Chapter 17

Categorizing verb-internal modifiers

Chenchen Song
University of Cambridge

In this chapter, I propose a novel theory to explain the syntactic and semantic characteristics of a class of previously lesser studied verb modifiers, namely the non-heads of compound verbs like *double-check* and *hand-wash*. Such verb-internal modifiers are more widely used in English contrary to common impression, and they are also a very productive compounding strategy in East Asian languages like Chinese and Japanese. However, in the familiar European languages they are either systematically missing (e.g. in Romance languages) or subject to odd movement constraints (e.g. in German), even when these languages do not equally lack compound nouns. The theory I propose makes use of a defective categorizer that bears a lexically unvalued categorial feature. It agrees with the categorial feature on the base verb and results in a word-internal adjunction structure. The model is solely based on Simplest Merge without resorting to Pair Merge or Root incorporation, can be readily extended to the nominal domain, and nicely relates the typology of verb-internal modifiers to the parametrization of verb movement.

1 Introduction

There is a long line of syntactic research on verbal modifiers (VMs, É. Kiss's 2002 term), most fruitfully on verbal particles, as represented by those in Germanic languages (e.g. German ***ein**-kaufen*[1] 'in-buy; to shop', cf. Dehé et al. 2002, Haiden 2006, and references therein) and Hungarian (e.g. ***ki**-mos* 'out-wash; to wash out', cf. É. Kiss 1987, 2002, 2008, Hegedűs 2013). A standard view on the particle-like VMs is that they are base-generated as V-complements, e.g. in small clauses (Taraldsen 1983, Kayne 1985 et seq.). They are modifiers in the broad sense that non-heads in a phrase enrich the head's meaning.

[1]The hyphen is used for expository convenience and does not indicate orthography.

Chenchen Song. 2020. Categorizing verb-internal modifiers. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences I: Syntax inside the grammar*, 357–384. Berlin: Language Science Press.
DOI: 10.5281/zenodo.3972866

There is still another type of VM which has received comparatively less attention in traditional generative studies. Observe the examples in (1).

(1) *double*-check, *second*-guess, *proof*-read, *dry*-clean, *hand*-wash, *stir*-fry, *sleep*-walk, *window*-shop, *baby*-sit, *breast*-feed, *hitch*-hike …

While the italicized components in (1) are intuitively also modifiers, these complex verbs are traditionally treated as compounds, i.e. lexical items, whose internal structures are a matter of morphology rather than syntax. In other words, the VMs in (1) are word-internal; call them verb-internal modifiers (VIMs). Unlike verbal particles, VIMs are neither V-complements nor secondary predicates. Rather, they modify the base verbs in the same way adverbs modify VPs.

Contrary to the common impression that compound verbs are unproductive in English, English speakers are evidently no less capable of creating items like (1)[2] than e.g. speakers of Chinese, which is considered very productive in compound verbs.[3] If compounding is part of our language competence, it should be subject to general linguistic principles and, crucially, only rely on computational mechanisms made available by Universal Grammar (UG), hence no compounding-specific rule. Distributed Morphology (DM, Halle & Marantz 1993 et seq.) treats syntax (essentially Merge, Hauser et al. 2002) as the only generative engine in the human language faculty (single engine hypothesis, Marantz 2001). I take this as my point of departure.

With these theoretical advances, many issues about compounding need to be carefully rethought, as witnessed by the numerous works within DM (i.a. Zhang 2007; Harley 2009; Hu 2013; Nishiyama & Ogawa 2014; Bauke 2016; de Belder & van Koppen 2016; Song 2017b). This chapter furthers this exploration by putting forward a new perspective on the structure of VIMs. To be specific, I categorize VIMs via a lexically unvalued "defective categorizer" and assign them the categorial value of the base verbs via Agree. This new model has three major advantages. First, it is solely based on Simplest Merge and labeling (Chomsky 2013), making no use of Pair Merge or Root incorporation. Second, it can be extended to the nominal domain, unifying verbal and nominal compounding. Third, it relates the typological availability of VIMs to the parametrization of verb movement.

This chapter is organized as follows. In §2, I illustrate the categorial properties of VIMs with cross-linguistic data, concluding that they are simultaneously

[2] Syntacticians are contributing quite a bit to this list. A quick Google search finds the following examples in the published literature: *set/pair/self-merge*, *head/phrasal/A/Ā/wh-move*, *left/right-adjoin*, etc. All are attested in the PRS.3SG form, so they are unequivocally used as verbs.

[3] The productivity of compound verbs is influenced by multiple factors, e.g. (1)-type compounds in Chinese are extremely productive because they form standard prosodic words (Feng 1997).

lexical and functional and qualify as word-internal adjuncts. In §3, I review two minimalist approaches to adjunction, arguing that the labeling-based model is more favorable. In §4, I propose and motivate such a model, featuring a defective categorizer and a Root-joining schema. In §5, I further discuss the theoretical and typological predictions of the model. §6 concludes.

2 The categorial status of verb-internal modifiers

As a general observation, VIMs can be of any lexical category, as in (2).[4]

(2) a. English
hand$_{N}$-wash, *stir*$_{V}$-fry, *dry*$_{A}$-clean, *under*$_{P}$-score

b. Chinese
shǒu$_{N}$-xǐe 'hand-write; to handwrite', *zǒu*$_{V}$-dú 'walk-read; attend a day school', *dà*$_{A}$-xiào 'big-laugh; to laugh loudly'

c. Japanese
se$_{N}$-ou 'back-carry; to carry on back', *oshi*$_{V}$-taosu 'push-topple; to push down (topple by pushing)', *chika*$_{A}$-zuku 'close-attach; to get near'

One may be tempted to conclude that VIMs simply belong to their separate lexical categories. This conclusion is problematic in several ways. First, it misses the generalization that VIMs, whatever their lexical source, all perform the same function (i.e. modification). This issue does not arise in traditional studies where VIMs have no syntactic relevance whatsoever, but in the single-engine approach, we need to syntactically formalize this "beyond-lexical" equivalence class.

Second, even the lexical labels themselves may not be tenable, for VIMs and the respective lexical categories do not have much in common beyond the superficial resemblance. Consider the "N" modifiers in (2). They repel typical nominal distributions such as pluralization and quantification in English (3a), classification in Chinese (3b), and adjective modification in all the three languages (3c).

(3) a. * hands-wash, * one hand-wash

b. * yì zhī shǒu-xiě 'one CLF hand-write'

c. * pretty hand-wash, * qiǎo shǒu-xiě 'skillful hand-write', * aoi se-ou 'blue back-carry'

[4]Chinese and Japanese have no P-origin VIMs because they lack the English-type P items (cf. Huang et al. 2009; Tsujimura 2013; Song 2017a). I leave P-related issues aside due to space limit.

