

Lexical and Postlexical Word Formation

Author(s): Uthaiwan Wong-opasi

Proceedings of the Eighteenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on The Place of Morphology in a Grammar (1992), pp. 467-482

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

Lexical and Postlexical Word Formation¹

Uthaiwan Wong-opasi

University of Illinois at Urbana-Champaign

O. Abstract: This study is intended to lend support to the modular approach to word formation (e.g. Wong-opasi 1987, 1991; Shibatani & Kageyama 1988). Contrary to the strong lexicalist view (e.g. Di Sciullo and Williams 1987) which excludes word formation (WF) outside the lexicon and the extreme nonlexicalist position (e.g. Lieber 1988) which advocates entire WF from syntactic principles, it will be proposed that both theories must be relaxed to allow word formation both inside and outside the morphological component. Evidence will be drawn from compounding and derivation in Romance languages and Thai. Section 1 deals with the structure of words. In the treatment of words as morphological objects, researchers have singled out "headedness" as a basic criterion in the characterization of word structure. However, missing generalizations related to headship in current theories of WF will be pointed out. In particular, it will be shown that D&W's "universal" Right-hand Head Rule (RHR) must remain language-specific. The problems encountered are the directionality of headedness, left vs. right, which differs cross-linguistically and within the same language, as well as various productive Romance compounds which are headless on the surface. Nor can the extreme nonlexicalists handle the complex concord system--internal and external--to the construct, or the exclusion of internal inflection by advocating entire WF from purely syntactic principles. The proposal advanced in this paper is to account for mixed headship and surface headless structure by combined morphological, phonological, and syntactic principles. The interactive principles of morphology, phonology, and syntax during WF are illustrated in sections 2 and 3 through the formations of composite and derived forms, respectively. The autonomy of morphology in words composed of syntactic constituents discussed in section 4 provides strong arguments against exclusive WF in the lexicon or postlexicon. In section 5, a revised version of the theory of Lexical Phonology and Morphology (LPM) is proposed as a model for a modular approach to word formation as it provides comprehensive explanations for synchronic and diachronic data. A wide range of complex data show a continuum of words formed by either morphological, phonological, or syntactic principles separately or a combination of any of the three principles. The last section concludes the present study and offers issues of interest for future research.

1. Morphosyntax of words:

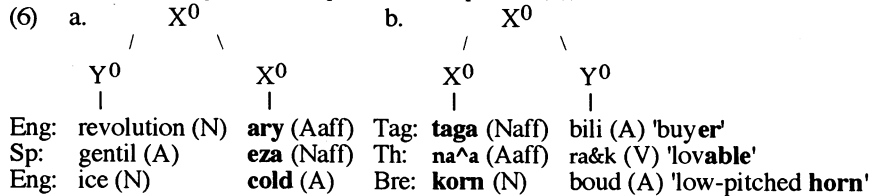
1.1. D&W's HEAD criterion: In the description of the structure of a word, for example in D&W 1987, the notion of "head" is employed to denote a constituent that structurally "dominates" its co-constituent and so "determines" the category of the construct. Thus, D&W's (1987: 23) conception of morphological heads is an extension of the notion of HEAD from syntax into morphology through a general principle called "Percolation". By assigning a lexical category to a suffix, or the "head", and through Percolation, the morphosyntactic features of the head are percolated or encoded onto the whole construct. In this fashion, the category of a construct and the category of its head plus complete agreement in morphosyntactic features between the head and the whole are maintained. Moreover, from the observation that lexical derivation is uniformly carried out through suffixation in English and the Romance languages, D&W proposed a

gender and number agreement in compounds were left unexamined. My data, on the other hand, indicate that the majority of Romance compounds are of the types *timbres-poste* and *essuie-glace*. Contreras 1985, for example, speculates that V+C compounds are perhaps the most productive type in Spanish. The productivity of English phrasal compounds is also unrefutable as defended by Lieber. In addition, WF in isolating languages (e.g. Thai and Chinese) is realized primarily through compounding, rather than affixation since most derivational affixes in these languages are arguably free-morphemes. In fact, the Thai equivalents of various English bound-derivational affixes, e.g. the nominalizers *-er*, *-ist*, *-ation*, *-ness*; the adjectival suffixes *-ble*, *-ed*; the adverbializer *-ly* can function as independent lexical words, and not bound-morphemes (see Wong-opasi, in preparation). Hence, most derived lexical items in isolating languages can be argued to contain phrasal constituents and, thus, the highly productive compounding processes involving phrasal elements are expected. In effect, the major part of the Thai and Chinese lexicons (Chao 1968) comprise compound lexical items. To attribute syntactic compounding to the periphery of grammar is, then, counter-intuitive since the core is reduced to a handful of derivational processes.

Another controversial issue generated by D&W's strong lexicalist position concerns the universal application of the RHR in lexical derivation. If headship is determined by adding a category-changing affix, then a priori, either a LHR or RHR can apply to derive words by affixation on either end of a construct (or in the case of infixation, in the middle of a word). Thus, contrary to English head-final affixal derivation, for languages which derive their lexical categories by prefixation, such as Thai, morphological heads are located on the left.

- | | | | |
|-----|--------------------------------|---|--------------|
| (5) | plot + pløøy
(undo+release) | phu^u + plot + pløøy
(agent. nominalizer+undo+release) | 'liberator' |
| | ra&k
(love) | na^a + ra&k
(adjectivizing prefix+love) | 'loveable' |
| | nu^m + nuan
(soft+nicety) | khwaam + nu^m + nuan
(stative nominalizer+soft+nicety) | 'gentleness' |

1.2. Lieber's Generalized Phrase Structure Principles: The other extremist theory of WF is advocated in Lieber 1988, among others. For extreme nonlexicalists, all word formations are executable solely by syntactic principles which can replace morphological WF principles altogether. Hence, Lieber (1988: 211-213) proposes the structures in (6a) for right-headed affixed forms and Modifier+Head compounds as found in English and (6b) for their left-headed counterparts in Breton, Tagalog, and Vietnamese.² The abundance of phrasal compounds in English is explained in Lieber 1988 by the structure in (6c) where YP covers all phrasal modifier constituents in phrasal compounds. (6d), on the other hand, shows only a slight variation in the direction of the head, as evidenced in postlexical compounds in Spanish (examples in (7)) and Thai.



