiction’). Compounds can also undergo meaning adjustments over time as they become lexicalized (Lang 1990: 72–3), can be subject to recursion, and can be invented at any time in order to account for newly emerging terminology in a field or cultural and/or social phenomena (Moyna 2011: 30–4, Díaz Pérez 2014: 90). Another important classification scheme divides compounds into ENDOCENTRIC and EXOCENTRIC types, according to whether the syntactic and semantic characteristics of the compound come from the internal head (for example, cartón piedra ‘paper mache’; cartón ‘cardboard’ + piedra ‘stone’) or from a head outside of the internal members (for example, rabucocandil ‘long-tailed bee-eater’; rabuco/rabo ‘tail’ + candil ‘oil lamp’), respectively (Bustos de Gisbert 1986: 195–200, Moyna 2011: 54–61).2

Moyna (2011: 74–80) also discusses morphological, syntactic, and semantic means she uses to identify items that qualify as compounds. For example, in cases where one member of a compound is a bound stem with a distinct or absent final vowel that facilitates connecting the two members (for example, ajiaceite ‘garlic and olive oil sauce’), there is morphological support for compound status. Concerning such morphological properties, Moyna (2011) demonstrates that historically varying spellings do not impede our ability to confirm compoundhood (for example, músico
Terapia and musicoterapia are both attested forms of ‘music therapy’; músico ‘music’ + terapia ‘therapy’). Furthermore, in terms of syntax, Moyna (2011) devises three tests to distinguish compounds from phrases in cases where they superficially appear to have the same structure: inseparability, deletion under coordination, and fixity of constituent order. Inseparability disallows the insertion of, for example, the adverb muy ‘very’ between individual members of a compound such as pavo real ‘peacock’; pavo ‘turkey’ + real ‘royal’. If the members of this compound were to form a phrase, an intervening adverb would be grammatical. Deletion under coordination deems ungrammatical the omission of the second of two identical nouns in compounds, such as pez ‘fish’ in el pez luna y el (pez) espada ‘the ocean sunfish and the sword (fish)’. Finally, fixity refers to the inability to change the order of members of a compound. For example, balomano ‘handball’; balón ‘ball’ + mano ‘hand’ cannot be reversed to *manobalón. Next, from a diachronic point of view, Moyna (2011) claims that the meaning of a compound settles over time. That is, initially, a compound refers to a specific type of the concept to which its head more generally refers but, over time, its meaning narrows in scope. For example, malvivir ‘to live badly’; mal ‘bad’ + vivir ‘to live’ used to be associated with badness in general, but now it specifically deals with badness as a result of feeling indignant.

2.2 Relevant prosodic terminology and theoretical notions

Stress concerns the relative metrical prominence of lexical items, phrases, and utterances in discourse (Liberman & Prince 1977: 249–50, Ladd 1996: 42–6, Hualde 2006/2007: 60). In individual words, syllables that are lexically specified for stress at the phonological level, which in Spanish belong to content rather than function words (Quilis 1999 [1993]: 390–5), are realized with increased acoustic prominence compared to those that are not. One way of signaling such prominence is through accent, which in this context refers to fundamental frequency (hereafter, F0) movement located in or near lexically stressed syllables (Quilis & Fernández 1985: 153, Garrido et al. 1993: 574–5, Ladd 1996: 46–50, Quilis 1999 [1993]: 384–5, Face 2003: 117–18, Hualde 2006/2007: 60, 2009: 200–1). For this reason, Spanish is deemed a stress-accent language. While authors such as Face (2003: 121) and Rao (2009: 36) claim that accent is the most important indicator of stress, other recent experimental work has demonstrated that duration and intensity at times play a role in cuing stress, regardless of the occurrence or lack of accent (Ortega-Llebaria 2006: 104–6, Ortega-Llebaria & Prieto 2007: 155–7, Ortega-Llebaria & Prieto 2010: 73–7, among others).

Accent is particularly central to the AM approach to intonational phonology, in which individual or sequences of high (hereafter, H) and low (hereafter, L) F0 movements, or tones, associated with stressed syllables form pitch accents, which are surrounded by intervening phonetic interpolation (Hualde 2003). Above the syllable, AM proposes that F0 is associated with other metrically prominent phonological targets at the edges of the hierarchically organized constituents in (1).
IPs are generally demarcated by longer, easily perceptible pauses at their left and right boundaries. They are non-isomorphic with regard to syntactic structure and are considered to be meaning-bearing groups (Rao 2009: 34). In general, declarative utterances used in previous work on Spanish from a lab perspective are regarded as containing one IP whose right edge is cued by F0 suppression corresponding with an L% boundary tone, lengthening effects, and a clear disjuncture in speech (>400 milliseconds, based on Rao 2010). However, the termination of an IP can show variation. For example, CIRCUMFLEX movement, or a rise-fall configuration, is attested in the declaratives of some varieties (Sosa 1999: 190, Martín-Butragueño 2004: 349–56, 2006: 1–9, Prieto & Roseano 2010: 4–13, among others). Furthermore, ips are shorter phrases housed within IPs that do not always contain completed thoughts. Ip boundaries in Spanish, which are often transcribed as H– boundary tones, are mainly cued by F0 continuation rises to the final syllable of words, sustained pitch, longer duration of words, stressed syllables and vowels, drastic modifications to F0 range, decreased intensity, and short pauses (Elordieta et al. 2003: 487, 489, D’Imperio et al. 2005: 66–7, Prieto 2006: 41–2, Frota et al. 2007: 134–5, Rao 2007a: 31–6, 2007b: 82–95, 2010: 69–70). Below both phrase levels is the PW, which is a domain associated with main stress in content words. A lexical item achieves PW status if it displays one or more of the previously mentioned acoustic cues to stress through its lexically specified stressed syllable. In particular, the authors who claim that F0 is the most salient cue to stress in Spanish employ methods in which DEACCENTED (that is, failure to bear a pitch accent) words also lose their status as PWs (Rao 2009: 36–41, 63–7).

Previous work has revealed that ips usually contain between one and four PWs, with two being the ideal length (Prieto 2006: 44–5). Variation in PWs/ip has been attributed to factors such as satisfying overlap with syntactic boundaries, pragmatic functions, and speech rate. In order to illustrate the relevant levels of the hierarchy discussed to this point, this paper provides example (2), which contains a simple subject–verb–object declarative and a phrasing division separating the noun phrase (hereafter, NP) and verb phrase (hereafter, VP; stressed syllables are underlined).

(2) *Maria come la manzana*

‘Mary eats the apple’

[[MariaPW]ip [comePW la manzanaPW]ip]IP

The remaining commentary in this subsection elaborates on the major prosodic characteristics of Spanish broad focus declaratives, with an emphasis on differences between prenuclear (that is, non-final) and nuclear (that is, final) phrase or utterance
positions. With regard to prenuclear positions, PWs in phrase initial position generally manifest an ascent in F0 through their stressed syllable in the majority of dialects of Spanish. This results in the highest peak of the utterance or phrase, which is followed by gradual peak decay (that is, downstepping) across the utterance (Prieto et al. 1996: 445–7, Prieto 1998: 261–2). In prenuclear words in general, F0 valleys are typically anchored to the stressed syllable onset and are followed by a rise to a peak in the post-tonic syllable. This tendency is phonologically transcribed as an L+ > H* sequence, where > denotes post-tonic peak alignment and * references the tone more strongly associated with the stressed syllable.

Another prenuclear pattern identified in previous work is F0 staying low through the stressed syllable, beginning its rise at the offset, and peaking post-tonically (Face & Prieto 2007: 119–23). This sequence is labeled L* + H. Only recently have perceptual studies called for a declarative/interrogative distinction, based on where the rise originates, between L+ > H* and L* + H (Prieto & Roseano 2010: 2–3). Overall, post-tonic alignment characterizes situations in which there are enough intervening unstressed syllables to temporally facilitate F0 transitions between the excursions associated with stressed syllables. In cases in which a stressed syllable’s peak is closely followed by another stressed syllable, the second movement can be a high F0 plateau (that is, H*) or, if the relevant movement takes place in the lower region of the overall F0 range, it can evidence deaccenting. In sum, F0 has a very active role in and around stressed syllables of prenuclear PWs, especially in phrase-initial position.