Since no distributional criterion can tell us *hand*, *shǒu*, and *se* are nouns, the label N can only come from the impression that they are usually used as nouns elsewhere (the same is true for the other VIM labels). However, such impression-based categorization is unreliable, because the same form may be reused in different categories, e.g. *a* ***hand***$_N$ vs. *to* ***hand***$_V$ *in the essay*. The invariant part here is the Root √HAND rather than its categorized products.

Third, some VIMs do not fall in any existing lexical category, such as the prefixes in ***re****-build*, ***un****-fold*, ***dis****-close*, ***mis****-understand*, etc. They perform the same "adverbial" function as the other VIMs we have seen but cannot be categorized by impression. Similarly, in some Japanese V-V compounds, the first component is so bleached[5] that its assumed category V becomes vague, as in (4).

(4) *sashi*-semaru 'put-come.close; be imminent', *tori/tott*-tsuku 'take-attach; cling to, be obsessed', *hin*-mageru 'pull-bend; bend, distort', *butt*-taosu 'hit-topple; violently topple'…

According to Kageyama (1993), these italicized forms have become intensifying prefixes. Like English *re*-, *un*-, etc., they cannot be classified into any category.

In sum, if we want to identify a unified syntactic category for VIMs, the ordinary lexical categories are not a good place to look; the more plausible place is their functionality instead. That is, albeit counterintuitive, VIMs may form a functional category. This said, however, they are not inflectional, because canonical inflectional categories are closed classes, often with dedicated exponents, e.g. *-ed* for past tense. Being an open class with no fixed exponents, VIMs are again more like lexical categories.

This categorial status is reminiscent of the functionally "recycled" lexical items in Biberauer (2016a, 2017). According to Biberauer, recycling effects such as grammaticalization and multifunctionality are a distinctive property of natural languages, reflecting the domain-general third factor maximize minimal means. I illustrate this point with Chinese light verbs (5) and classifiers (6) (see Biberauer's works for more cross-linguistic examples).

(5) a. *dǎ*-rén 'hit-person; to hit someone' vs. *dǎ*-yú 'DO-fish; to fish'
    b. *bǎ*-zhù fúshǒu 'hold-still handrail; to firmly hold the handrail' vs. *bǎ*-shū dǎ-kāi 'DISP-book hit-be.open; to open the book'

(6) a. bǐjì-*běn* 'note-book' vs. yì *běn* shū 'one CLF book; a book'
    b. shuǐ-*bēi* 'water-glass' vs. yì *bēi* shuǐ 'one CLF water; a glass of water'

[5]The bleaching is not only semantic but also phonological, e.g. *tott*<*tori*, *hin*<*hiki*, *butt*<*buchi*.

Light verbs and classifiers have lexical origins, and they still keep much idiosyncrasy as function words, as evidenced by the numerous same-function items in (7) which are nonetheless non-interchangeable.

(7) a. 'Do' light verbs: *dǎ* 'hit', *zuò* 'make', *nòng* 'play around' …
    b. Disposal light verbs: *bǎ* 'hold', *jiāng* 'lead, support', *guǎn* 'manage' …
    c. Classifiers: *běn* 'for books', *bēi* 'for liquid in glass', *tóu* 'for animals' …

Similar flexibility exists in other languages, e.g. there are at least four productive light verbs in English: *do*, *take*, *make*, and *have*. The cross-linguistic widespreadness of semi-functional items implies some basic generative strategy. Biberauer (2016b: 5) identifies this strategy as adjoining featurally underspecified elements (effectively Roots) to null functional heads.[6] Following this idea, the functional heads behind light verbs and classifiers are Larsonian VP-shells (e.g. Voice, Appl, cf. Lohndal 2014) and Cl (Borer 2005; Feng 2015; Huang 2015). By comparison, the head H behind VIMs is much less clear-cut. It cannot simply be VIM, for that would entail an ad hoc formal feature (FF) [VIM] which makes little sense in our feature system.[7] Nor can it be any VP-shell category, because on the one hand, VIMs are inside the complex verbs rather than above VP; on the other hand, while VP-shells and Cl only recycle from V and N sources respectively (in line with Roberts & Roussou's 2003 observation that grammaticalization is always upwards in a functional hierarchy), H can recycle from any contentful morpheme without categorial restriction, which makes the process more like *lexicalization*, with H systematically converting various concepts into lexical items, just like categorizers. This effectively bears out the DM view that the non-heads of primary compounds are bare Roots (cf. de Belder 2017), though I deviate from (almost all) previous DM approaches to compounding from RootP incorporation (e.g. Harley 2009) to Root–Root merger (e.g. Zhang 2007; Bauke 2016), for reasons to be spelled out in §4.

In fact, since the VIM is merged as a non-complement non-projecting sister of V, it is essentially a V-adjunct, which means H, if existent, systematically creates head adjunction. As such, a proper syntactic model of VIMs relies on an adequate theory of adjunction. I briefly review theories of adjunction in the next section.

[6] This idea deviates from DM. First, it relies on a conception of Root broader than that in DM (but closer to that in Borer 2013), for not only lexical but also functional forms can be recycled (e.g. $\textit{that}_{\text{D/C}}$). Second, it violates the DM assumption that Roots cannot appear without being categorized by one of the category defining heads (the categorization assumption, Embick & Marantz 2008; see Song 2017c for a less restrictive version compatible with Biberauer's idea).

[7] According to Zeijlstra (2008) and Biberauer (2016b, 2017), FFs piggyback on substantive features, so [Person] and [Gender] are legitimate FFs while [Affix] and [Complement] are not.

3 Minimalist approaches to adjunction

3.1 Pair Merge

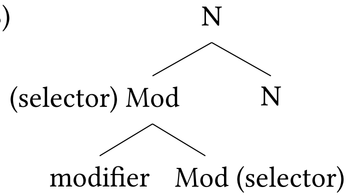
One may wonder: if VIMs are adjuncts, why do we need to give them any functional head at all? Shouldn't their modifier role be self-evident? These questions implicitly take adjunction and its asymmetric effect for granted, which is undesirable given the (beyond-)explanatory goal of the minimalist program.

The standard minimalist approach to adjunction is Pair Merge (Chomsky 2000, 2004), which takes two syntactic objects α, β and yields an ordered pair ⟨α, β⟩. α (the adjunct) is attached to β from a separate plane, which is invisible to and thus cannot interfere with the primary-plane derivation. Following this idea, adjunction does not need any functional head but is a special operation. However, Pair Merge sacrifices the minimalist and evolutionary advantages of the theory, because, as Collins (2017: 52) points out, it has to be stipulated as an independent UG operation, which goes against the strong Minimalist thesis (SMT, "language keeps to the simplest recursive operation", Berwick & Chomsky 2016: 71). Chomsky (2013: 40) also criticizes the "extension of Merge", arguing that there is no remerge, multidominance or late Merge (among others), but only simple Merge.