Sp: mus (N) **araña** (N) Th: **kaa** (N) na&am (N) 'water pot =
kettle'

c.

		X ⁰	
	/		\
	YP		X ⁰
CP:	[a [who's the boss]		wink]
VP:	[a [didn't eat enough]		look]
PP:	[a [right to the bottom]		plunge]
NP:	*[those [a salad and some soup]		suppers]
cf.	[those [salad and soup]		suppers]

Note: According to Lieber (1988: 206), the absence of determiners in the NPs of phrasal compounds is because of some independent reason.

d.

		X ⁰		(Examples are from Thai compounds exhibiting phrasal forms.)
	/		\	
	X ⁰	YP		
CP:	[s ^s a 'shirt'	[twa (classifier) kœ 'good, capable']]		'favorite shirt'
VP:	[khr ^s œ 'machine'	[la&œ 'wash' caan 'dish']]		'dishwasher'
PP:	[khon 'fellow'	[kha ^a œ 'beside' ba ^{an} 'house']]		'neighbor'
NP:	[raœ wan 'prize'	[kha ⁿ 'dipper' na ^{am} 'water' pha ^{an} 'long-stemmed tray'		
	thœœ 'gold']]	'prize of gold-plated water dipper and long-stemmed tray'		

(7) (examples (7a, b, d) from Vallbona 1982: 21, 24, 22; (7e) from travel ads.)

- VP: a. [su 'her' e [limpia-casa] 'cleans-house', [barre-suelos] 'sweeps-floors', [cambia-pañales] 'changes-diapers', [lavaplatos] 'washes-dishes', [marido-en-la-cama] 'husband-in-(the)-bed', y [final-de-día-vacío-rotatoda-por-dentro]] 'end-of-day-empty (i.e. 'an empty day')-crushed-completely-(by)-inside' (all of these composite forms describe the wearisome chores of a housewife) 'her house cleaning, floor sweeping, diaper changing, **dish washing**, husband-in-(the)-bed, and empty and wearisome day'
- PP: b. [una 'a' sarta 'series/string' [de 'of' no-se-puede 'one-cannot-do-it', no-se-hace 'one-does-not-do-it', no-se-tiene 'one-does-not-have-it', no-no-no-no.....'no-no-no-no'.....]] 'a series/string of one-cannot-do-it,
- c. [marido 'husband' [en-la-cama 'in-(the)-bed']] 'in-bed husband'
- AP d. [la nave-[soñada-hacia-el-paraiso]] 'the dreamed-(to)-paradise-ship'
- NP: e. [traslado [aeropuerto-hotel-aeropuerto]] 'airport-hotel-airport transfer'

The rules that generate structures in (6) are as outlined in (8). (8a) is a modified version of X-bar theory, one which allows recursion at the lexical level while (8b) shows the options of having either a phrase or a word as a base constituent. (8c) replaces the usual language-specific Head initial/Head final parameter with more explicit parameters, i.e., the setting whether complements, as well as specifiers and modifiers, precede or follow their heads.

(8) **Lieber's Generalized Structural Principles:** (Lieber 1988: 209)

- n-1
- a. Xⁿ → Xⁿ⁻¹, recursion at least for n=0
- b. Pre- and post-head position can contain Y^{max} or Y⁰
- c. Licensing: i) All and only complements are final/initial
All specifiers and modifiers precede/follow the head

- All complements are Y^{\max} , but not all Y^{\max} are complements
- ii) Case assignment right/left;
case assigned only to maximal projections under adjacency

One strength in Lieber's theory is to capture WF processes from syntactic constructs by a UNIFIED set of principles that applies also to WF from single syntactic elements. However, my data argue for an irreducible autonomy of morphology in WF as evidenced in surface headless compounds in Spanish.

1.3. A compromise: I propose that the "mixed head parameter" is explained when words are generated by different sets of morphological and syntactic principles. That is, lexical compounds are right-headed whereas postlexical composite forms show a left-headed structure following Zwicky's MORPHOSYNTACTIC LOCUS, where inflectional morphemes are located, and MORPHOLOGICAL DETERMINANT, the determining constituent of the resulting category of the construct, criteria for headship, namely, both the morphological loci and morphological determinants MAY be found to coincide on the right-edge of lexical compounds as opposed to the left-hand member of postlexical compounds. Such a distinction could well be suggested to be accommodated in Lieber's theory by claiming these lexical constructs to be "listed" or "memorized" items. Nevertheless, the fact that lexical compounds do manifest word-internal concord is not accounted for by claiming listedness alone. If syntactic principles were ruled out in the formation of lexical compounds, it would be accidental that ill-formed constructs showing no internal concord, e.g. **hierba+bueno* (cf. *hierba+buena*), are never listed. Additionally, Spanish compounds with word-internal concord are counterexamples to Zwicky's claim that the determinant of concord inside a word is not required "because parts of words do not exhibit concord with one another" (Zwicky 1985: 15). The mixed head parameter is most obvious in the assignment of external morphosyntactic agreement to lexical compounds, e.g. *hierbabuena* has an overall feminine gender following the gender of the left-hand member but plurality is only assigned to the right-hand member *hierbabuenas* (**hierbasbuena*). Zwicky's provision that morphological determinant is a property of RULES, and thus not necessarily localizable on a constituent, while correctly captures the discontinuous affixation, infixation, and process morphology, cannot explain this split morphosyntactic agreement, either. Again, we may say that these lexical items are frozen, unanalyzable units. Still, in the same way that plural agreement is prohibited in **hierbasbuena*, plural verbal inflection in Romance V+C compounds is barred, e.g. **tocandiscos* 'record-players' (<- *toca+discos* plays+records), yet the V+C compounds are the most productive compounding process in Spanish. Nor does the listedness argument hold for Spanish diminutives provided that the diminutive formation is extremely productive as well. Furthermore, on the assumption that Romance V+C compounds have an underlying head which may be deleted on the surface, neither syntactic nor morphological principles alone can account for the absence of plural agreement on the V, with or without concomitant spelling out of an overt head on the surface, or the ungrammaticality of deleting the head from similar underlying composite structures in other pro-drop languages like Thai and Chinese. The final point is that the derivation of Spanish specifiers will not be complete without provisions for morpho-phono-syntax interface which should replace Zwicky's 1985a referral rules. In all, my main concern is the shortcomings found in the extreme lexicalist and nonlexicalist positions in advocating WF in

ONE single component in view of the complexity in my data which fosters a modular approach to WF to which I now turn.