Concerning nuclear position, Romance languages closely follow the Nuclear Stress Rule (Chomsky & Halle 1968: 16–18) in constituents of the prosodic hierarchy, meaning the rightmost element of a prosodic domain is the most prominent. However, the cue to such salience can vary by constituent. For example, in IP-nuclear position of declaratives, word salience is often achieved through final lengthening, with F0 being suppressed (that is, final lowering). While such low F0 levels correspond with deaccenting in prenuclear position, a phonological target transcribed as L* is attached to words in IP-nuclear position due to this position’s prominence (Prieto & Roseano 2010: 3). Furthermore, the fact that F0 continuation rises or plateaus cue ip boundaries shows that F0 is more active in nuclear position of the ip level, though often the most drastic excursions here are related to a boundary tone rather than the stressed syllable itself (D’Imperio et al. 2005: 79, Prieto 2006: 42).6

Finally, in nuclear phrase position in general, whenever F0 peaks are present, they are typically pushed leftward into the stressed syllable due to an upcoming phrase boundary, and thus do not show the delay in alignment seen in prenuclear position. This pattern is phonologically denoted as L+ H*. It is noteworthy that this tonal sequence does occur in prenuclear position in broad focus in some dialects (Colantoni & Gurlekian 2004: 109). It is also cited as a prenuclear manner of prosodically conveying narrow focus, as opposed to the more common means of syntactic movement (Face 2000: 45–7, 2001: 226–8, 2002: 71–6). Overall, based on the current comparison of prenuclear and nuclear positions, one would expect F0 to consistently exhibit more movement through stressed syllables in the former position rather than in the latter, especially in shorter utterances, in which multiple ips may not be present.
General schematics of frequently occurring pitch accents in Spanish, illustrating F0 movement in relation to stressed syllables, are shown in Figure 1 (modeled after Aguilar et al. 2009). The only pitch accent not previously detailed in this subsection is H+L*, which is most common in nuclear position of some interrogatives, but can occur in declaratives in some dialects based on pragmatic variation (Prieto & Roseano 2010: 4–9). When individual H and L tones in any of these pitch accents are associated with an abrupt F0 increase or decrease from their preceding identical tone in the same ip, upstep (¡) or downstep (!) notation can be implemented, respectively. An example of peak upstep is seen in Figure 1 for L+H*. Such upstep can break the expected downstepping in a declarative in order to convey narrow focus on a word, at times prior to a phrase boundary.

2.3 Previous non-acoustic observations of stress in Spanish compounds

According to Bustos de Gisbert (1986: 182–92), the stress properties of compounds are dictated by an interplay of factors such as agglutination, the internal relationship between members, their status as endocentric or exocentric, and their number of syllables. In relation to [NN]N compounds, he finds that almost 3/4 of the overall compound tokens in his corpus only bear one stress. In coordinate compounds, where the infix /-i-/ often takes the place of the left member’s gender marker, Bustos de Gisbert (1986: 184) suggests that there is a loss of independence in the left member, and, consequently, stress loss as well (for example, ajiaceite and not *ajiaceite). The
same tendency is attested in subordinate [NN]N compounds, or compounds that result from the loss of a preposition but do not undergo the level of fusion seen in the coordinate type (for example, **bocacalle** ‘intersection’; **boca** ‘mouth’ + **calle** ‘street’). By contrast, Bustos de Gisbert (1986: 189) cites a small percentage of examples in the subordinate category that seem to maintain two stresses, as well as other compounds in which the stress pattern is unclear (for example, **lápiz plomo** or **lápiz plomo** ‘lead pencil’; **lápiz** ‘pencil’ + **plomo** ‘lead’). Compounds with such atypical stress patterns can be semantically endocentric or exocentric. Bustos de Gisbert (1986: 190–1) also finds that attributive compounds, or those in which both members maintain their freestanding form and are not associated with the absence of an intervening preposition (but rather of an attributive verb, for example, **ser** ‘to be’), are the most challenging to classify as having one or two stresses. He mainly relies on reference to the number of unstressed syllables between the beginning of the word and the first stress in attempting to count stresses, and claims that three or more unstressed syllables prior to the first stress is undesirable. Assuming that right stress is typically present, compounds with longer (three or more syllables) left members and/or shorter (one or two syllables) left members and longer right members with paroxytonic or oxytonic stress may show increased evidence of two stresses in order to avoid having three or more unstressed syllables at the beginning (for example, **pájaro polilla** ‘kingfisher’; **pájaro** ‘bird’ + **polilla** ‘moth’). Finally, while the semantic division between endocentric and exocentric does not affect stress in the [NN]N compounds in Bustos de Gisbert’s (1986) corpus, he does find that endocentric compounds in adjective + noun compounds, especially those in which the adjective comes second (for example, **Noche vieja** ‘New Year’s Eve’; **noche** ‘night’ + **vieja** ‘old’), manifest two stresses at a much higher frequency than exocentric compounds of the same variety do.

Hualde (2006/2007: 66–76) discusses stress maintenance and deletion in Spanish compounds comprised of words belonging to various grammatical categories. In general, in stress-deleting compounds (hereafter, SDCs), the left member acoustically resembles a word that is lexically unstressed, such as a preposition. Left members are those that show stress deletion because Spanish favors a rightward erasure of prominence, that is, the tendency is to first remove prominence in domain-initial position. As such, Hualde (2006/2007) states that an SDC like **lavaplatos** ‘dishwasher’, in which stress is deleted on the verb **lava** ‘he/she washes’, is similar in stress pattern to a morphologically derived, non-compound like **lavadora** ‘washing machine’. According to Hualde (2006/2007), words with stress omission similar to the verb + noun **lavaplatos** can also be combinations of nouns (for example, **aguanieve** ‘sleet’; **aguas** ‘water’ + **nieve** ‘snow’), adjectives (in which the first member ends in /o/, for example, **anglosajón** ‘Anglo-Saxon’; **anglo** ‘Anglo’ + **sajón** ‘Saxon’), and nouns + adjectives (for example, **campsanto** ‘cemetry’; **campo** ‘field’ + **santo** ‘holy’).

In the cases of SDCs outlined by Hualde (2006/2007: 66–7), the loss of prominence is attributed to distinctions at the word and phrase level (following Liberman & Sproat 1992: 175–8). For example, stress deletion in **lava** ‘he/she washes’ within **lavaplatos** ‘dishwasher’ shows that both members of the compound are part of one and the same phonological, word-level domain ([lavaplatos]PW). On the other hand,
maintenance of both stresses when both words stand alone, such as in lava platos ‘he/she washes dishes’, suggests that the verb lava ‘he/she washes’ is a PW head of a phrase-level constituent structured as [[lava] PW [platos] PW] ip. Finally, when examining combinations of two nouns, an SDC, like puerco espín ‘porcupine’; puerco ‘pig’ + espín (from espina) ‘thorn’, only allows for plural formation at the end of the entire word domain. Thus, puerco espines ‘porcupines’ is grammatical, but *puercos espín and *puercos espines are not. Conversely, none of the evidence concerning stress deletion mentioned to this point holds true for a non-stress-deleting compound (hereafter, NSDC), such as hombre lobo ‘wolf man’, which can pluralize in more than one way and does maintain stress in both members as if they were freestanding.

Overall, Hualde’s (2006/2007: 64–76) observations with respect to NSDCs and SDCs provide a useful point of departure for studying the intonational phonology and stress properties of compounds. Joining his analysis with a couple of areas he does not address (that is, acoustic analyses of compound data, and the relationship between orthographic representations of compounds and their stress patterns) provides inspiration for follow-up work on compounds. Regarding orthography, one possibility mentioned in Moyna (2011: 81) is that compounds in which members are joined as one orthographic word demonstrate stress loss of the left member, while those spelled with a space between members (or perhaps a hyphen) maintain both stresses. Several [NN]N examples in Bustos de Gisbert (1986: 186–90) support this claim, albeit from non-acoustic approaches. However, Moyna (2011: 81–2) identifies two problematic aspects of this approach: many compounds have historically been represented as both one and two orthographic words, and non-compounds that are actually variants of syntactic phrases have been historically spelled as one single word.

2.4 Current agenda

Based on the topics discussed in previous sections, the remainder of this paper focuses on answering the following research questions related to the phonological representation of [NN]N compounds in Spanish:

i Do these compounds generally show evidence, through accent, of bearing one or two stresses?

ii If variation exists, what morphological, (morpho-)syntactic, semantic, orthographic, and/or phonetic/phonological concepts influence the number of stresses?

iii What are the phonological transcriptions of these compounds and what contributes to any variation seen in their phonological representations?

3. Methods

This section provides information on the participants in the study, the design of the data elicitation task, the way in which recordings were carried out and, finally, the techniques employed to analyze and code the compound data.
3.1 Participants

Eight native speakers of Spanish originally from Mexico, but currently residing in the Upper Midwest of the United States, were recruited to participate in the study. The group consists of four males and four females ranging between 30 and 55 years of age. They all came to the United States to attend graduate school, and have been in the country for at least five years, but continue visiting Mexico multiple times annually. Overall, even though they live in an English-dominant environment and are proficient speakers of English, they maintain consistent use of Spanish on a daily basis due to their academic affiliations and social circles.

3.2 Materials

In order to elicit production of \([NN]_N\) compounds demonstrating variation in a number of ways, the eight speakers carried out a reading task containing 30 such compounds taken from Bustos de Gisbert (1986: 187–90) and Hualde (2006/2007: 66–76). The majority of the selected compounds are also referenced in Moyna (2011: 331–56). The sample of 30 was selected carefully, through consultation with the aforementioned sources concerning variables that may affect compound stress patterns, in order to obtain variation with respect to factors such as endocentricity versus exocentricity, one versus two orthographic units, number of unstressed syllables prior to the stressed syllable of the right member, intervening unstressed syllables between the potentially stressed syllables of both members, and the attributive versus subordinate distinction in the syntactic-semantic categories of Bustos de Gisbert (1986: 187–90). The specific selection of attributive and subordinate compounds was motivated by Bustos de Gisbert’s (1986) general uncertainty about their stress patterns. Some examples of compounds in the elicitation task are listed in Table 1, along with their features of interest, which shed light on their similarities and differences. A list of all the compounds used can be found in the Appendix at the end of this paper.