Also note that the motivation of Pair Merge is empirical ("it is an empirical fact that there is also an asymmetric operation of adjunction", Chomsky 2004: 117), but its problem is conceptual. As such, if we could give the "empirical fact" an alternative explanation, Pair Merge would no longer be needed. I will discuss such alternatives in §3.2. For now, let's turn to another problem of Pair Merge, raised in Rubin (2003):

> We need to avoid circularity here, so we cannot simply say that we want adjuncts to be adjuncts, so we invoke pair-Merge, which creates adjuncts. Before any two expressions are merged, relational terms such as *adjunct*, *complement*, and *specifier* are premature. (Rubin 2003: 663)

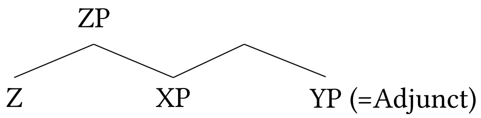
The problem is essentially how syntax can determine Pair Merge is appropriate for adjuncts. Rubin's solution is a dedicated functional head Mod, which "forms an extended projection around all base adjuncts" such that "[a]ny phrase headed by Mod is subject to pair-Merge" (p. 664). This idea is not so different from our functional head H in §2 and also compatible with the Borer–Chomsky conjecture (BCC, Baker 2008), which highlights the fundamental role of features. However, the solution is not optimal. First, as Arsenijević & Sio (2009: 2) notice, when Mod connects a modifier to a noun (both are phrases in Rubin 2003), it selects twice – first the modifier and then the noun, as in (8) – but Pair Merge only happens in the second selection, which makes the triggering effect of Mod inconsistent.

(8)


Second and more relevant to us, Mod has no substantive featural basis. Though Rubin (2003: 666) specifies its semantic type as ⟨⟨e,t⟩,⟨⟨e,t⟩,⟨e,t⟩⟩⟩ ("a function from predicates to properties of predicates"), this only describes the function we want Mod to perform but does not relate it to any conceptual interpretation. So Chomsky's (1995) criticism of Agr (that it is present only for theory-internal reasons) also applies to Mod. The above two problems may not be unsurmountable, but they do show that Rubin's intuition can be further developed.

3.2 Labeling

While Rubin (2003) "determines" Pair Merge and justifies its role in adjunction, Hornstein (2009) and Oseki (2015) dispense with it and derive adjunction via Simplest Merge (Hornstein's "concatenate") plus labeling. Following Epstein et al. (2012), Oseki (2015) assumes when two phrases XP and YP merge but share no feature, the merger cannot be labeled. Adopting the label accessibility condition (LAC, Hornstein 2009: 90, Epstein et al. 2012: 254),[8] which states that unlabeled syntactic objects cannot be accessed by Merge,[9] Oseki further claims that at this stage the derivation can only proceed by letting one of XP and YP participate in further Merge, thus yielding the "two-peaked" structure in (9). In Hornstein's terms, YP "dangles off" the [$_{ZP}$ Z XP] complex.

(9)


Epstein et al. (2012: 261) conceive this structure as "two intersecting set-theoretic SOs". Crucially, one peak must be removed (via Transfer) from the narrow syntax, which then becomes inaccessible to later derivation, rendering the island effect.[10]

[8] Epstein et al. (2012: 262) deduce LAC from minimal search and conceive it as a third factor consequence in the sense of Chomsky (2005).

[9] This view is not unanimous, e.g. for Chomsky (2013) labels are only needed by the interfaces. As such, the indispensability of LAC in Epstein et al.'s model may turn out to be a disadvantage.

[10] Epstein et al.'s main focus is the Spec-TP subject. Oseki extends their model to adjuncts.

Several issues remain unclear. First, the definition of "peak" is vague. Geometrically, a peak consists of two branches, but then removing a peak amounts to removing an entire {XP, YP}, which means XP cannot stay in syntax to merge again. Second, even if XP could stay, since Transfer cannot undo Merge, the removal of YP cannot save the second merger of XP from violating *no tampering*, and since the intersected element is contained in two sets, set intersection inevitably leads to multidominance. Third, for Epstein et al. the removed peak is consistently the phase-head-complement, but this causes trouble for Oseki, as it wrongly predicts that adjuncts are only ever adjoined to phase heads.

While the two-peaked model is far from ideal, the labeling idea behind it is indeed more advantageous than Pair Merge: (i) it obeys the SMT and is evolution-friendly, (ii) it reduces the specialness of adjunction to specific features, in line with the BCC. Remember that Rubin's (2003) idea was also to reduce adjunction to a specific category, which makes it potentially compatible with a labeling-based approach. Thus, instead of resorting to "unlabelable" scenarios (e.g. the two-peaked model), we could also seek a solution from scenarios where labeling can normally proceed (as in Rubin's model). I propose a new model along this line in §4.

3.3 Interim summary

To recapitulate §§1–3, the structure of verb-internal modifiers (V-level adjuncts) is a tricky issue in syntactic approaches to word-formation, partly due to the elusive categorial status of VIMs and partly due to the unavailability of a satisfactory theory of adjunction. The two problems point to the need of a categorial account of adjunction, e.g. via a mediating functional head H. As such, among previous approaches to adjunction, those based on labeling (manipulating categories) are more advantageous than those based on Pair Merge (a specialized UG operation). In addition, among potential labeling-based theories those featuring "labelable" scenarios are more coherent than those featuring "unlabelable" ones.

4 Deriving verb-internal modifiers

4.1 How *not* to merge a Root

As mentioned in §2, the relation between H and VIMs is similar to that between categorizers and Roots. Note that I did not prove the necessity of H, but only speculated that it could potentially replace Rubin's (2003) Mod. Two points could make this speculation suspicious. First, labeling (essentially minimal search) does not need any special head to proceed. FFs on the Merge input alone are enough. Second, if VIMs are Roots, then H can be nothing but a categorizer (à la categorization assumption, cf. footnote 6), which leads to a dilemma, for no existing lexical category is adequate for VIMs.[11]

This dilemma is faced by all models applying ordinary categorizers to compound non-heads (e.g. Harley 2009 for compound nouns), but it does not force us to resort to uncategorized "floating Roots" (e.g. de Belder & van Koppen 2016) or postsyntactic Root operations (e.g. fission, de Belder 2017), especially if those solutions rely on unwarranted definitional extension of Root, which is no more desirable than extension of Merge. Below I will defend the conservative position that Roots are bare (FF-less), syntactically inert (no √P), and must be categorized.