2. A modular analysis of Spanish compounds: Lexical compounds differ from postlexical compounds according to their semantic, morphological, and phonological properties. We outline the structures shared by both lexical and postlexical endocentric compounds in (9) and (10).

2.1. Shared composite word structure:

- (9) Right-headed or Lexical Compounds: (accent marks denote prosodic stress)
- a. N+N musaráña, f. 'shrewmouse; bug, worm; foggy illusion'
(mus, n.m. 'mouse' + araña, n.f. 'spider')
 - b. N+N aspaviénto, m. 'exaggerated gestures or feelings'
(aspa, n.f. 'hand (of windmill)³ + viento, n.m. 'wind')
 - c. N+N bocamánga, f. 'cuff, wristband'
(boca, n.f. 'mouth' + manga, n.f. 'sleeve')
 - d. N+N carricóche, m. 'rickshaw, covered wagon'
(carro, n.m. 'cart' + coche, n.m. 'car')
 - e. Head+Modifier⁴ hierbabuéna 'mint'
(hierba, n.f. 'grass' + buena, adj.f. 'good')
 - f. Head+Modifier tióvivo 'merry-go-round, carousel'
tío, n.m. 'uncle' + vivo, adj.m. 'alive'
 - g. V+C lavaplátos, n.m. 'dishwasher'
(lava, v. 3sg. 'washes' + platos, n.m.pl. 'plates')
 - h. V+C cuentagótas, n.m. 'eye-dropper'
(cuenta, v. 3sg. 'counts' + gotas, n.f.pl. 'drops')
- (10) Left-headed or Postlexical Compounds:
- a. N+N hómbr-araña, m. 'spider-man'
(hombre, n.m. 'man' + araña, n.f. 'spider')
 - b. N+N fál-da-pantalón, f. 'culotte'
(falda, n.f. 'skirt' + pantalón, n.m. 'pants')
 - c. N+N búque-escuéla 'training ship'
(buque, n.m. 'ship' + escuela, n.f. 'school')
 - d. N+N diós-héroe, m. 'hero-god'
(dios, n.m. 'god' + héroe, n.m. 'hero')
 - e. Head+Modifier tarjéta-verde 'green card'
(tarjeta, n.f. 'card' + verde, adj.f. 'green')
 - f. Head+Modifier cárro-usádo 'used car'
(carro, n.m. 'car' + usado, adj.m. 'used')
 - g. V+C lávaplátos, n.m. 'dishwashing (chore)'
(lava, v. 3sg. 'washes' + platos, n.m.pl. 'plates')
 - h. V+C árma líos, n.m. 'trouble maker'
(arma, v. 3sg. 'starts, stirs up' + líos, n.m.pl. 'mess')

2.2. Semantic interpretation: On semantic grounds, postlexical compounds unanimously retain the meaning of the head which is on the left while lexical compounds may no longer contain a visible semantic head since lexical compounds may have acquired a non-compositional meaning. The right-headedness of lexical compounds and the left-hand headship of postlexical compounds are reflected by morphological processes such as gender assignment and suffixations of plural and diminutive morphemes.⁵ For example, lexical compounds such as *musaráña* contrast sharply with postlexical compounds like

hombre-araña in that the right-hand member determines the overall gender of lexical compounds, e.g. *una musaraña* (**un musaraña*), while the resulting gender of postlexical compounds follows that of the left-hand member, e.g. *un hombre-araña* (**una hombre-araña*). Likewise, the plural marker attaches to the right edge in lexical compounds, e.g. *musarañas*⁶ (**musesaraña*) but to the left-hand member in postlexical compounds, e.g. *hombres-araña* (**hombre-arañas*). Regarding diminutive formation, the diminutive suffix attaches in a similar contrasting fashion as well, e.g. *bocamanguíta* (**boquitamanga*) vs. *hombrecito-araña* (**hombre-arañita*).

2.3. Phonology of compounds: Turning to phonological distinctions, lexical compounds carry only one primary stress whereas postlexical compounds possess two, one on each composing constituent, as indicated by the prosodic accent marks, suggesting, thereby, that they should be treated as two phonological words. However, despite their double phonological word status, postlexical compounds like *timbres-poste* are single morphological words since they undergo morphological processes as a coherent unit. Although Spanish speakers may have uncertainty in determining the locus of pluralization and diminutivization in postlexical compounds, as they may place a given morpheme on both members of the compounds (Wong-opasi 1987), the majority of speakers suffix the plural marker and the diminutive morpheme onto the left-hand member. In contrast, such options are not available in syntax since all pluralization and diminutivization is governed by syntactic principles and the semantics of each syntactic element. That is, *sofacito-cama*; *sofá-camita*; *sofacito-camita* are all interpreted as 'a little sofa-bed' while *perrito y gata* can only denote 'a little dog and a (regular) cat'. In the same vein, *perro y gatita* just means 'a dog and a little cat'. The same is true for pluralization, i.e., *perros y gata* is not equivalent to *perro y gatas*. Consequently, words like *timbres-poste* cannot be simplistically analyzed as syntactic phrases, as postulated by D&W.