Furthermore, since phonological accounts of initial and nuclear position differ, each of the 30 compounds was produced in two short minimal pair carrier phrases (60 total productions/participant). Thus, for each compound, the only difference between its two carrier phrases was that in one of them the compound appeared in initial position, while in the other it was in nuclear position. An example of this type of minimal pair is given in (3). In 56/60 of the carrier phrases, the expression \(\text{esta´ allı´} \{\text{definite noun}\} \text{ is there’}\) came before or after each compound. The compounds \(\text{ba\l\on}\text{mano} \{\text{handball’}\}; \text{bal\on} \{\text{ball’} + \text{mano ‘hand’}\} \) (henceforth, potential stresses are underlined) and \(\text{lengua madre} \{\text{mother tongue’}\}; \text{lengua ‘tongue’} + \text{madre ‘mother’}\) are not conducive to appearing with \(\text{esta´ allı´} \) and, therefore, \(\text{es divertido} \{\text{noun’] is fun’}\) and \(\text{es interesante} \{\text{noun] is interesting’}\) were substituted, respectively. All compounds were presented to participants in carrier phrases in their singular form and preceded by a definite article.
Table 1. Examples of compounds in the data elicitation task and their variation in terms of their syntactic-semantic relation between members (subordinate or attributive), semantic classification as endocentric or exocentric, orthographic unity or separation between members, number of syllables between left and right potential stresses, and the number of syllables from the beginning of the word to the right member’s stress.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Relation</th>
<th>Endo/Exo</th>
<th>Orthography</th>
<th>Syllables between stresses</th>
<th>Syllables to right stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>baño maría</td>
<td>subordinate</td>
<td>endo</td>
<td>separate</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>‘double boiler’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bauil mundo</td>
<td>subordinate</td>
<td>endo</td>
<td>separate</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>‘large trunk’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bulbocastaña</td>
<td>subordinate</td>
<td>exo</td>
<td>together</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>‘great pignut’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cartón piedra</td>
<td>attributive</td>
<td>endo</td>
<td>separate</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>‘paper mache’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lápiz plomo</td>
<td>subordinate</td>
<td>endo</td>
<td>separate</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>‘lead pencil’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nochehíelo</td>
<td>attributive</td>
<td>exo</td>
<td>together</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>‘stained cloth’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pájaro polilla</td>
<td>attributive</td>
<td>endo</td>
<td>separate</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>‘kingfisher’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pejemuller</td>
<td>attributive</td>
<td>exo</td>
<td>together</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>‘manatee’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(3) Minimal pair from the stimuli

a. Compound in initial position: El hombre lobo está allí (‘The wolf man is there’, /l/ of el resyllabifies to be realized as [lom]).

b. Compound in nuclear position: Está allí el hombre lobo (same meaning as (a), but literally, ‘Is there the wolf man’).

The participants were asked to read all the utterances in broad focus and at their natural speech rate, or as if they were speaking to another native speaker of Spanish. All 60 utterances were randomized when presented to participants. All the recording was done using Praat (Boersma & Weenink 2013), at a 44.100 Hertz (Hz) sampling frequency, a head-mounted microphone, and a laptop computer. Overall, the elicitation task yielded 480 [NN]N compound productions (30 compounds × 2 productions × 8 speakers = 480) and 960 total syllables of interest (480 compounds × 2 members/compound).

3.3 Data analysis

The 960 syllables of interest were first inspected acoustically for the occurrence of F0 excursions that might serve as evidence of a cue to stress. When such evidence was
detected, namely through F0 movement of at least 7 Hertz (Hz) (following O’Rourke 2005, 2006, Rao 2009) or, less commonly, through a flattened F0 in the upper regions of the pitch range, a pitch accent was expected to occur since these phonological targets associate with stressed syllables. As such, the second step of the analysis involved transcribing F0 movement at or near stressed syllables in compound members as pitch accents and, when relevant, boundary tones, using the AM conventions outlined in the previous section.

Based on the general intonational tendencies cited in Section 2, the attestation of pitch accents and their AM notation in members of compounds depended on alignment of F0 peaks and valleys with respect to stressed syllables’ offset and onset, respectively, as well as the degree of F0 excursion through stressed syllables, and the relative F0 level in stressed syllables (O’Rourke 2005, 2006). Alignment values relative to offsets and onsets were tabulated in milliseconds (ms). Regarding peak height, in order for a rise from a valley to be considered significant enough to culminate in a peak, it had to satisfy the threshold value of 7 Hz just mentioned. When this amount of movement was not present and F0 traces were in the lower region of the F0 range for an utterance, compound members were labeled either DEACCENTED (that is, there was evidence of stress across speakers, but no pitch accent in particular instances) or UNACCENTED (that is, there was no evidence of stress across speakers and no pitch accent). Due to nuclear prominence in the prosodic domains of Spanish, right members were generally expected to bear a pitch accent. Therefore, the F0 activity of left members of each compound is of particular interest because they are predicted to exhibit the most F0 variation, which, in turn, will convey whether each compound carries one or two stresses and, thus, one or two pitch accents.

Additionally, even though the stimuli were short sentences comprising one IP, Spanish declaratives often manifest ip boundaries between subjects and verbs, as seen in (2). Such a division in the data set would have ramifications for AM notation because utterance-medial ip boundary tones would then have to be incorporated. Also, if certain compounds with evidence of two stresses were to consistently be phrased separately from the other two PWs of elicitation utterances, two ips would occur, each with two PWs. According to previously proposed prosodic well-formedness conditions, this is a highly desirable structure (Prieto 2006, Rao 2007a). In order to determine whether ip boundaries occur and where they occur, each utterance was examined for the previously cited F0 correlates of ip junctures. These non-terminal breaks with rising F0 movement were transcribed as H-.

Figure 2, which is a male’s production of an utterance including the compound bienes raíces ‘real estate’; bienes ‘properties’ + raíces ‘roots’ in initial position, summarizes the steps of the acoustic analysis. Here, the F0 excursions through the prominent syllables of both bienes and raíces signal evidence of stress in both members. In bienes, the valley-to-peak rise begins at 100.6 Hz and culminates at 164.4 Hz, in the post-tonic syllable, indicating an L+ > H* pitch accent. The second valley, prior to the rise through the stressed syllable of raíces, is higher than its preceding one, and is therefore transcribed as ¡L to mark upstep. The second peak occurs within the stressed syllable
of raíces, at 127.1 Hz, and is followed by an additional rise to an H- boundary tone at 153.7 Hz. This boundary separates the compound from the remainder of the carrier sentence. These comments on raíces translate to an ¡L+H*H- sequence for the right member of this compound.

Finally, the data were tabulated according to the presence/absence of a pitch accent in right and left members, transcriptions of pitch accents and phrase boundary tones, as well as the factors exemplified in Table 1.

Before detailing the results, it is worth noting at this juncture that, while acoustic analyses were clearly important in determining the presence/absence and type of pitch accent in members of compounds, as well as the location of ip boundaries, the remainder of this paper focuses on the phonological implications of the acoustic analysis (similar to Prieto & Roseano 2010) rather than on an in-depth statistical analysis of F0 activity.

4. Results

This section is divided into two subsections because the outcome of the analysis of F0 yielded two clear groupings of [NN]N compound stress patterns. It should be noted that right members categorically contain evidence of stress and pitch accents, as can be expected from previous literature. Thus, all tables in this section focus on variation in left members based on utterance position. In our data, these left members are typically syntactic heads modified by the right members. The first subsection details Group 1 (hereafter, G1), which shows evidence of two stresses, and the second discusses Group 2 (hereafter, G2), which only shows evidence of right member stress. For each

Figure 2. The F0 contour corresponding with a male informant’s production of Los bienes raíces están allí ‘The real estate is there’.

<table>
<thead>
<tr>
<th>Syllables</th>
<th>los</th>
<th>bie</th>
<th>nes</th>
<th>ra</th>
<th>i</th>
<th>ces</th>
<th>ces</th>
<th>tá</th>
<th>na</th>
<th>lli</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tones</td>
<td>L-&gt;H*</td>
<td>¡L+H*</td>
<td>H-</td>
<td>L+H*</td>
<td>L<em>L</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

96 RAJIV RAO
group, F0 evidence of stress in left members in both initial and nuclear positions, and the factors that appear to influence such evidence or a lack thereof, are addressed first. This section then overviews characteristic pitch accents manifested in left and right members of compounds in each group, and examines factors that influence pitch accent variation.

4.1 G1 compounds

4.1.1 Comments on left member stress

This subsection details the 15 [NN]N compounds that form G1, which shows the highest frequencies of F0 excursions through the stressed syllables of individual members. When examining this group’s results, as outlined in Table 2, a number of general observations can be made with respect to the variables of interest for this paper. Regarding the positional distinction, 88% of left members show F0 evidence of stress and possess pitch accents in initial position, while 76% do so in nuclear position. The higher frequency in the former position makes sense given that F0 is typically more active toward the beginning of an utterance. However, 76% in nuclear position is noteworthy, given that final lowering often occurs in Spanish declaratives. In terms of phrasing patterns, 43% of compounds in initial position have an ip boundary at their right edge, that is, before the VP, meaning they are located within their own ip. In nuclear position, this tendency occurs in 32% of cases. Across G1, compounds occurring in their own ips are relatively balanced, that is, there is not one compound or set of compounds that is more conducive to having its own ip compared to others. Additionally, six of G1’s compounds are subordinate (baño maría ‘double boiler’, baúl mundo ‘large, deep trunk’, bienes raíces ‘real estate’, lápiz plomo ‘lead pencil’, palo Brasil ‘Brazilwood’, palo Campeche ‘bloodwood tree’) and nine are attributive (camión tanque ‘tanker truck’, cartón piedra ‘paper mache’, escalera caracol ‘spiral staircase’, hombre lobo ‘wolf man’, hormiga caballo ‘horse ant’, lengua madre ‘mother tongue’, pájaro mosquito ‘hummingbird’, pájaro polilla ‘kingfisher’, pez luna ‘ocean sunfish’). Therefore, there does not seem to be convincing evidence of two stresses being specifically characteristic of one of these two syntactic-semantic classifications.