To begin with, the bare Root view is faithful to the original purpose of the Root theory, i.e. lexical decomposition.[12] Lexical decomposition targets non-primitive lexical items (LIs) and submits that any composite LI, be it a pure FF bundle or an FF-equipped Root, has to be assembled from smaller atoms rather than appearing as such all of a sudden. This is evidenced in language acquisition/change, where feature bundles are gradually formed and further alterable.[13] To wit, any theory working with bundled features has to assume some LI forming mechanism, including DM.[14] However, as Collins (2017) remarks, this poses a conceptual problem, because "that mechanism is not Merge":

> This state of affairs seems undesirable for two reasons. First, humans have an unlimited capacity to learn and to coin new lexical items, just like they have an unlimited capacity to form new phrases [...] Second, adding a new mechanism (to form lexical items) would increase the complexity of UG, going against the SMT. (Collins 2017: 61)

Collins concludes that LIs are formed by Merge. So, FF-equipped Roots, if any, must also be products of Merge, which takes bare Roots and FFs as input. In short, the single engine hypothesis and SMT together force a bare Root view.

[11]Similar considerations led de Belder & van Koppen (2016) to conclude that the non-heads of some Dutch compound nouns are bare Roots without any functional category, not even categorizers.

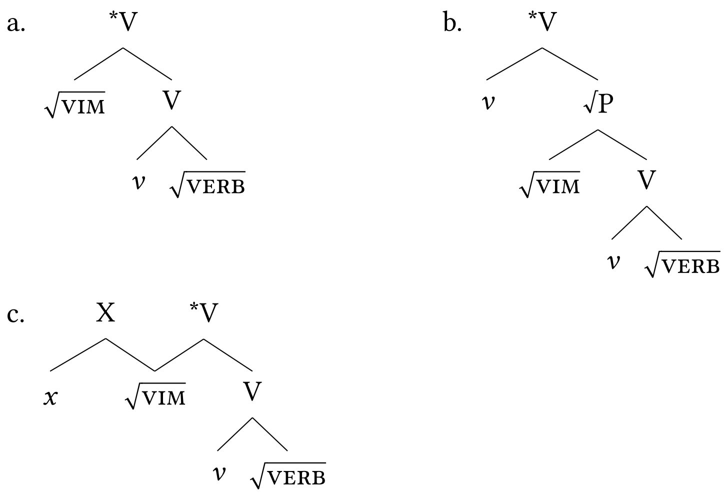
[12]See Ramchand (2008: 11) and Gallego (2014: 192) for summaries of various Root views.

[13]Despite the intuition that we use LIs as whole units, the existence of sub-LI knowledge has never been denied (hence the branch "morphology") – it has simply been handled by a separate generative engine (the lexicon). In this sense, lexical decomposition is not introducing anything new but merely aims to capture the sub-LI knowledge in the single-engine framework.

[14]Marantz (1997: 203) conceives the DM narrow lexicon as "generative", as it contains "atomic bundles of grammatical features [that are] freely formed, subject to principles of formation".

Following this line of thought, if Roots are stored bare presyntactically,[15] they must be inert in narrow syntax, which only manipulates FFs. Among others, this means Roots cannot head or project/label, hence no √P (in line with i.a. Acquaviva 2009, Borer 2009, 2014, Chomsky 2013, de Belder 2011 et seq., Alexiadou 2014; contra Cuervo 2014, Harley 2014). Moreover, since no featural dependency could ever be established on Root nodes, nor could they be moved or host movement,[16] hence there can be no Root incorporation (contra Harley 2009). The only way a Root may participate in syntactic derivation is via categorization, either exclusively by the lexical categorizers (as in standard DM) or by any functional category (as in Borer 2005; 2013; Biberauer 2016b; Song 2017c). What matters here is that there can be no floating Root, i.e. every Root must be the most deeply embedded element in its workspace (a conclusion compatible with Marantz 2001 and Boeckx 2014). As such, the model in (10a) is infelicitous, for it is impossible to categorize the VIM Root without letting it project (10b) or remerge (10c).[17]

(10)



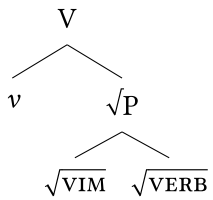
Note that (10c) is the two-peaked structure in §3.2. Despite its infelicity, the idea that √VIM may be categorized in adjunction is insightful. Therefore, if we could overcome the multidominance problem, (10c) may well become a felicitous model.

[15]This does not rule out the possibility that non-bare Roots (just like other composite LIs and even larger phrases) could be lexicalized and stored postsyntactically (in DM lists 2 and 3) or extra-syntactically (as general experience, cf. Marantz 2013).

[16]Thus, Roots may be conceived as adjuncts (à la Marantz 2013).

[17]Strictly speaking, √VIM can only be categorized via the multidominance structure in (10c), because in (10b) what the upper *v* categorizes is √P rather than √VIM. Besides, (10b) wrongly predicts VIMs can only be V-origin.

I will further pursue this route in §4.2. For now, let's turn to another infelicitous structure in (11).

(11) 

This is the compounding model adopted in i.a. Zhang (2007), Borer (2013), Bauke (2014; 2016) and de Belder & van Koppen (2016). A clear problem with it is the symmetric relation between the two Roots, which means there is no way to determine which Root is the modifier and which is the verb at logical form (LF), nor can they be algorithmically linearized at PF. Borer (2013) resorts to Root incorporation to yield the asymmetry, but this operation is illegitimate under the bare Root view, as FF-less objects cannot be moved.[18] For more thorough arguments against direct Root-Root merger see Song (2017c).

With (10–11) ruled out, we are left with only one structure to derive VIMs, i.e. [H √VIM]-[*v* √VERB], where the two Roots are separately licensed before being joined together. The necessity of a functional head H is thus proved, not by the requirement of labeling but by the nature of Root.

4.2 Defective categorizer

Further examination of the structure [H √VIM]-[*v* √VERB] reveals that H and *v* must share some feature(s), for otherwise the structure is unlabelable.[19] However, H cannot simply be *v*, because that would make the structure symmetric just like (11) and its two branches formally undistinguishable (distinctness is an important interface principle, cf. Richards 2010). Rather, H and *v* should be simultaneously homogeneous and non-identical, and ideally the distinction should not be achieved by bundling extra features into H/*v*, for that would go against the

[18] De Belder (2017) proposes a fission-based variant of (11), where the two Root nodes are "split" postsyntactically and the asymmetry is yielded by "the order of insertion". I do not have space to evaluate this approach, but *ceteris paribus* the model I will propose later is free from postsyntactic operations and thus potentially more parsimonious.

[19] Here I follow Chomsky's (2013, 2015) conception that all branching nodes (i.e. all products of Merge) must be equipped with a label at the interfaces. See Bošković (2016), Bauke & Roeper (2017) for looser positions and Collins (2002 et seq.) for a label-free system.

spirit of lexical decomposition. Remember that in §2 H was likened to categorizers, and that in §4.1 the ordinary DM categorizers were ruled out. As such, a simple hypothesis about H is that it is a special categorizer.