Another morphophonological rule operative solely in the Spanish lexicon is the spelling out of the **terminal element** (TE) of a lexical item. All Spanish lexical entries can be claimed to contain an overt or null ending. For nonverb forms, the TE is the **word marker** (WM) containing such morphosyntactic features as N, V, A, etc., including gender for nouns and adjectives (but excluding the plural morpheme which follows the word marker). According to Harris 1985, Spanish WMs are floating morphemes which is realized by a morphophonological rule spelling out all and only the (universally) accessible WMs, namely-- only one WM for simple nouns, and two for compound lexical items. (11) illustrates the spelling out of only the last WM in simple derived lexical words.

$$(11) \begin{array}{l} [\text{sed}] + \mathbf{a}]_{\text{N}} \rightarrow [\text{sed}] + \mathbf{a}]_{\text{N}} \text{ os} + \mathbf{o}]_{\text{A}} \rightarrow [\text{sed}] + \mathbf{a}]_{\text{N}} \text{ os} + \mathbf{o}]_{\text{A}} \text{ idad} + \mathbf{0}]_{\text{N}} \\ [\text{'silk}]_{\text{N}} \rightarrow [\text{'silk}]_{\text{N}} \text{ y}]_{\text{A}} \rightarrow [\text{'silk}]_{\text{N}} \text{ i}]_{\text{A}} \text{ ness}]_{\text{N}} \end{array}$$

Wong-opasi 1987 proposes a uniform morphological structure for all Spanish lexical items⁷--verbs and nonverbs--as in (12) where the derivational stem(DS) can be the bare root morpheme or be affixed further by one or more derivational stems to derive new words and ends with a TE. (For verbs, the TE is the tense/mood/aspect morpheme, excluding the subsequent person/number marker which may be phonetically filled or zero as well.) (13) shows the morphological analysis of both verbal and nonverbal forms.

(12) [(DS)₁ (DS)₂ (DS)_n + TE]_X

(13) Nonverbs:

[metal]DS₁ + °]_N

[metal]DS₁ {ic}DS_n + o]_A

[metal]DS₁ íst)DS_n + a]_N

[metal]DS₁ ist)DS₂ erf) DS_n + a]_N

Verbs:

[habl]DS₁ á)DS_n + mos] v

[habl]DS₁ á)DS₂ {ba})DS_n + is] v

[habl]DS₁ á)DS₂ r)DS_n + °] v

[habl]DS₁ a)DS₂ r)DS₃ í{a})DS_n + n] v

With respect to WMs in compounds, certain lexical compounds are distinguished from postlexical compounds by the replacement of the internal WM with an *-i-* infix, e.g. *carricoche* (cf. (9d)) vs. **carri-usado* (cf. (10f)); *verdinegro* 'dark green' (<*-verde*, adj.m. 'green' + *negro*, adj.m. 'black') vs. (*televisor*) **blanquinegro* 'black and white (T.V. set)' (<*-blanco*, adj.m. 'white' + *y* 'and' + *negro*, adj.m. 'black'); *patiuerto* 'bow-legged' (<*-pata*, n.f. 'foot' + *tuerto*, adj.m. 'crooked'). Neither can the internal WM of lexical compounds be inflected for gender whereas such a restriction does not apply to postlexical compounds, e.g. **sordamuda* (cf. *sordomuda*), adj.f. 'deaf-mute' vs. *sorda y tonta* 'deaf and dumb'. Conversely, only postlexical compounds, but not lexical compounds, undergo postlexical phonological rules, formulated in accordance with Kiparsky's Structure Preservation Hypothesis which requires that rules creating allophones be operative only in the postlexicon. In effect, the aspiration of the phoneme [s] to yield [h], the velarization of the nasal [n], among other rules do not entail a change in meaning. For example, *dios*, pronounced as *dio[s]* or *dio[h]* equally denotes 'god'. Similarly, both *pa[n]* and *pa[œ]* mean 'bread'. Thus, we conclude that compounds exhibiting postlexical phonological phenomena are postlexical compounds whereas the absence thereof confirms the lexical status, e.g. postlexical compounds *dio[s]/-(h)éroe* (initial *h-* is silent in Spanish) → *dio[h]/-éroe* vs. lexical compounds *mu[s]jaraña* → **mu[h]jaraña*; and *pa[n] y queso* → *[una merienda de] pa[œ] i queso* 'a bread and cheese snack' vs. *pa[n]jiqueso* → **pa[œ]jiqueso* 'a kind of dish'.

3. Morpho-phono-syntactic interface in derivation:

3.1. Derivational vs. diminutive affixation: In D&W, lexical derivation is described as the addition of a category-changing suffix which produces a head-final structure and is purported to be correctly captured by the RHR. However, we have seen that neither the RHR nor Zwicky's morphological determinant criterion is obtained in the case of Spanish diminutive suffixation. Alternatively we would like to suggest treating lexical derivation and diminutive formation as distinct processes. The goals are, on the one hand, to find independent explanations for the "relativized head" phenomenon, and, on the other, to expand Zwicky's notion of morphosyntactic loci to include distinctions between inflection which exhibit word-internal concord as opposed to word-external agreement for a better description of WF in Spanish. That is, such morphosyntactic information as inherent gender must be separate from gender and number received from external morphosyntactic sources in so far as Spanish is concerned. The differentiation between derivational and diminutive suffixes is similar, but not limited, to Sproat's 1985 STEM-level vs. WORD-level affixation distinction.⁸ On morphological grounds, we have seen in the preceding section that derivational suffixes are right-hand heads while the diminutive suffix and its variants generate head-initial structures. Also, initial headship in diminutives differs crucially from