On the other hand, a striking tendency emerges from the point of view of orthography: G1 compounds always appear as two separate units. Furthermore, in terms of intervening unstressed syllables between members’ individual stressed syllables, there is F0 evidence of stress in both members of compounds with clash (for example, baúl mundo ‘large, deep trunk’), one intervening syllable (for example, lengua madre ‘mother tongue’), two intervening syllables (for example, palo Brasil ‘Brazilwood’) or three intervening syllables (for example, pájaro polilla ‘kingfisher’). G1 compounds demonstrating the highest frequencies of evidence of left member stress have more syllables across the compound in general and in the left member in particular, as well as more intervening unstressed syllables between left and right stress (for example, escalera caracol ‘spiral staircase’ versus pez luna ‘ocean sunfish’).
Table 2. Frequencies of F0 evidence of left member stress in G1 compounds by utterance position. Syllables that are typically stressed when both members exist on their own are underlined.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Initial</th>
<th>Nuclear</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>bienes raíces</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>‘real estate’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pájaro mosquito</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>‘hummingbird’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>escalera caracol</td>
<td>100%</td>
<td>88%</td>
<td>94%</td>
</tr>
<tr>
<td>‘spiral staircase’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pájaro polilla</td>
<td>88%</td>
<td>100%</td>
<td>94%</td>
</tr>
<tr>
<td>‘kingfisher’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hormiga caballo</td>
<td>100%</td>
<td>75%</td>
<td>88%</td>
</tr>
<tr>
<td>‘horse ant’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>baño maría</td>
<td>75%</td>
<td>88%</td>
<td>81%</td>
</tr>
<tr>
<td>‘double boiler’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lápiz plomo</td>
<td>75%</td>
<td>88%</td>
<td>81%</td>
</tr>
<tr>
<td>‘lead pencil’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>palo Brasil</td>
<td>75%</td>
<td>88%</td>
<td>81%</td>
</tr>
<tr>
<td>‘Brazilwood’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>palo Campeche</td>
<td>75%</td>
<td>88%</td>
<td>81%</td>
</tr>
<tr>
<td>‘Bloodwood tree’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bail mundo</td>
<td>75%</td>
<td>75%</td>
<td>75%</td>
</tr>
<tr>
<td>‘large trunk’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>camión tanque</td>
<td>88%</td>
<td>63%</td>
<td>75%</td>
</tr>
<tr>
<td>‘tanker truck’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cartón piedra</td>
<td>88%</td>
<td>63%</td>
<td>75%</td>
</tr>
<tr>
<td>‘paper mache’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lengua madre</td>
<td>100%</td>
<td>50%</td>
<td>75%</td>
</tr>
<tr>
<td>‘mother tongue’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hombre lobo</td>
<td>88%</td>
<td>50%</td>
<td>69%</td>
</tr>
<tr>
<td>‘wolf man’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pez luna</td>
<td>88%</td>
<td>38%</td>
<td>63%</td>
</tr>
<tr>
<td>‘ocean sunfish’</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All the above makes sense because it relates to speakers having more time available to produce an F0 excursion. The exception to the comment on longer left members occurs in bienes raíces ‘real estate’, whose left member only has two syllables. Also, the last four rows of Table 2 show that shorter left members with one or no intervening syllables relatively reduce frequencies in nuclear position. Furthermore, there is no evidence of a clear pattern with respect to Bustos de Gisbert’s (1986) concern about having three or more unstressed syllables before the right member’s stress. One would anticipate that all of the compounds with longer left members and/or oxytonic right members would be prevalent in this table in order to satisfy this condition. By contrast,
the general results in Table 2 reveal that anywhere from two to six syllables can occur before the right members’ stress (for example, baúl mundo ‘large, deep trunk’, escalera caracol ‘spiral staircase’, respectively). However, a top to bottom glance at the table shows that having three or more unstressed syllables before the right stress seems to increase left stress frequencies, especially in nuclear position.

4.1.2 Pitch accent patterns

This section presents a phonological analysis of G1’s left and right members based on an overview of characteristic pitch accents and any relevant phrase boundary tones. Each member is described according to utterance position, since pitch accents typically vary in prenuclear and nuclear positions of declaratives in Spanish, as well as the most frequently occurring tonal sequence across speakers.

From a cursory glance at Table 3, it is quite clear that G1’s left members in initial position are mainly transcribed as L+>H*, that is, they are realized with an F0 valley at the stressed syllable onset, followed by a rise through the stressed syllable that reaches its peak post-tonically. However, a closer look reveals that this pitch accent only characterizes examples in which there is at least one unstressed syllable between left and right stresses. In such cases, F0 has enough space to displace its peak to a post-tonic syllable before beginning its excursion associated with the subsequent stressed syllable. On the other hand, in the four compounds with stress clash (baúl mundo ‘large, deep trunk’, camión tanque ‘tanker truck’, cartón piedra ‘paper mache’, pez luna ‘ocean sunfish’), the F0 excursion associated with the right member’s stressed syllable blocks the potential displacement of the left member’s peak and, as such, the left member has an earlier aligned peak, within its stressed syllable. These movements are phonologically classified as an L+H* pitch accent.

Furthermore, G1’s right members in initial position are produced with two tonal patterns. One is ¡L+H*H-, with optional valley upstep from the left member, and an H- ip boundary tone, which forces its preceding F0 peak to align within the stressed syllable and divides the compound from the following VP. The other is L+>H*, without the ip boundary tone, which allows peak displacement. The latter case occurs when entire utterances are articulated as just one ip. In nuclear position, left members yield variability that seems to be tied to their total number of syllables and the occurrence of stress clash or a lack thereof. The words with three syllables, as well as the exceptional, bisyllabic lápiz ‘pencil’, favor a bitonal pitch accent with peak displacement, L+>H*, just as in initial position.

Another tendency is an F0 plateau associated with a monotonal, H* pitch accent. This is typically a lower plateau than the F0 high of the previous word, and thus carries downstep notation (!) in its transcription. These plateaus occur mainly in bisyllabic words with initial stress, possibly because of the crowding of stresses created by the preceding oxytone allí ‘there’ of the majority of carrier phrases, and the following right member, which temporally prevent movement down to a valley and back up to a peak. In cases of stress clash, two tendencies are observed: the same pitch accent as in initial position, L+H*, and deaccenting, the latter of which also
occurs in lengua madre ‘mother tongue’, where an intervening unstressed syllable occurs between its left and right member. Nuclear left member deaccenting is particularly noteworthy because it is exclusive to this position. Finally, F0 in right members in nuclear position yields a clear final lowering tendency that is phonologically demarcated as L*L%.

Figures 3 and 4 show F0 contours illustrating examples of the compound stress and tonal tendencies cited in this section. Figure 3 complements Figure 2 from Section 3 by showing a production of bienes raíces ‘real estate’ in nuclear utterance position that

<table>
<thead>
<tr>
<th>Compound</th>
<th>Position</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>bienes raíces</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>(j)L + H*H-</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + &gt;H*</td>
<td>L*%</td>
</tr>
<tr>
<td>pájaro mosquito</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>L + H*(H-)</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + &gt;H*</td>
<td>L + H*%L%, L*L%</td>
</tr>
<tr>
<td>escalera caracol</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>(j)L + H*(H-)</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + &gt;H*</td>
<td>L*L%</td>
</tr>
<tr>
<td>pájaro polilla</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>L + &gt;H*</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + &gt;H*</td>
<td>L*%</td>
</tr>
<tr>
<td>hormiga caballo</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>L + H<em>H-, L + &gt;H</em></td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + &gt;H*</td>
<td>L*L%</td>
</tr>
<tr>
<td>baño maría</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>L + &gt;H*</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + &gt;H*</td>
<td>L*%</td>
</tr>
<tr>
<td>lápiz plomo</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>(j)L + H*(H-), (j)L* + H(H-)</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + &gt;H*</td>
<td>L*%</td>
</tr>
<tr>
<td>palo Brasil</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>L + H*(H-)</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + &gt;H*</td>
<td>L*%</td>
</tr>
<tr>
<td>palo Campeche</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>(j)L + &gt;H*</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + &gt;H*</td>
<td>L*%</td>
</tr>
<tr>
<td>baúl mundo</td>
<td>initial</td>
<td>L + H*</td>
<td>¡L + H*, H*(H-)</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + H*</td>
<td>L*%</td>
</tr>
<tr>
<td>camión tanque</td>
<td>initial</td>
<td>L + H*</td>
<td>(j)H*</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + H*</td>
<td>L*%</td>
</tr>
<tr>
<td>cartón piedra</td>
<td>initial</td>
<td>L + H*</td>
<td>H*(H-)</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + H*, deaccented</td>
<td>L*L%</td>
</tr>
<tr>
<td>lengua madre</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>(j)L + H*H-</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>deaccented</td>
<td>L*%</td>
</tr>
<tr>
<td>hombre lobo</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>(j)L* + H</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>deaccented</td>
<td>L*L%</td>
</tr>
<tr>
<td>pez luna</td>
<td>initial</td>
<td>L + H*</td>
<td>H*(H-)</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>deaccented</td>
<td>L*L%</td>
</tr>
</tbody>
</table>
Figure 3. A male informant’s production of the utterance *Esta´n allı´ los bienes raı´ces.* The F0 contour shows two ips and pitch accents on both members of the G1 compound in nuclear position of the utterance. As expected, F0 levels decay across the utterance.