To identify H, therefore, we need a better understanding of categorizers and their place in the inventory of functional categories. A first point to note is that terms like "categorizer", "categorial feature", and "categoryless" are used loosely in the literature, because if items without a categorizer are categoryless, then categoryless items would include not only Roots, but also T, Asp, Num, etc. In other words, if categorial features (largely limited to [N], [V], [A]) are what define categories (as the term literally suggests), then various functional categories would end up being non-categories. Obviously, these are not what DM is expected to predict; what the above terms really mean are "lexical categorizer", "lexical categorial features", and "lexical-category-less" instead. So, our mission is to identify a special *lexical* category.

Despite their intuitive straightforwardness, lexical categories are a notoriously disputed area in minimalism. As content words are decomposed into categorizers and Roots, the previously held lexical categories become functional in nature. However, "lexical", "noun", "verb", etc. do not follow the nomenclature of functional categories (FF-based, piggybacking on substantive features, cf. note 7) and need to be either renamed or redefined. Two representative approaches exist in this regard. Borer (2005) denies the existence of dedicated categorizers and treats traditional lexical categories as distributional contrasts that are only definable as "categorial complement spaces" of functional projection series, e.g. D-Num-Cl is "nominal" while C-T-Voice is "verbal" (Biberauer 2016b has a similar view). On the other hand, Panagiotidis (2015, 2017) endows the categorial features [N] and [V] with interface substantiveness, letting them represent two "fundamental interpretive perspectives" (FIPs) – "sortality" and "extending into time":

> Sortality will have to be associated with *individuation*, number, quantification etc. — realised as functional categories Number, Determiner etc. "Extending into time" will be the seed of events and causation, and will require event participants, a way to encode length of event and relation between time intervals etc. – realised as an event projection / argument, Voice, Aspect, Tense. (Panagiotidis 2017: lecture 1, p. 4)

The two approaches are not necessarily incompatible. Considering many conventional labels have turned out to be mixtures of heterogeneous concepts (e.g. IP/CP are extended domains, $\text{Merge}_{\text{MP}} = \text{Merge}_{\text{PoP}} + \text{labeling}$[20]), lexical categorial labels like "noun" and "verb" may also have multiple dimensions that could

[20] MP = *Minimalist program* (Chomsky 1995), PoP = *problems of projection* (Chomsky 2013).

(and should) be unbundled. Specifically, we can conceive "noun", "verb", etc. as distributional patterns following Borer while having an FIP-introducing functional layer in each pattern following Panagiotidis. This layer may be identified as the "categorizer" but is not really the original DM categorizer, for it does not turn a Root into a conventional noun/verb but merely turns it into an FIP-bearing item. Other nominal/verbal properties (e.g. referentiality, argument structure) are introduced by additional functional layers in later derivation. Featurally speaking, the FIP-introducer is not so different from other functional heads such as T and Gen in that they are all FF-based[21] and interface-motivated, as in Table 17.1.

Table 17.1: Parallelism between FIP and other functional categories

Category	FIP	T	Asp	Num	Gen
Value	sortal ([N]) ext-in-time ([V]) ?([A])	present past future	perfective progressive habitual	singular plural dual	masculine feminine neuter

Following Adger & Svenonius (2011), a valued feature is a pair of attribute and value ⟨att, val⟩ – or [att:val] in more popular notation – which may be a UG-given template (in the sense of Biberauer 2016b). The attribute is a feature class (i.e. a subset of all features) and the value a feature belonging to that class. Thus, [N] and [V] are more precisely [FIP: sortal/ext-in-time] (henceforth [FIP: N/V] for expository convenience), similar to [T: pres/past]. Adger & Svenonius argue that since the feature classes themselves can be referred to by rules or principles (e.g. agreement copies φ-features), they are grammatically active independently of concrete values. This means there can be valueless attributes – an unsurprising conclusion given the fundamental syntactic role played by unvalued features, or more exactly feature classes (the term "feature" is variably applied to features and feature classes, Adger & Svenonius 2011: 35).

Previous discussions of unvalued features are largely limited to "parasitic" ones, i.e. unvalued features bundled on heads defined by valued features, such as [*u*T] on V and [*u*φ] on T. But in the context of lexical decomposition, there may well be standalone unvalued features making up their own heads.[22] I postulate an unvalued FIP-introducer, consisting of a single [*u*FIP] feature (more vividly [FIP:__]), which declares an FIP interpretation but leaves its value underspecified.

[21]Strictly speaking, there can be no non-FF-based differences among functional heads.

[22]In fact, this is the only possibility if Collins (2017) is on the right track. Unvalued and valued features can still be bundled, but that can only be done in syntax via Merge (cf. §3.1).

Assuming the lexically valued [FIP: N] and [FIP: V] correspond to the ordinary categorizers *n* and *v*, we may call the unvalued FIP-introducer a "defective categorizer" (Cat for short).

4.3 Cat and verb-internal modifier

Cat counts as a non-ordinary lexical categorizer in that it is lexically unvalued. As a result, the Root material it introduces has no concrete FIP interpretation and appears categoryless. This is precisely what we need from H in [H $\sqrt{\text{VIM}}$]-[*v* $\sqrt{\text{VERB}}$], so I identify H as Cat. In this section, I will show how Cat derives VIM.

I adopt the following theoretical assumptions. First, categorizers (however defined) are phase heads (à la Marantz 2001). But unlike Chomskyan *v**P (though maybe like CP), the categorizer phase is spelled out as whole, including both the Root and the categorizer. This is because the Root cannot be properly interpreted without the categorizer. Second, spelled-out constituents do not necessarily vanish from the syntax. Some (e.g. complex "satellites" like specifiers/adjuncts) leave their labels behind as "bookmarks" that behave as terminal nodes (X^0s) for linearization purpose (Nunes & Uriagereka 2000; Fowlie 2013). Third, the bookmarkish "new" lexical items may be derived by spellout plus "renumeration" (Johnson 2003). That is, satellite substructures may be separately derived (perhaps via lexical subarrays in separate workspaces), labeled, and put back in the numeration, so that they can participate in the next cycle of derivation. With these technical devices, we can now derive modificational compound verbs.

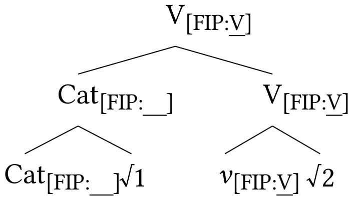
To begin with, Cat and *v* separately categorize a Root. Since the Roots are not lexically marked as VIM or V, I simply write them numerically as √1 and √2.

(12)
  a. Select Cat and √1 into a lexical subarray LA_i.
  b. Merge Cat and √1. LA_i is exhausted. Transfer.
  c. Since the Root is FF-less, Cat labels {Cat, √1} as Cat (featurally [*u*FIP]).
  d. Renumerate the Root-supported Cat (notated as $Cat_{\surd}$).
  e. Repeat steps a-d for *v* and √2.