right-headed derivation in that derivational suffixes may alter both the category and the gender of the derived words, e.g. *seda* (n.f.) 'silk' -> *sedoso* (adj.m.) 'silky' -> *sedosidad* (n.f.) 'silkeness' and neither the base nor the suffix trigger word-internal morphosyntactic concord whereas the diminutives must agree in gender with their base forms, e.g. *chico* (n.m.sg.) -> *chiquito* (dim.n.m.sg.) vs. *chica* (n.f.sg.) -> *chiquita* (dim.n.f.sg.). This finding counterclaims Zwicky's dismissal of relevance of determinant of concord in morphology. Moreover, while derivational suffixes have semantic correlates of semantic arguments, diminutive suffixes, on the other hand, do not change the semantic features of the head. The only semantic feature that the diminutive suffix carries is to modify the base, thereby, assuming the role of a semantic functor. Thus, the structure and the relations between the left-element and the diminutive suffix are those of the Head+Modifier, in an order parallel to regular Head+Modifier sequence in Spanish noun phrases. The only parallelism between derivational and diminutive suffixes is that diminutives pattern like derivational suffixes in locating the external gender (and number) marker(s) on the rightmost elements, e.g. *poco(s)* (adj.m.sg.,pl.) -> *poquito(s)* (dim.adj.m.sg.,pl.) vs. *poca(s)* (adj.f.sg.,pl.) -> *poquita(s)* (dim.adj.f.sg.,pl.).

On phonological grounds, we find that derivational affixes attach to stems whereas diminutive suffixes only affix onto word bases and never to a stem. The criteria distinguishing a STEM-base from a WORD-base include two phonological rules. At the stem-level or within the derivational domain in Lexical Phonology's terms, we find alternations between vowels and diphthongs (i.e. [e] -> [ye] and [o] -> [we]). Such alternations are conditioned by stress on the vowel. In effect, alternating mid vowels are diphthongized when stressed, e.g. the stem *herb* -> [*hyérb*] + *a* 'grass' (n.f.) vs. -> [*herb*] *ós* + *o* 'grassy' (adj.) which can be derived further -> [*herb*] *os* *idád* + *o* 'grassiness' (n.f.); the stem *cont* -> [*cwént*] + *o* 'story' (n.m.) or -> [*cwént*] + *a* 'account' (n.f.) vs. -> [*cont*] *á* *r* + *o* 'count, tell' (v.) -> or [*cont*] *ábl* + *e* 'countable' (adj.) and further derivation -> [*cont*] *abil* *idád* + *o* 'accounting' (n.f.). On the contrary, the diminutive suffix is added **after** words are derived. Therefore, it follows that when the base forms which have undergone diphthongization are diminutivized, the diphthongizing effect is carried over even when the vowel in question is no longer in stressed position, e.g. [*hyérb*] + *a* 'grass' (n.f.) -> [*hyerb*] + *ecít* + *a* 'little grass' (dim.n.f.); [*cwént*] + *o* 'story' (n.m.) -> [*cwént*] + *ecít* + *o* 'story' (dim.n.m.); [*cwént*] + *a* 'account' (n.f.) -> [*cwént*] + *ecít* + *a* 'account' (dim.n.f.). Another phonological process which differentiates stems from words is the desyllabification of a high vowel at the edge of a stem during lexical derivation, when it is unstressed (i.e. [i] -> [y] and [u] -> [w]) as in [*Perú*] + *o* 'Peru' -> [*perw*] + *án* + *o* 'Peruvian'. However, these high vowels never desyllabify at the edge preceding the diminutive suffix, e.g. [*rí*] + *o* 'river' -> [*rí*] + *ít* + *o* 'small river'. Furthermore, diminutive suffixes are unlike derivational suffixes in that the former, but not the latter, have a more predictable correspondence between the gender and the vocalic shape of the final morpheme. That is, diminutives have a fixed two-way gender distinction with -*o* predominantly indicating the masculine gender while -*a* signals feminine lexical items.⁹ Derivational suffixes, in contrast, are more idiosyncratic in allowing an optimal number of 3 genders:--masculine, feminine, and neuter, unpredictable from the phonetic shape of the final vowels, e.g. *lino*, m. 'flax vs. *mano*, f. 'hand' vs. *esto*, ntr. (demonstrative pronoun) 'this (matter, stuff)'; *poeta*, m. 'poet' vs.

plata, f.; *padre*, m. 'father' vs. *madre*, f. 'mother'; *marroquí*, m., f. 'Morrocan'; *tabú*, m. 'taboo' vs. *tribu*, f. 'tribe'. Suffixes ending with a consonant also show similar idiosyncrasies.

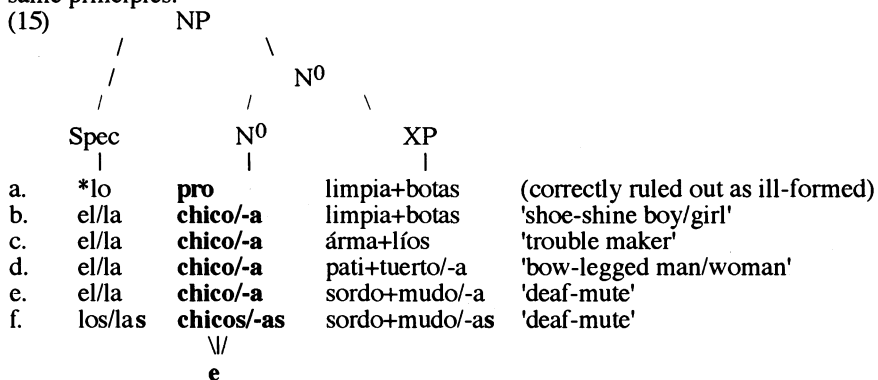
3.2. Derivation of specifiers: Last, but not least, the derivation of the whole class of Spanish specifiers which include determiners, quantifiers, and certain adjectives in the Specifier position is shown to reflect a morpho-phonology-syntax interface. (Wong-opasi 1991). The allomorphic variants of specifiers such as the definite article *el*, *la*, *los*, *las* for m.sg., f.sg., m.pl., and f.pl., in that order, plus the neuter form *lo* as well as strong pronouns and clitics are argued to be derivable from a single source. Of interest here is that one can debate whether these allomorphs should be listed in the lexicon and are associated with lexical items by morpholexical rules. However, the clear distinction between the masculine definite article *el* and its neuter counterpart *lo* and the applicability and nonapplicability of the rule that changes the feminine definite article *la* to *el* in front of a stress [á] can only be explained on the basis of interactive morpho-phonology-syntaxic processes during WF, and not accountable for by Zwicky's 1985a "referral rules". Specifically, *lo* is the derived masculine singular form in front of an underlyingly null head of a noun or noun phrase. While the presence of an underlying head yields *el* as a result of apocope which elides the final [o] and the epenthesis of the default vowel [e] to compensate for lack of syllabicity after apocope. The apparent switch in form (*la*-> *el*), due to a vowel truncation rule which deletes the final [a] preceding a f.sg. head of a noun or noun phrase and beginning with a stressed [á] and a subsequent [e] epenthesis, occurs without affecting the inherent feminine gender of the base form, i.e. *el* is still used in the context requiring a f.sg.def.art. (*el alma* (n.f.sg.)). The postulation of an underlying head N is supported by an identical head deletion phenomenon in noun phrases. The suppression of overt heads is effected by coindexing the features such as [CASE, NUMBER, GENDER] of the heads onto the determiners. (14) shows the morpho-syntaxic structure for these two morpho-phonological rules to apply.