Figure 4. A female informant’s production of *Esta´ allı´ el hombre lobo* ‘The wolf man is there’, where the G1 compound occurs in nuclear position of this utterance that is manifested as one ip/IP. This speaker combines the two /a/ segments of *está* and *allí* to produce an elongated [a:] and, as expected in Spanish, resyllabifies the sequence *el hombre.* The F0 suppression through *hombre* is evidence of deaccenting of this left member.
appears in its own phrase, while Figure 4 depicts a realization of *hombre lobo* ‘wolf man’ with deaccenting of its left member in nuclear utterance position.

4.2 G2 compounds

4.2.1 Comments on left member stress

Table 4. Frequencies of F0 evidence of left member stress in G2 compounds by utterance position. Syllables that are typically stressed when both members exist on their own are underlined.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Initial</th>
<th>Nuclear</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>maestrepasquín</td>
<td>50%</td>
<td>25%</td>
<td>38%</td>
</tr>
<tr>
<td>‘anti-authority expression’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pejemuller</td>
<td>38%</td>
<td>38%</td>
<td>38%</td>
</tr>
<tr>
<td>‘manatee’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cepacaballo</td>
<td>14%</td>
<td>38%</td>
<td>26%</td>
</tr>
<tr>
<td>‘spiny cocklebur’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bulbocastania</td>
<td>38%</td>
<td>13%</td>
<td>25%</td>
</tr>
<tr>
<td>‘great pignut’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nochehilo</td>
<td>38%</td>
<td>13%</td>
<td>25%</td>
</tr>
<tr>
<td>‘black-stained cloth’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>varagavilán</td>
<td>25%</td>
<td>25%</td>
<td>25%</td>
</tr>
<tr>
<td>‘scratch plough’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>brocamantón</td>
<td>13%</td>
<td>25%</td>
<td>19%</td>
</tr>
<tr>
<td>‘crotchet of diamonds’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rabucocardil</td>
<td>25%</td>
<td>13%</td>
<td>19%</td>
</tr>
<tr>
<td>‘long-tailed bee-eater’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bocatijera</td>
<td>25%</td>
<td>13%</td>
<td>19%</td>
</tr>
<tr>
<td>‘pole socket’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>estrellamar</td>
<td>13%</td>
<td>13%</td>
<td>13%</td>
</tr>
<tr>
<td>‘starfish’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>balonmano</td>
<td>13%</td>
<td>13%</td>
<td>13%</td>
</tr>
<tr>
<td>‘handball’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rataconejo</td>
<td>13%</td>
<td>0%</td>
<td>6%</td>
</tr>
<tr>
<td>‘coney rat’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aguanieve</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>‘sleet’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>reinaluisa</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>‘lemon verbena’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bocamanga</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>‘cuff’</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The 15 G2 compounds in Table 4 display F0 movement in left members’ freestanding stressed syllables at rates of just 0–38%, which are substantially lower than those of G1.
The frequency of left stress is slightly higher in initial position (20%) than in nuclear position (15%). The overall phrasing patterns are similar in G1 and G2: in initial G2 compounds, 39% are housed in their own ip, as are 34% of nuclear compounds in this group. In the infrequent cases in which left members evidence stress, 31% of such compounds were contained in their own ip. In terms of syntactic-semantic category, Table 4 displays six compounds classified as attributive (cepacaballo ‘spiny cocklebur’, nochehíelo ‘cloth stained with black’, pejemuller ‘manatee’, rataconejó ‘coney rat’, reinaluisa ‘lemon verbena’, varagavílan ‘ard/scratch plough’; the syllables which are not in small capitals do not show evidence of stress in compound form) and nine labeled as subordinate (aguaniéce ‘sleet’, balonmano ‘handball’, bocamanga ‘cuff’, bocatjera ‘pole socket on a carriage’, brocamantón ‘crotchet of diamonds’, bulbocastaña ‘great pignut’, estrellamar ‘starfish’, maestrepasquin ‘written expression against authority placed in public’, rabucocandil ‘long-tailed bee-eater’). The distribution between the categories indicates that neither of these two categories is particularly characteristic of G2 compounds. Recall that, similarly, the G1 compounds from Table 2 did not seem to use the attributive/subordinate divide as an indicator of higher levels of left member F0 activity. It is also unclear whether G2 compounds are connected more with endocentricity or exocentricity because both types are present in Table 4: four are exocentric (bulbocastaña ‘great pignut’, nochehíelo ‘cloth stained with black’, rabucocandil ‘long-tailed bee-eater’, reinaluisa ‘lemon verbena’), while the others are endocentric. These results further highlight the lack of a strong distinction between these two categories shown in G1 compounds. Regarding orthography, Table 4 shows that G2 compounds are categorically spelled as one word. This tendency is the exact opposite of what was observed for G1 in Table 2. Therefore, thus far, the distinction between one versus two orthographic units appears to show the strongest relation with left stress or a lack thereof in the current set of compounds. With respect to syllable-related issues in G2, the number of intervening unstressed syllables between left and right individual stresses is not a distinguishing feature of G2 compounds. Table 4 lists compounds with a variety of intervening unstressed syllables, ranging from zero (for example, balloon mano ‘handball’) to three (for example, varagavílan ‘ard/scratch plough’), with one (for example, bocamanga ‘cuff’) and two (for example, brocamantón ‘crotchet of diamonds’) being the most common. The comparison between the top nine compounds and the bottom six suggests a general pattern of two or more intervening syllables relating to higher frequencies of F0 evidence of left stress. Exceptional cases within the top and bottom levels are nochehíelo ‘cloth stained with black’ and rataconejó ‘coney rat’, respectively. Additionally, the condition concerning three or more unstressed syllables before the right member’s stress necessitating left member stress is rejected by both high and low ranking items within Table 4 (for example, bulbocastaña ‘great pignut’, brocamantón ‘crotchet of diamonds’, maestrepasquin ‘written expression against authority placed in public’, rabucocandil ‘long-tailed bee-eater’ and rataconejó ‘coney rat’. This suggests a lack of a G2 intragroup pattern for this variable as well.
4.2.2 Pitch accent patterns

Table 5. Characteristic pitch accents in left and right members of each G2 compound in initial and nuclear utterance positions. Right member pitch accents in initial position followed by an H- boundary tone indicate a compound in its own ip. Tones or upstep/downstep notation in parentheses accompanying pitch accents only occurred in some productions. M% (that is, mid) IP boundaries have a final sustained pitch rather than a rise (H%) or a fall (L%).

<table>
<thead>
<tr>
<th>Compound</th>
<th>Position</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>maestrepasquín</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>L + H*(H-)</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + &gt;H*</td>
<td>L + H*(L%/M%)</td>
</tr>
<tr>
<td>pejemuller</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>L + H*(H-)</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>(!)H*</td>
<td>L*L%</td>
</tr>
<tr>
<td>cepacaballo</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>L + &gt;H*</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>!H*</td>
<td>L*L%</td>
</tr>
<tr>
<td>bulbocastaña</td>
<td>initial</td>
<td>L + H*</td>
<td>L + &gt;H*</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + H*</td>
<td>L + H*L%</td>
</tr>
<tr>
<td>nochechico</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>L + &gt;H*</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + &gt;H*</td>
<td>L*L%</td>
</tr>
<tr>
<td>varagavilán</td>
<td>initial</td>
<td>L + &gt;H*, L + H*</td>
<td>L + H*(H-)</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + &gt;H*</td>
<td>L*(L%/M%/LM%)</td>
</tr>
<tr>
<td>brocamantón</td>
<td>initial</td>
<td>L* + H</td>
<td>L + H*(H-)</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + &gt;H*</td>
<td>L + H*(L%/M%/H%)</td>
</tr>
<tr>
<td>rabucocandil</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>L + H*(H-)</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + &gt;H*</td>
<td>L*L%</td>
</tr>
<tr>
<td>bocatijera</td>
<td>initial</td>
<td>L + H*</td>
<td>L + H*(H-)</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>!H*</td>
<td>L*L%</td>
</tr>
<tr>
<td>estrellamar</td>
<td>initial</td>
<td>H*</td>
<td>L + H*(H-)</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>!H*</td>
<td>L*L%</td>
</tr>
<tr>
<td>balonmano</td>
<td>initial</td>
<td>L + H*</td>
<td>L + &gt;H*</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>L + H*</td>
<td>L*L%</td>
</tr>
<tr>
<td>rataconejo</td>
<td>initial</td>
<td>L + &gt;H*</td>
<td>L + &gt;H*</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>Unaccented</td>
<td>L*L%</td>
</tr>
<tr>
<td>aguamieve</td>
<td>initial</td>
<td>Unaccented</td>
<td>L + &gt;H*</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>Unaccented</td>
<td>L<em>L%, L + H</em>L%</td>
</tr>
<tr>
<td>reinaluisa</td>
<td>initial</td>
<td>Unaccented</td>
<td>L + &gt;H*</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>Unaccented</td>
<td>L + H*(L%/M%)</td>
</tr>
<tr>
<td>bocamanga</td>
<td>initial</td>
<td>Unaccented</td>
<td>L + &gt;H*</td>
</tr>
<tr>
<td></td>
<td>nuclear</td>
<td>Unaccented</td>
<td>L<em>L%, L + H</em>L%</td>
</tr>
</tbody>
</table>