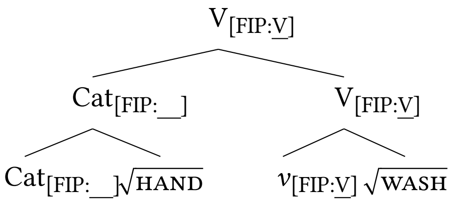
After (12), the numeration contains the two "recycled" lexical items $Cat_{\surd}$ and $V_{\surd}$. This is the end of word-internal derivation and the beginning of the Chomskyan derivation, where lexical items are equipped with categorial information.

Then, $Cat_{\surd}$ and $V_{\surd}$ are selected into another lexical subarray LA_j together with other *v**P-phase items and merged accordingly. Upon the next Transfer, the unvalued FIP feature on Cat probes for a value and finds one on V. It is thus valued via Agree, and the Cat$_{\surd}$-V$_{\surd}$ merger is labeled as V by feature sharing (Chomsky 2013), as in (13a).[23] See (13b) for a concrete example.[24]

(13) a.



b.



Suppose the system can distinguish intrinsically valued features from features valued via Agree,[25] there would be a derivational asymmetry between Cat$_{\surd}$ and V$_{\surd}$, with the former's interpretation depending on the latter's. This dependency may be reflected in semantics as variable sharing, which I briefly illustrate below.

Under the bare Root view, I assume that the denotation of a Root is radically underspecified, to the extent that it is not only grammatically void, but also does not make a complete function. Instead, a Root merely denotes a vague property – a "function template" whose domain (including variable type) is not yet defined, as in (14a). This information is only added when the Root is categorized, as in (14b).

(14) a. $[\![\sqrt{\text{WASH}}]\!] = \lambda$__ .WASH(__)
= 'encyclopedically related to wash and compositionally __'

b. $[\![[v \sqrt{\text{WASH}}]]\!] = \lambda e.\text{WASH}(e)$
= 'encyclopedically related to wash and compositionally an extending-into-time FIP (i.e. an event)'

[23] I remain agnostic as to whether feature sharing in labeling is the same mechanism as that in agreement as proposed in i.a. Frampton & Gutmann (2000, 2006) and Haug & Nikitina (2016).

[24] I assume the pairing of Roots and categorizers to be a matter of pre-linguistic planning. As Chomsky (1995: 227) remarks, there is "no meaningful question as to why one numeration is formed rather than another". What matters here is merely that each LA only contain one Root.

[25] I leave aside the technical details, but any adequate theory would be compatible. See Rooryck & Vanden Wyngaerd (2011: 10) for a proposal based on feature sharing.

The event variable *e* in (14b), which defines eventuality, is introduced by the verbalizer (cf. Marantz 2013). Since the verbalizer is featurally [FIP:V], *e* is presumably encoded in the value [V]. More generally, I assume all variable types to be functionally introduced rather than an inherent part of the Root. Being valueless, Cat does not introduce any variable type, though it does endow the Root with interface interpretability (as an FIP).[26] So, $\text{Cat}_{\surd}$ has the denotation in (15).

(15) a. ⟦[Cat √1]⟧ = λ⟦[FIP:__]⟧.1(⟦[FIP:__]⟧)
= ‘encyclopedically related to 1 and compositionally a __ FIP’

b. ⟦[Cat √HAND]⟧ = λ⟦[FIP:__]⟧.HAND(⟦[FIP:__]⟧)
= ‘encyclopedically related to hand and compositionally a __ FIP’

After Agree, $\text{Cat}_{\surd}$ is equipped with the event variable introduced by *v*. However, since the categorial interpretation of √1 has been fixed in the previous spell-out cycle, the newly obtained *e* can no longer turn √1 into an independent event, but only connects it to another event, i.e. that denoted by $\text{V}_{\surd}$. As such, √1 effectively becomes a modifier of $\text{V}_{\surd}$, as in (16).

(16) a. ⟦(13a)⟧ = λ<u>e</u>.1(<u>e</u>) ∧ λ*e* .2(*e*)
= ‘encyclopedically related to 1 and compositionally connected to an event’ ∧ ‘encyclopedically related to 2 and compositionally an event’
= ‘an event of 2, encyclopedically related to 1’ (event identification)

b. ⟦(13b)⟧ = λ<u>e</u>.HAND(<u>e</u>) ∧ λ*e* .WASH(*e*)
= ‘encyclopedically related to hand and compositionally connected to an event’ ∧ ‘encyclopedically related to wash and compositionally an event’ = ‘an event of washing, encyclopedically related to hand’

Since {$\text{Cat}_{\surd}$, $\text{V}_{\surd}$} and $\text{V}_{\surd}$ have identical labels, $\text{Cat}_{\surd}$ is in effect an adjunct. Since $\text{Cat}_{\surd}$ is dominated by V, it is verb-internal. The modificational compound is thus derived solely by Simplest Merge and labeling, with no need of Pair Merge, Root incorporation, postsyntactic operation or multidominance. In effect, the structure in (13) unifies two Roots under one ordinary categorizer without violating the DM tenet that one categorizer can only categorize one Root (cf. Embick 2010).

[26] A consequence of the single engine hypothesis is that unvalued features must not be deleted by the end of the categorizer phase (i.e. when the categorized Roots are renumerated), because they are still required in the Chomskyan numeration and the next cycle of derivation. I merely acknowledge this point but do not attempt to account for it in this study.

5 Some implications

5.1 Noun-internal modifiers

In §4, I illustrated how VIMs are derived by Cat, but the application of the defective categorizer hypothesis is not confined to the verbal domain. In fact, since all Cat$_{\surd}$ needs is an FIP value, it may well be merged with a noun and become a noun-internal modifier (NIM). While leaving NIMs to future research, in (17) I illustrate the flexibility of Cat by items that can be used as both VIM and NIM.