(14)

		NP	
	/		\
	/		N ⁰
	/	/	\
	Spec	N ⁰	X/XP
I.	a. lo/*el	pro (3 sg.m.)	maravilloso (A)/maravilloso y puro (AP)
	b. el/*lo	chico (m.sg.) 'guy'	maravilloso (A)/maravilloso y puro (AP)
	c. la/*el	chica (f.sg.) 'girl'	álta (A) 'tall/high'/ álta y religiosa (AP)
	d. el/*la	álma (f.sg.) 'soul'	(pura) (A) /(pura y generosa) (AP)
	e. la/*el	álma (f.sg.) 'soul'	álta (A) 'tall/high'/ álta y religiosa (AP)
	f. los/las	chicos/-as (m./f.pl.)	áltos/-as / áltos/-as y religiosos/-as
II.	g. el	vaso (m.sg.) 'glass'	lleno de agua (AP) 'filled with water'
	h. la	chica (f.sg.) 'girl'	de uñas rojas (PP) 'with red nails'
	i. los/las	turistas (m./f.pl.)	que llegaron (S) 'who arrived'
		∨∨	
		e	

4. The autonomy of morphology in phrasal compounds: Contrary to Contreras 1987, I propose that V+C compounds have an underlying head which can be deleted from the surface or left intact. Examples with filled heads are taken

from newscasts from Spain (via SCOLA):-- *la máquina quitanieve* (the machine removes-snow) 'snow-plow' and *la ministra portavoz* (the mistress carries-voice) 'spokeswoman minister'. The head deletion is licensed by feature coindexation between the deleted heads and the determiners in essentially the same way as head deletion discussed above. Similarly, Contreras 1987 has proven that determiners in Spanish, but not English, are functional heads which license the absence of the heads of an NP. Given this distinction, it follows that the V+C compounds are less productive in English. It also accounts for the ungrammaticality of deleting the head nouns in Thai phrasal compounds despite the fact that Thai is a pro-drop language, e.g. *khɯŋ^sœ la&œ caan* (machine wash dish) 'dishwasher' -> **la&œ caan*. The head-omission option is excluded in Thai due to the lack of determiners in Thai in the general sense and as functional heads. Finally, postulation of a surface deleted head is the most effective solution to account for the lack of internal agreement in compounds like *patituerto* (e leg-crooked) 'bow-legged man'. On the assumption that the *-i* infix replaces the internal word marker of these compounds, thereby erasing the inherent gender of the base noun and enabling it to be encoded with the morphological features of the null head. Consequently, we can extend the unified structure in (14) to Romance V+C compounds and Spanish surface headless compounds. Thus, V+C compounds with gender variants like *el/la limpiabotas* (e cleans+boots) 'shoe-shine (boy/girl)' as well as headless compounds such as *el/la patituerto/-a* 'bow-legged (man, woman)' plus the choice of definite articles are accounted for by the same principles.



Despite the syntactic origin of these compounds, the fact that inflectional morphemes appear on the right edge only lends very strong support to the morphological status of these constructs since all WF must precede inflection as shown in Diagram 1 (Kiparsky 1982; Zwicky 1985; Wong-opasi 1987).

5. A "modular" approach to word formation via a revised model of Lexical Phonology and Morphology: Wong-opasi (1987, 1991, in preparation) and Shibatani & Kageyama 1988 independently advanced a theory of word formation across the three major linguistic modules, namely--the morphological, the phonological, and the syntactic components. Diagram 1, a revised model of Lexical Phonology, demonstrates the interactions between the three linguistic components. Specifically, not all derivational processes are limited to the confinement of the lexical component. In the same fashion, neither is

compounding found exclusively in the lexical or postlexical components. Moreover, historical data, e.g. *un no-sé-qué* 'a je ne sais quoi' (cf. details in Wongopasi 1987); (Italian) *un nonsoche* (Zuffi 1981: 23) as well as recursion of compounding and further affixation on fully-lexicalized composite forms are allowed in our modular model of word formation. I find examples of recursion such as *limpiaparabrisas* (cleans-stops-winds) 'windshield-wiper', a compound on a composite V+C base form *parabrisas* (stops-winds) 'windshield'; *su gente-buitre-come-entrañas* (his/her people-vulture-eat(s)-bowels) (Vallbona 1982: 17), which mirrors superimposed N+N and V+C structures; affixation on compounds: *paracaídas* (stops-falls) 'parachute' → *paracaidista* 'parachutist'. Finally, the "relativized head" and the "featural atomic head" of V+C compounds are accounted for by the overlapping effects of morphology, phonology, and syntax in WF. These generalizations are inevitably missed in an exclusivist lexicalist or nonlexicalist theory of D&W and Lieber, respectively.