All the G2 compounds in Table 4 had relatively low frequencies of F0 evidence of stress in left members, so their left members should not generally be expected to serve as anchoring points for pitch accents. However, Table 4 shows that pitch accents can be attested in cases that are above 0%. Table 5 illustrates the phonological representations characterizing G2’s left and right members according to position. This table shows that
five different types can be noticed when G2’s left members bear a pitch accent, which demonstrates increased variation compared to G1’s left members. In G2, in initial position, L+H* occurs when unstressed syllables follow, and L+H occurs in the one clash situation (balon mano ‘handball’); however, the generally less common H* and L*+H are also found, as well as L+H* in a non-clash context, even though this early aligned pitch accent is not expected in prenuclear broad focus conditions. In nuclear position, G2 left member pitch accents exhibit less variation than in initial position. The top two pitch accents observed are L+H* and the downstepped plateau !H*. This distinction may be due to speech rate differences. In nuclear position, since the PW before the left member of all but one compound is the oxytone allí ‘there’, a faster speech rate would not allow an F0 drop to a valley prior to the realization of the left member’s pitch accent, resulting in a drop to a flatter F0 that is still not in the lowest region of the F0 range, rather than a drop to a valley and then a rise to a peak, which requires more time. Furthermore, L+H* is found only in nuclear position in the one left member in a clash context: balon mano ‘handball’.

In terms of G2’s right members, in general, pitch accent frequencies favor L+H* in initial position. In such cases, there is typically only one ip/IP in the utterance and compound peaks displace to the carrier phrase’s verb. However, when the right member is an oxytone (for example, brocamantón ‘crotchet of diamonds’, estrellamar ‘starfish’, rabucandil ‘long-tailed bee-eater’, zaragavilán ‘ard/scratch plough’), the peak tends to align within the stressed syllable, signaling an L+H* pitch accent, which is optionally followed by an H- boundary tone. The exception is bocati ‘pole socket on a carriage’, whose right member is a paroxytone, but still favors L+H*.

Additionally, Table 5 also illustrates that nuclear final lowering, corresponding with the sequence L*L%, in addition to the circumflex configuration L+H*L%, are the two most common terminal sequences. There does not seem to be a clear G2 right member characteristic that distinguishes the preference for one of these tonal sequences or the other. Finally, Table 5 also shows that four G2 examples have right members whose productions are followed by different types of IP boundaries. Such examples were not present for G1 in Table 3, where L% was the only IP boundary attested.

Figures 5, 6, and 7 provide F0 contours that exemplify some of the patterns discussed in Section 4.2. The first two of these figures illustrate F0 contours corresponding with productions of bocaManga ‘cuff’; in Figure 5, this compound occurs in initial utterance position, and in Figure 6 it is in nuclear position. The contour in Figure 7, containing the individually phrased compound bulbocastaña ‘great pignut’, depicts the less common case of a pitch accent associating with a left member.

5. Discussion and conclusions

5.1 Summary of findings

This paper’s analysis of 30 [NN]N compounds produced by eight speakers of Mexican Spanish was intended to identify which exhibit F0-based evidence of left member
stress, to identify factors that contribute to the occurrence or absence of left member stress, and to provide phonological characterizations of the compounds. The acoustic analysis of left and right members of the compounds produced yielded results that are most strongly divided according to orthographic convention: when compounds were written as one orthographic entity, as in G2, the left member typically failed to demonstrate evidence of stress, whereas when they were written as two separate items, as in G1, both members displayed F0 movement evidencing two stresses. These two tendencies were most frequently attested when compounds were in initial position in carrier phrases. Furthermore, G1 also demonstrated evidence of an increased distance between left and right member stress generally increasing the frequency of left member stress. However, the data did not support Bustos de Gisbert’s (1986) argument on the undesirability of three or more stressed syllables occurring at the beginning of a compound. Similarly, the endocentric/exocentric and attributive/subordinate oppositions did not appear to distinguish G1 from G2 compounds.

Regarding phrasing, compounds belonged to their own ip in just over one-third (36%) of all productions across the data, regardless of group and position. Finally, left member pitch accents, when present, favored L+>H* when post-tonic unstressed syllables were available to receive peak displacement, and gravitated toward L+H* in cases of tonal crowding. For right members in initial position, selection of L+>H*, L+H* or H* depended on whether or not there was an upcoming ip boundary, and/or on the number of pre-tonic or post-tonic unstressed syllables. Right members in nuclear position overwhelmingly favored L*L%, or the final lowering tendency typical
Figure 6. A male informant’s production of *Está allí la bocamanga* ‘The cuff is there’ as one ip. The G2 compound, in nuclear position here, exhibits final lowering. This relative F0 low is evidence against left member stress.

Figure 7. A male informant’s production of *El bulborcastaña está allí* ‘The great pignut is there’, with the G2 compound in initial position of the first of two ips. The F0 excursion in the left member’s syllable /bul/ is evidence of secondary stress with a corresponding L+H* pitch accent that normally signals narrow focus in prenuclear position.
of terminal junctures of Spanish declaratives. Overall, the tonal characterizations of G2 demonstrated greater variation than those of G1.

5.2 Implications

At this point, the implications of the aforementioned major findings will be discussed in terms of expansions of previous research. Throughout the remainder of this section, it is important to bear in mind that this is one of the first studies of its kind for Spanish and, as such, further research is needed in order to elaborate on the current commentary concerning stress and the phonology of compounds.

5.2.1 Orthography

It is tempting to conclude that orthography appears to have the strongest influence on compound stress in the current data; however, linguistic analysis principally addresses speech and not writing. From this perspective, the findings presented here imply that the speakers’ mental representations of the compounds they saw (or recognized) in the elicitation task dictated their prosodic interpretations of the compounds. This implication arises from an acoustic analysis and complements Moyna’s (2011: 31) remark: ‘Access to prosody and sounds in general is mediated by orthography and thus, by the criteria of the scribe and/or lexicographer’. This means that words orthographically represented as one or two words are the result of previous mental prosodic interpretations of compounds by those producing written documents. The same argument can be implemented with the current data, meaning orthographic separation or union of compound members did not drive the speakers’ productions. Rather, these orthographic forms that ended up corresponding with one or two stresses were non-speech-based representations shared by the mental prosodic representation of all G1 and G2 compounds, respectively. As such, one could propose that the current findings reinforce the prosodic interpretation of the scribes/lexicographers who used the written forms used in our elicitation task.

Conversely, it can be argued that orthography plays a role in conditioning the realization of compound stress, and a task similar to the one used here could be created, using exclusively compounds with alternate orthographic representations (for example, hombre lobo versus hombr el obo ‘wolf man’), in order to test this proposal. Along these lines, Juhasz et al. (2005), concluded that a space between the two members of English compounds helped participants gain access to each member separately. They also found that joining two members into one orthographic unit resulted in processing the compound with its full meaning. The implications of these findings would suggest that when our speakers read compounds with two orthographic units (especially those they did not recognize), they processed and produced each one with its individual phonological characteristics (that is, as two PWs). Conversely, when they saw G2 items, their processing merged the individual members into one unit, which led to a phonological output of just one PW with a nuclear prosodic head. Overall, while we tentatively assume the first rationale presented in relation to orthography due to the
nature of linguistic inquiry in general, there is plenty of room for further exploration of perception and production tasks that could lend further support for the two perspectives just overviewed.

5.2.2 Syntax and semantics

The fact that the majority of our compounds are syntactically head-initial, but only final, modifying members (that is, prosodic heads at the word or phrase level) consistently exhibited evidence of prominence implies that location at higher syntactic nodes does not clearly influence compound stress patterns. The minimal effect of syntactic as well as semantic variables was further highlighted through our inability to distinguish G1 and G2 compounds via reference to the attributive versus subordinate or the endocentric versus exocentric comparisons. According to Bustos de Gisbert (1986), one would have expected these classification schemes to play a more important role in compound stress.

5.2.3 Prosodic classifications

The results also have implications for Hualde’s (2006/2007) categorizations of stress deletion or retention in Spanish compounds. The current examination of F0 movement reveals that its two groupings can be translated into Hualde’s terminology: G1 compounds can be classified as NSDCs and G2 compounds as SDCs. Therefore, G1 compounds’ left and right members were produced as two PWs that only share prosodic domains at the ip level and above. On the other hand, G2 compounds’ members were articulated as one PW, meaning they are unified at the lower, word level, similar to an unstressed pronoun or preposition and its following stressed word. This distinction affects the autonomy of left and right members and helps explain why, for example, G1 compounds can be pluralized in multiple ways, while those of G2 only accept the realization of the plural morpheme at the end of the compound.