(17) a. English
hand$_{\text{[FIP:}\underline{\text{V}}\text{]}}$-wash$_{\text{[FIP:V]}}$ vs. hand$_{\text{[FIP:}\underline{\text{N}}\text{]}}$-gel$_{\text{[FIP:N]}}$,
sleep $_{\text{[FIP:}\underline{\text{V}}\text{]}}$-walk$_{\text{[FIP:V]}}$ vs. sleep$_{\text{[FIP:}\underline{\text{N}}\text{]}}$-mode$_{\text{[FIP:N]}}$,
breast$_{\text{[FIP:}\underline{\text{V}}\text{]}}$-feed$_{\text{[FIP:V]}}$ vs. breast$_{\text{[FIP:}\underline{\text{N}}\text{]}}$-bone$_{\text{[FIP:N]}}$

b. Chinese
shǒu$_{\text{[FIP:}\underline{\text{V}}\text{]}}$-xǐ$_{\text{[FIP:V]}}$ ‘hand-wash; to handwash’ vs.
shǒu$_{\text{[FIP:}\underline{\text{N}}\text{]}}$-jī$_{\text{[FIP:N]}}$ ‘hand-machine; mobile phone’,
xīn$_{\text{[FIP:}\underline{\text{V}}\text{]}}$-suàn$_{\text{[FIP:V]}}$ ‘heart-calculate; to do mental calculation’ vs.
xīn$_{\text{[FIP:}\underline{\text{N}}\text{]}}$-lì$_{\text{[FIP:N]}}$ ‘heart-force; mental efforts’

c. Japanese
se$_{\text{[FIP:}\underline{\text{V}}\text{]}}$-ou$_{\text{[FIP:V]}}$ ‘back-carry; to carry on back’ vs.
se$_{\text{[FIP:}\underline{\text{N}}\text{]}}$-bone$_{\text{[FIP:N]}}$ ‘back-bone’,
oshi$_{\text{[FIP:}\underline{\text{V}}\text{]}}$-taosu$_{\text{[FIP:V]}}$ ‘push-topple; to push down’ vs.
oshi$_{\text{[FIP:}\underline{\text{N}}\text{]}}$-bana$_{\text{[FIP:N]}}$ ‘push-flower; pressed flower’

The Cat-licensed Roots $\sqrt{\text{HAND}}$, $\sqrt{\text{SLEEP}}$, $\sqrt{\text{BREAST}}$, etc. have no fixed FIP interpretation – they become VIMs when merging with V$_{\surd}$s and NIM when merging with N$_{\surd}$s. Admittedly, whether or not a specific Cat-item has both verbal and nominal uses is a matter of language-specific lexicalization, e.g. while all of *hand-wash*, *hand-gel*, and *foot-gel* are fine in English, there is no ?*foot-wash* (‘wash with foot’) by the time this chapter is written (though it could easily be coined). The defective categorizer hypothesis does not aim to predict which VIMs/NIMs actually exist in a certain language, but merely captures the capacity of human beings to create such language units.

5.2 Universality of compounding

The proposed theory can not only be extended to the nominal domain but also predict the widespreadness of modificational compounds. Given the mutual dependence between feature classes (attributes) and features (values), the FIP class (i.e. the set of FIP values) should be as widespread as its values. Moreover, if "no value" can be conceived as the empty set, i.e. [F:__] = [F: ∅], then unvalued features are in effect free-riders of their valued counterparts, for the empty set is a subset member of all sets, which include all feature classes (conceived as subsets of all features, cf. §3.2). This means any language with at least one FIP value also has a grammatically active defective categorizer. In other words, modificational compounding as a generative mechanism is as widespread in human languages as conventional lexical categories, i.e. universal (cf. Baker 2003; Panagiotidis 2015).

This conclusion is supported by typological studies. According to Bauer (2009: 344), (modificational) compounding has been suggested to be a language universal (Fromkin et al. 1996: 54–55; Libben 2006: 2), as evidenced by language acquisition (Clark 1993) and contact (Plag 2006). A caveat here is that universality may be masked by varied terminology and classification in descriptive grammars. For example, descriptions of Ainu (e.g. Refsing 1986; Shibatani 1990) do not mention compounding at all, though the language does have de facto compounds, as in (18a). Similarly, Evenki has also been claimed to lack compounds (Nedjalkov 1997: 308), but a quick look into alternative sources reveals many of them, as in (18b).

(18) a. Ainu (language isolate; via Bauer 2009)
atuy asam 'bottom sea; sea bottom', kamuy napuri 'mountain god; holy mountain', supuya kur 'trace smoke; smoke trace'

b. Evenki (Tungusic; cf. Hu & Chao 1986)
eyji shee 'brick tea', aaxin jolo 'liver stone; marble', unaaji ute 'girl son; daughter'

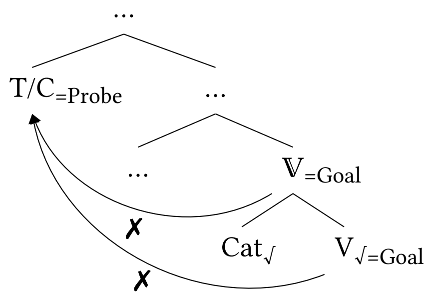
5.3 Compound verb typology

Despite the universality of modificational compounds, compound nouns are cross-linguistically a lot more common than compound verbs. Take the familiar European languages for example: while modificational compound nouns exist in all of Germanic, Romance, and Slavic languages (cf. Bauer 2009), compound verbs like *hand-wash* are only seen in English with some productivity. One might take this to be an areal phenomenon, for compound verbs are more widely used in e.g. East Asia. However, as Bauer (2009: 355) comments, the areal preferences are not clearly correlated "with anything linguistic in the appropriate languages".

The defective categorizer hypothesis provides a new perspective on modeling this unbalanced typology. Since the node dominating [$\text{Cat}_{\surd}$ $\text{V}_{\surd}$] (call it $\mathbb{V}$[27]) has

[27] Similar to Booij's (1990) V*, which is more than V^0 but less than V′ (cf. Vikner 2005).

exactly the same label as the $V_{\surd}$ node, operations targeting one node also target the other. As a result, in languages requiring V-to-T/C movement, the T/C probe is unable to access the real V^0 (i.e. $V_{\surd}$, which becomes a terminal lexical item after renumeration, cf. §4.3) due to the intervening 𝕍, as in (19). This is presumably a minimal search effect, as formulated in the minimal link condition (20).

(19)


(20) Minimal link condition (Chomsky 1995: 311):
K attracts α only if there is no β, β closer to K than α, such that K attracts β.

In addition, since 𝕍 is not a minimal category (head) on the clausal spine, it cannot undergo head movement either (see 19). Therefore, in the end nothing moves to T/C, and the derivation crashes. This means Cat–V compound verbs are only well-formed in languages/contexts without verb movement requirement. So Romance languages, where V systematically moves to T (cf. Biberauer & Roberts 2010), are incompatible with such compound verbs. For instance, the concepts in (1) are expressed periphrastically in Spanish, as in (21).[28]

(21)

English	Spanish
double-check	*volver a revisar* 'to inspect again'
dry-clean	*limpiar en seco* 'to clean in dry'
hand-wash	*lavar a mano* 'to wash by hand'
sleep-walk	*caminar dormido* 'to walk asleep'
window-shop	*mirar escaparates* 'to look at shop windows'
baby-sit	*hacer de canguro* 'to do kangaroo'
hitch-hike	*hacer autoestop* 'to do car-stop'…

However, the prediction as such is too strong, for apart from V-to-T/C, there is also V-to-*v** (or more generally V-to-VP-shell) movement, e.g. in English (cf.