6. Conclusion and direction for future research: I have proposed, in this study, a modular approach to account for various aspects of word construction and illustrated how the proposal can be accommodated by the universal theory. That is, the generalized X-bar schema can be revised to license both single lexical words and phrasal strings as base forms for WF and to permit recursion at the lexical level while the directionality of the head of a lexical item follows the parametric settings for individual languages. I found that Spanish is head-final in the morphological component which is consistent with Spanish language-specific morphological criteria while left-headedness follows Spanish syntactic principles. The mixed head parameter in Spanish compounds is due to interactions between morphological, phonological, and syntactic properties as expected when WF involves materials from more than one module. Recursion that brings syntactic WF processes back into the lexicon is facilitated by the lexicalization process through time with and without concomitant sound change, resulting in the opacity in the internal structure of lexicalized words. Synchronic WF principles are most evident in erroneous translation, e.g. an English head-final compound such as *dealer's maintenance* is interpreted as having an initial head by speakers with less exposure to English and is rendered as **el dealer de mantenimiento* instead of the correct *el mantenimiento de dealer*. It is my hope that the theory of Lexical Phonology is complemented by a thorough study on interventions from syntax, in addition to morphology and phonology, during word formation in this paper. Additionally, along this line of proposals, Shibatani & Kageyama 1988 postulate four characteristic properties which mark postlexical/postsyntactic compounds, from lexical compounds, namely:--referentiality, non-registry, semantic compositionality, and lack of lexical status. Although the last three criteria seem to hold in our analysis of word formation, particularly in Romance and Thai, I would like to add that natural languages may differ in the degrees of conformity to these four criteria, and further studies are encouraged to refine the proposals. For example: the referentiality criterion may apply to only Japanese postsyntactic compounds while Thai lexical compounds are found to violate the anaphoric islandhood, with the use of an empty category (instead of an overt pronoun) to refer to part of a compound, i.e. 'tea' in 'teapot', as in. *kha\$w_i yo&ok kaa+na&m+chaa_j kh\$^n lE&Ew e_ithee e_jloce nay thu^ny* (he_i lift pot+water+tea; up then e_i pour e_j down in cup) 'he lifted the teapot and poured (it) into the cup'.

Notes

¹ An earlier version of the present study was presented at BLS. I wish to thank Michael Kenstowicz, Dieter Wanner, and Hye Suk James Yoon for their insightful comments. Special gratitude goes to Chin-Woo Kim, Jerry Morgan, H-S James Yoon, the Advisory Committee of the UIUC Dept. of Linguistics, and Kevin F. Miller for their support during the writing of this article. Also, thanks to W. Curtis Blaylock for editing an earlier version of this paper. All errors of interpretation are the author's.

² Left-headedness in English is pointed out in Lieber (1988: 214). The prefix *en-* creates verbs from nouns and adjectives, e.g. *encase*, *enrage*, *enable*, *endear*, etc., though *en-* prefixation is an unproductive WF process in current English.

³ The etymology of *aspaviento* is disputed in two sources. According to the *Diccionario etimológico de la Real Academia Española* (1983: 141), it is derived from the infinitive *aspaventar*, (from Latin **expaventa:re*, which in turn, is from *expa:vens*, *-entis*, meaning, 'one who fears'). Corominas (1981: 381), in contrast, claims an Italian origin, i.e. It. *spavento* -> Old Sp. *espaviento* -> Modern Sp. *aspaviento* 'exaggerated gestures or feelings'. The Italian *spavento* also gives rise to Modern Sp. *espanto* 'fear' and is derived from inf. *spaventare* yielding Sp. inf. *espantar* 'to fear'. The change of the initial vowel from *e-* to *a-* was said to be contaminated by the Spanish infinitive *aspar* 'to torment', the form which caused reinterpretation in Spanish as a compound of *aspa*+*viento* (with the first element taken as meaning the hand of a windmill).

⁴ The reversed order is more rare. Examples are *gentilhombre* 'gentle-man', (from *gentil+hombre*, equivalently 'gentle+man'); *buenaventura* 'good luck' (from *buena+aventura*, lit. 'good+adventure/daring enterprise'). These compounds have the structures: Adj.+N alongside *bienvenida* 'welcome' (from *bien* (adv.) +*venida* (n.f.), lit. 'well+coming'), and Adv.+N/Inf. such as *bienestar* 'well-being, welfare', among others.

⁵ Apart from the diminutive suffix *-(e(c))itV*, and such regional variants as *-ikV*; *-ín/a*, we also find augmentative suffixes, e.g. *-ote/a*; *-ón/a*, with the same morphological and phonological behaviors. In Wong-opasi 1987, the covered term employed for word-level suffixation of this nature is "productive formation".

⁶ The correct gender and number assignment is confirmed by the use of the word in the feminine plural form to mean 'foggy illusions' as in *mirar a las musarañas* (coll) to stare vacantly (literally, to look at foggy illusions)' (Williams 1986: 243).

⁷ The morphological plane is given in this form, following Wong-opasi 1987, primarily for the purpose of stress assignment where prefixation is permitted, although it plays no role in stress placement since stress is predictably allocated on the last syllable of the rightmost derivational stem, unless the said syllable is lexically marked as extrametrical, as shown here in curly brackets.

⁸ Wong-opasi 1987 argues for more-constrained distinctions between stem- and word-level affixations than the one Sproat proposed. That is, the lexicon of Spanish is stratified into the Derivation, Compounding, Productive Formation (including Diminutives), and Pluralization strata (cf. Diagram 1). The diminutives exhibit word-level suffixation behaviors despite their apparent single WM which might suggest a stem-level process (see also Wong-opasi (1992)).