5.2.4 Suprasegmental phonology

When pitch accents occurred in compounds in this study, they generally reflected previously noted Spanish declarative patterns, that is, the peak displaced to a post-tonic syllable (L+H*) unless adjacent stresses or a following phrase boundary tone were attested, in which case peaks aligned with tonic syllables (L+H*). Also, in prenuclear position, ip boundaries were all of the H− variety, signaling the continuation of a thought, and in nuclear position of the IP as a whole, F0 suppression, represented as L*L%, was most common. Overall, the two members of compounds in G1 behave as if they were individual content words. Therefore, our data demonstrate that, while compounding the individual members of G1 creates semantic and syntactic adjustments from freestanding forms, their individual phonological targets are generally maintained. While these tonal tendencies were similar on the whole, declaratives with G1 compounds typically contain one more significant F0
excursion, and thus, pitch accent and PW, compared to those with G2 compounds. However, the fact that both groups had compounds with three or more syllables before the right member’s stress weakens Bustos de Gisbert’s (1986) hypothesis that such a threshold may force longer compounds to carry two stresses. This hypothesis was not corroborated, but another issue related to stress spacing merits further commentary. Regarding intragroup frequency comparisons in G1, fewer intervening unstressed syllables, in particular, stress clash situations between members, reduce left member stress frequencies, especially in utterance nuclear position. Stress clash is not favored in Spanish, but it may occur, especially in particular speech styles (see Hualde 2010: 11–13). By contrast, in Catalan, a closely related language, resolution in such contexts has been found to take place by deleting the first stress and optionally prosodically restructuring the items in question into one minor phrase (Prieto 2005: 217–20, 2011: 1928–31). A similar claim (at the PW, rather than the phrase level) may shed light on G1’s data with clash: in such environments, speakers may optionally delete left member stress and restructure the two compound members into one PW. On the other hand, G1’s decreased frequencies of left member F0 movement specific to nuclear position may also have been because duration tends to have more of a role in this position or because moving a grammatical subject to nuclear position in a Spanish declarative is a syntactic means of focalization, which may somewhat detract from the role of prosody. Left members of G1 with more deaccenting are also relatively shorter in terms of number of syllables, which echoes one of Rao’s (2009) findings for spontaneous speech. The exception to this case is bienes raíces ‘real estate’, which is categorically realized with evidence of two stresses, even though its left member only has two syllables. This compound’s evidence of two stresses can be explained by morphological means: It is a unique, marked form in this data set for its left and right pluralization, which is interpreted here as increased autonomy for the left member and, thus, the ability to bear its own stress. In sum, the discussion above outlines some explanations of intragroup frequency differences in G1’s left members, but these issues merit future research.

Additionally, in the transition between pitch accents of left and right members of G1, a noteworthy, optional tendency that has rarely been cited for Spanish declaratives appears: valley upstep (¡L) in the right member, particularly when followed by an ip boundary. The H- boundary in right members of compounds, which has been cited as perceptually highlighting an element immediately preceding it (Face 2002: 75), may be accompanied by valley upstep in an attempt to distinguish the ip head from its non-head. This could have implications for conditions on ip length (that is, two PWs/ip is ideal, but one to four is an acceptable range at most speech rates) and balance in length of adjacent ips (Prieto 2006, Rao 2007a) with respect to the additional PW in G1. For example, if G1 compounds had drastically higher rates of being contained within their own ip, it could be argued that they do so to satisfy the aforementioned prosodic well-formedness criteria. However, the similar frequencies of compounds in individual ips in both G1 and G2 do not allow for such claims, nor does the brevity of our carrier phrases (at maximum, four PWs). As such, embedding compounds within syntactically more complex utterances would be a fruitful object of inquiry.
If G1 compounds are claimed to be NSDCs and G2 to be SDCs, it is also necessary to account for how pitch accents are manifested at times in the latter group. This appears to be counterintuitive given that, by definition, pitch accents are targets associated with a stressed syllable. To account for the results of left members in Table 5, this paper resorts to Hualde’s (2006/2007, 2010) reference to postlexical secondary stress, where a syllable to the left of the one carrying main stress also bears prominence. This must be specified as postlexical in order to indicate that it occurs at a higher, phrasal level, after lexical rules have been applied at the PW level. In the present data, these G2 left member pitch accents appear in several varieties, some of which are relatively uncommon in the literature on Spanish prenuclear broad focus declaratives (for example, $L+H^\ast$). This implies that compound secondary stress has a distinct nature in terms of tonal targets associated with left members. Hualde (2010) finds that this phenomenon is possible even directly to the left of main stress. This covers why a case such as *balón mano* ‘handball’ was realized with secondary stress (albeit minimally), just as most other left members of G2 were, at varying frequencies. Overall, the limited set of examples proposed as bearing secondary stress in our compounds does not support or reject this idea.

5.3 Limitations and future directions

We now turn to specific characteristics of our stimuli that raise further questions about stress in Spanish compounds. First, *maestrepasquín* ‘written expression against authority placed in public’ is the only clear case of a syntactically head-final compound in our data set. In this compound, the syntactic and prosodic heads overlap. The fact that this particular compound demonstrates the highest frequency of F0 evidence of stress in G2 hints that the comparison of compound stress in head-initial and head-final compounds is a relevant variable to pursue. Second, alternate written forms can change syntactic head-ordering, as well as the number of orthographic units. For example, the corpora cited in Bustos de Gisbert (1986) and Moyna (2011), as well as simple Google searches, demonstrate that *hombre lobo* ‘wolf man’ can be written with the left and right members reversed (*lobo hombre*), as one or two words in either order (*hombrelobo, lobombre*), and even, in some cases, with a hyphen between members (*hombre-lobo, lobo-hombre*). Along similar lines, *pez mujer* also appears as two words (*pez mujer*). Additionally, it is important to note that, in Spanish compounds, orthographic accent may affect a speaker’s interpretation of stress placement. For example, when the compound *balón mano* ‘handball’ is written in this form, the left member does not have an orthographic accent, but it does (*baloñ*) when it is freestanding. Therefore, if a speaker produces this word as the left member of a compound and sees that there is no written accent, he/she may be influenced to reduce its normally stressed syllable’s prominence. Additionally, in related work, instances of mismatch between the gender of individual members and that of the compound as a whole could be controlled. For example, in *aguas nieve* ‘sleet’, both *agua* ‘water’ and *nieve* ‘snow’ are feminine nouns, but the compound as one entity is masculine. The masculine definite article *el* in front of this compound may have influenced the
informants’ conceptualization of the compound as one prosodic word. Another concept to consider that was not incorporated into this paper is stress patterns in compounds with allomorphs with diphthongs (for example, \textit{cuentacuentos} ‘storyteller’), which, due to the evolution of the Spanish vocalic system, typically signal a stressed syllable.

Finally, a set of more general methodological issues related to this study can motivate further research on the phonology of Spanish compounds. The data collection procedure for this study involved a reading task in an experimental setting, which, by definition, involves orthographic conventions. However, other considerations, outside of or in conjunction with orthography (for example, comparisons with perception and/or picture description tasks), may help explain the F0-based differences between G1 and G2. Also, the stimuli used here were prepared mainly with respect to the corpora and analytical comments of Bustos de Gisbert (1986), Hualde (2006/2007), and Moyna (2011), but there may be other explanations outside of our variables of interest that account for the patterns observed here. Possibilities for such further exploration could come from the rich body of previous work on compounding in English (Bauer 1983, 1998, Ladd 1984, Sproat 1994, Olsen 2000, Payne & Huddleston 2002, Juhasz et al. 2005, Giegerich 2004, 2009, Gagné & Spalding 2006, Plag 2006, 2010, Plag et al. 2007, 2008, Bell & Plag 2012, among many others).

Other ideas are not pursued here but merit consideration, for example lexicalization of compounds over time and frequency-based approaches to language variation (Bybee 2002, Duñabeitia et al. 2007, Inhoff et al. 2008, Desrochers et al. 2010, among many others). In this paper, compounds that are more frequent in Spanish may have appeared in G2 because they were processed as lexicalized (for example, \textit{aguani}ve ‘sleet’, \textit{bocamanga} ‘cuff’). Furthermore, according to Bybee (2002), high frequency items become automated over time and, in turn, often lead to reductive processes in the realization of language, which, in our data, would apply to F0 activity. Concerning G2, some of its compounds as a whole, as well as their individual members, are highly infrequent in Spanish (for example, \textit{brocamantón} ‘crotch of diamonds’, \textit{bulbocastaña} ‘great pignut’, \textit{varagavilán} ‘ard/scratch plough’), which may have led to processing difficulties in speakers and could, to some extent, explain the greater variation of pitch accents (when present) and IP boundary tones and configurations found in G2, including those that are generally indicative of narrow focus.

In sum, this paper scratches the surface of issues related to the prosody of Spanish compounds. The hope is that its novel findings, as well and the further questions arising from them, may serve as a point of departure for future related studies.

Notes

1. I would like to express my sincere gratitude to four anonymous reviewers for their constructive feedback on an earlier version of this paper, which led to significant improvements.

2. Moyna (2011: 34–44) provides a series of cases that superficially appear to be compounds but that fail to meet her criteria.

3. Spanish in the Tones and Break Indices Framework (Sp_ToBI) is a useful transcription system that has emerged from AM in more recent years (Beckman et al. 2002, Face & Prieto
In this paper, reference is mainly made to AM; however, notation arising from Sp_ToBI analyses has been adopted and the paper's contents are relevant for the Sp_ToBI transcription system as well.