⁹ Exceptions to this generalization are also predictable. Diminutives inherit the idiosyncrasies of the base derivational forms in two systematic ways. Either

retaining the gender and the final vowel of the base, e.g. *poema*, **m.** -> *poemita*, **m.**, e.g. *radio*, **f.** -> *radiécita*, **f.**; *mano*, **f.** -> *manito*, **f.**, or the gender alone, e.g. *mano*, **f.** -> *manita*, **f.**

Bibliography

- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Contreras, Heles. 1985. Spanish Exocentric Compounds. *Current Issues in Hispanic Phonology and Morphology*. Frank Nuessel (ed.). Bloomington: IULC, 14-27.
- Contreras, Heles. 1987. On Spanish Empty N' and N. *Studies in Romance Linguistics*, C. Kirschner & J. DeCesaris (eds.). Amsterdam: John Benjamins, 83-95.
- Corominas, J. & J. A. Pascual. 1980. *Diccionario crítico etimológico castellano e hispánico*. Madrid: Gredos.
- Di Sciullo, Anna-Maria & E. Williams. 1987. *On the Definition of Word*. Cambridge: MIT Press.
- Harris, James W. 1985. Spanish Word Markers. *Current Issues in Hispanic Phonology and Morphology*. Frank Nuessel (ed.). Bloomington: IULC, 34-54.
- Jaeggli, Osvaldo. 1980. Spanish Diminutives. *Contemporary Studies in Romance Languages*. Bloomington: Indiana University Linguistics Club.
- Kiparsky, Paul. 1982. Lexical morphology and phonology. In I.-S. Yang (ed.), *Linguistics in the Morning Calm*. Seoul (Korea): Hanshin Publishing Co., pp. 3-91.
- Lieber, Rochelle. 1988. Phrasal Compounds in English and the Morphology-Syntax Interface. *Proceedings of the Chicago Linguistic Society* 24: 202-222.
- Real Academia Española. 1959. *Gramática de la lengua española*. (Nueva edición, reformada, de 1931) Madrid: Espasa-Calpe.
- Shibatani, Masayoshi & Taro Kageyama. 1988. Word Formation in a Modular Theory of Grammar. *Language* 64: 451-484.
- Sproat, Richard. 1985. *On Deriving the Lexicon*. MIT Ph.D. dissertation.
- Vallbona, Rima. 1982. *Mujeres y agonías*. Houston: Arte Público Press.
- Williams, Edwin B. 1986. *The Bantam New College Spanish and English Dictionary*. New York: Bantam Books.
- Wong-opasi, Uthaiwan. 1987. *Lexical Phonology and the Spanish Lexicon*. Bloomington: Indiana University Linguistics Club (1989).
- Wong-opasi, Uthaiwan. 1991. On Deriving Spanish Specifiers: Morpho-Phono-Syntactic Interactions. *New Analyses in Romance Linguistics*. Dieter Wanner & Douglas Kibbee (eds.). Amsterdam: John Benjamins, 101-122.
- Wong-opasi, Uthaiwan, in preparation. The ECP & the Head Parameter. ms., UIUC.
- Zwicky, Arnold. M. 1985 Heads. *Journal of Linguistics* 21: 1-29.
- Zwicky, Arnold. M. 1985a. Rules of Allomorphy and Phonology-Syntax Interactions. *Journal of Linguistics* 21: 431-436.

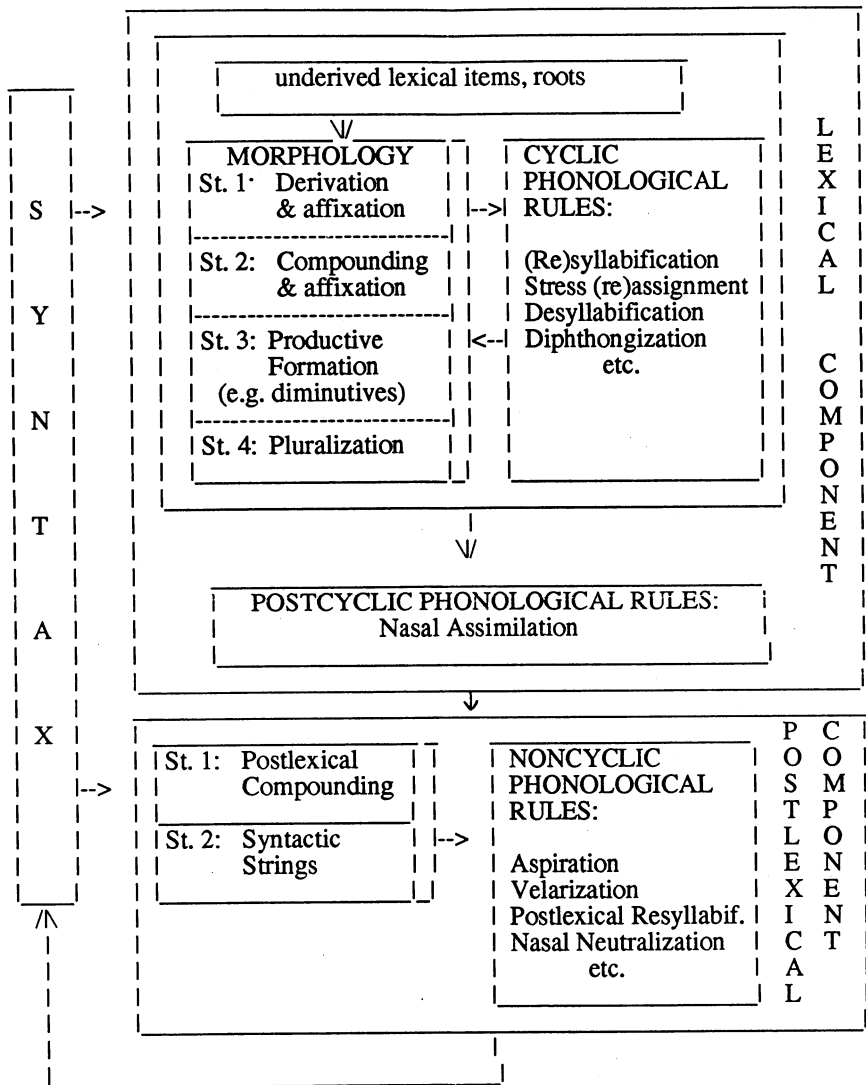


Diagram 1: Revised Model of Lexical Phonology