Details on the prosodic hierarchy in both AM and Prosodic Phonology are found in Selkirk (1984, 1995, 2000), Nespor & Vogel (1986), Ladd (1996, 2008), and Gussenhoven (2004). The clitic group and the mora are levels of constituents found in older work (most notably, Hayes 1989) that are less common in more recent studies.

Studies focusing on the syntax–prosody interface refer to this lower level phrase as a phonological phrase (PPH). PPHs refer to major syntactic phrases (XPs) in which heads dominate lower constituents, such as Noun Phrases (NPs), Verb Phrases (VPs), and Adjective Phrases (APs), while IPs are associated with larger syntactic clauses (Selkirk 1984, Nespor & Vogel 1986, Truckenbrodt 1999).

Main prominence may fall in prenuclear position of utterances or phrases in cases of narrow focus, though the unmarked method of conveying such emphasis in Spanish is via syntactic reconfiguration (Face 2000, 2001, 2002). Beyond this, another way of conveying narrow focus is housing words in their own ips, which has the same effect as nuclear position in that words occur just before a phrase boundary in both cases.

Five of the participants are from Mexico City and the remaining three are from northern Mexico, though all eight have spent a considerable amount of time in Mexico City. Based on a review of previous literature on the intonation of Mexican Spanish (for example, Prieto et al. 1996, Sosa 1999, Martín-Butragueño 2004, 2006, Willis 2005, de la Mota et al. 2010), the dialectal differences between participants were not expected to influence the results in the current data elicitation task.

While nominal movement to nuclear position is commonly done to express narrow focus in Spanish, there was no evidence of extra emphasis when compounds were at the end of utterances.

In previous literature on Spanish prosody, deaccenting a content word implies that a pitch accent was supposed to be manifested but actually is not (Rao 2009). It was not completely clear in this paper whether or not the left members of the 30 compounds were expected to carry stress or a pitch accent prior to conducting the data analysis. After the analysis, a spinoff of deaccenting was implemented in order to distinguish cases with evidence of left stress from those without it. Claiming such a distinction made sense thanks to personal communication with Timothy Face.

The apposition reinaluisa ‘lemon verbena’ is taken from Bustos de Gisbert (1986), whose discussion relates to the research questions of this paper. However, it should be noted that Moyna (2011) avoided constructions with proper names (for example, Luisa) because they can obscure which member is the syntactic head. Also, mentions of this compound in her study are spelled with a space between the two members: reina luisa. By contrast, in baño maría ‘double boiler’, another compound in our data set, baño is clearly the head.

While variation in the order of elements in this compound does exist, it should be noted that lobo hombre and lobo-hombre appear mainly in rock music lyrics that allegedly take the term from Boris Vian’s poem entitled Le Loup-garou (El Lobo-Hombre). Additionally, lombombre (Moyna 2011: 345) is attested in data from the 1200s.

References


Bauer, Laurie 1998. When is a sequence of two nouns a compound in English? English Language and Linguistics 2(1): 65–86.


Face, Timothy & Pilar Prieto 2007. Rising accents in Castilian Spanish: A revision of Sp._ToBI. 

Frota, Sónia, Mariapaola D’Imperio, Gorka Elordieta, Pilar Prieto & Marina Viegário 2007. The 
phonetics and phonology of intonational phrasing in Romance. In Pilar Prieto, Joan Mascaro & 
Maria-Josep Solé (eds), Segmental and prosodic issues in Romance phonology. Amsterdam and 

Gagné, Christina & Thomas Spalding 2006. Using conceptual combination research to better 

Garrido, Juan, Joaquim Llisterri, Carme de la Mota & Antonio Ríos 1993. Prosodic differences 
in reading style: Isolated vs. contextualized sentences. EUROspeech ’93: 573–576.

Giegerich, Heinz 2004. Compound or phrase? English noun-plus-noun constructions and the 


University Press.

Hayes, Bruce 1989. The prosodic hierarchy in meter. In Paul Kiparsky & Gilbert Youmans 


Portuguese Linguistics 5(2)/6(1): 59–89.


Hualde, José Ignacio 2010. Secondary stress and stress clash in Spanish. In Marta Ortega 
Llebaria (ed.), Selected proceedings of the 4th conference on laboratory approaches to Spanish 

Inhoff, Albrecht, Matthew Starr, Matthew Solomon & Lars Placke 2008. Eye movements 
during the reading of compound words and the influence of lexeme meaning. Memory & 

Juhasz, Barbara, Albrecht Inhoff & Keith Rayner 2005. The role of interword spaces in the 

266.


Lang, Mervin 1990. Spanish word formation: Productive derivational morphology in the modern 

Liberman, Mark & Alan Prince 1977. On stress and linguistic rhythm. Linguistic Inquiry 8: 
249–336.

Liberman, Mark & Richard Sproat 1992. The stress and structure of modified noun phrases in 
English. In Ivan Sag & Anna Szabolcsi (eds), Lexical matters. Stanford CA: Center for the 
Study of Language and Information. 131–181.


Martin-Butragueño, Pedro 2004. Configuraciones circunfléjicas en la entonación del español 

Martin-Butragueño, Pedro 2006. Proyección sintáctica-discursiva de la entonación circunfléjica 


Author’s address: (Rajiv Rao)

*Department of Spanish and Portuguese*
*University of Wisconsin-Madison*
*1142 Van Hise Hall*
*1220 Linden Dr.*
*Madison, WI 53706*
*USA*
*E-mail: rgrao@wisc.edu*
Appendix

Below is a list of the 30 [NN]N compounds included in the data elicitation task.

1. *aguanieve*: ‘sleet’; *agua* (‘water’) + *nieve* (‘snow’)
2. *balomano*: ‘handball’; *balón* (‘ball’) + *mano* (‘hand’)
3. *baño maría*: ‘double boiler’; *baño* (‘bath’) + *maría* (‘Mary’)
4. *baúl mundo*: ‘large, deep trunk’; *baúl* (‘trunk’) + *mundo* (‘world’)
5. *bienes raíces*: ‘real estate’; *bienes* (‘properties’) + *raíces* (‘roots’)
6. *bocamanga*: ‘cuff’; *boca* (‘mouth’) + *manga* (‘sleeve’)
7. *bocatijera*: ‘pole socket on a carriage’; *boca* (‘mouth’) + *tijera* (‘scissors’)
8. *brocamanto*: ‘crotchet of diamonds’; *brocha* (‘spool/spindle/reel’) + *manto* (‘shawl’)
9. *bulbocastaña*: ‘great pignut’ (type of plant); *bulbo* (‘bulb’) + *castaña* (‘chestnut’)
10. *camión tanque*: ‘tanker truck’; *camión* (‘truck’) + *tanque* (‘tank’)
11. *cartón piedra*: ‘paper mache’; *cartón* (‘cardboard’) + *piedra* (‘rock’)
12. *cepacaballo*: ‘spiny cocklebur’ (flowering plant); *cepa* (‘stump’ or ‘vine’) + *caballo* (‘horse’)
13. *escalera caracol*: ‘spiral staircase’; *escalera* (‘stair’) + *caracol* (‘snail’)
14. *estrellamar*: ‘starfish’; *estrella* (‘star’) + *mar* (‘sea’)
15. *hombre lobo*: ‘wolf man’; *hombre* (‘man’) + *lobo* (‘wolf’)
16. *hormiga caballo*: ‘horse ant’; *hormiga* (‘ant’) + *caballo* (‘horse’)
17. *lápiz plomo*: ‘lead pencil’; *lápiz* (‘pencil’) + *plomo* (‘lead’)
18. *lengua madre*: ‘mother tongue’; *lengua* (‘tongue’) + *madre* (‘mother’)
19. *maestrepasquín*: ‘written expression against authority placed in public’; *maestre* (‘grand master’) + *pasquín* (‘wall poster’)
20. *nochehielo*: ‘cloth stained with black’; *noche* (‘night’) + *hielo* (‘ice’)
21. *pájaro mosquito*: ‘hummingbird’; *pájaro* (‘bird’) + *mosquito* (‘mosquito’)
22. *pájaro polilla*: ‘kingfisher (type of bird)’; *pájaro* (‘bird’) + *polilla* (‘moth’)
23. *palo Brasil*: ‘Brazilwood’; *palo* (‘stick of wood’) + *Brasil* (‘Brazil’)
24. *palo Campeche*: ‘bloodwood tree’; *palo* (‘stick of wood’) + *Campeche* (‘Campeche’; a city in Mexico)
25. *pejemuller*: (also *pez mujer*): ‘manatee’; *pez* (‘fish’) + *muller/mujer* (‘woman’)
26. *pez luna*: ‘ocean sunfish’; *pez* (‘fish’) + *luna* (‘moon’)
27. *rabucocandil*: ‘long-tailed bee-eater (type of bird)’; *rabuco* (from *rabo* ‘tail’) + *candil* (‘oil lamp’)
28. *rataconejo*: ‘coney rat’; *rata* (‘rat’) + *conejo* (‘rabbit’)
29. *reina luisa*: ‘lemon verbena’ (medicinal plant); *reina* (‘queen’) + *luisa* (‘Louise’)
30. *varagavilán*: ‘ard’/ ‘scratch plough’; *vara* (‘long stick’) + *gavilán* (‘quillon’)

118 RAJIV RAO