[28]Retrieved from oxforddictionaries.com and wordreference.com (29 Dec 2017).

Roberts 2010, 2019). So, if Cat–V compounds and verb movement are totally complementary, then English becomes a major counterexample.

One possible solution lies in the design of Cat. Since it merely needs to merge with something that can provide it with an FIP value (and thus label the merger), which in the case of [V] is essentially an event variable, it can in theory merge with any *e*-equipped head. In a neo-constructionist event structure (cf. Acedo-Matellán 2016), this may be any subevental head (e.g. Init/Proc/Res in Ramchand 2008) or argument introducing head (e.g. Voice/Appl in Pylkkänen 2008). Considering Internal Merge occurs at phase level (cf. Citko 2014), i.e. after all steps of External Merge in a phase are done, and the $\text{Cat}_{\surd}$-$\text{V}_{\surd}$ merger is External Merge, here I make the conservative hypothesis that apart from the verbalizer, the next position Cat may attach to is the v^* phase head (whichever head that turns out to be in an elaborate verbal domain). Crucially, since Cat only merges in after all steps of Internal Merge in v^*P are done (i.e. as part of the next phase), Cat–V (more exactly Cat-v^*) compounds may well exist in a language with V-to-v^* movement. In sum, we can have a three-way typology of Cat–V compounds (and VIMs) regulated by the verb movement parameter, as in Table 17.2.

Table 17.2: Typology of Cat–V compound verbs

Type	Example	V-to-T/C movement	Cat–V compound
I	Romance	Yes	No
II	OV-Germanic	Main clause: Yes	No
II	OV-Germanic	Embedded clause: No	Yes
III	English, Chinese	No	Yes

Note that due to the inconsistent verb movement requirement, OV-Germanic languages may only have Cat–V compounds in non-V2 contexts, as in (22).

(22) German (via Vikner 2005)

a. * Spart er bau? / * Bau-spart er?
saves he building building-saves he
(‘Does he building-save?’)

b. Er will bau-sparen. / … weil er bau-spart.
he wants building-save.INF because he building-saves
‘He wants to building-save./…because he building-saves.’

The compound verb *bau-sparen* ‘building-save; to building-save’ cannot appear in finite main clauses but is only well-formed in situ, either in a sentence with a modal verb (which fulfills the V2 requirement) or in a subordinate clause (where there is no V2 requirement). Germanic compounds like *bau-sparen* are known as "immobile verbs" (cf. i.a. McIntyre 2002; Vikner 2005; Ahlers 2010; Song 2016). They have a natural explanation in the current model.

As a final remark, the typology in Table 17.2 only concerns Cat–V compounds. So, on the one hand, Type I–II languages may still have unhindered Cat-N compounds/NIMs, e.g. French *homme grenouille* 'man-frog; frogman', Spanish *bocacalle* 'mouth-street; street intersection'. On the other hand, they may also have other types of complex verb in all contexts, such as particle verbs (including their inseparable variants), e.g. German *ein-kaufen* 'in-buy; to shop' (V-PP), *er-warten* 'ER-wait; to expect' (*er*<OHG *ur* 'out'), Spanish *ex-traer* 'out-pull; to extract', and various phrasal verbs, e.g. French *mettre bas* 'put low; to give birth' (V+AP), Spanish *ponerse en camino* 'put.REFL on way; to set off' (V+clitic+PP), German *Schwein haben* 'pig have; to be lucky' (V+NP). I do not discuss these other types of complex verb (more exactly complex predicate) but merely distinguish them from Cat–V compounds. To wit, items like *ein*, *bas*, and *Schwein* are base-generated as V-complements, i.e. VMs in the broad sense (cf. §1), but they are not VIMs.

6 Conclusion

This chapter is a minimalist study of verb-internal modifiers (non-heads of modificational compound verbs). I have defended the position that compounding is a syntactic phenomenon based on the view that syntax is the only generative engine in the human language faculty. My main difference from previous syntactic models of compounding is that I have kept to the simplest definition of Merge (no Pair Merge or remerge) and the bare Root view (no RootP, Root-Root merger or Root incorporation), both of which are consequences of the SMT. Guided by the defective categorizer hypothesis, which is independently motivated in the minimalist feature system, I have derived VIMs in a labeling-based model. This new model not only avoids the conceptual problems in previous approaches, but also brings along a number of potential points of future research. First, it can be extended to the nominal domain and allows the same Root material to modify both verbs and nouns. Second, it predicts modificational compounding to be a language universal and relates the typology of Cat–V compounds to the verb movement parameter. In addition, beyond the verbalizer level, there may be further loci that Cat can attach to, e.g. the v^* phase head. As such, compounding is not only a natural part of syntax, but also sheds new light on "external" syntactic issues such as general head adjunction and phase-level modifiers.

Abbreviations

3	third person	LF	logical form
BCC	Borer–Chomsky conjecture	NIM	noun-internal modifier
CLF	classifier	PRS	present
DISP	disposal	SG	singular
DM	Distributed Morphology	SMT	strong Minimalist thesis
FIP	fundamental interpretive perspective	SO	syntactic object
		UG	Universal Grammar
INF	infinitive	VIM	verb-internal modifier
LAC	label accessibility condition	VM	verbal modifier

Acknowledgements

This study forms part of my PhD project “Flexibility of syntactic categories: A cross-linguistic study” (funded by Cambridge Trust and China Scholarship Council). I am grateful to Ian Roberts, Theresa Biberauer, Anders Holmberg and Víctor Acedo-Matellán for discussions and encouragement. Thanks to the anonymous reviewer for helpful feedback, and to the audience at the Cambridge SyntaxLab (2 February 2016, 7 February 2017), the “Linguistic variation in the interaction between internal and external syntax” workshop (Utrecht, 8–9 Feb 2016), and the graduate conference WoSSP13 (Barcelona, 30 June–1 July 2016) for questions and comments on earlier versions of this work. All remaining errors are my own.

References

Acedo-Matellán, Víctor. 2016. *The morphosyntax of transitions*. Oxford: Oxford University Press.

Acquaviva, Paolo. 2009. Roots and lexicality in Distributed Morphology. In Alexandra Galani, Daniel Redinger & Norman Yeo (eds.), *YPL2 — Issue 10 (May 2009) Special Issue — York–Essex morphology meeting*. University of York: Department of Language & Linguistic Science.

Adger, David & Peter Svenonius. 2011. Features in minimalist syntax. In Cedric Boeckx (ed.), *The Oxford handbook of linguistic minimalism*, 27–51. Oxford: Oxford University Press.

Ahlers, Timo. 2010. Komplexe C^0-phobe Verben des Deutschen [Complex C^0-phobic verbs of German]. MPhil thesis, University of Vienna.

Alexiadou, Artemis. 2014. Roots don’t take complements. *Theoretical Linguistics* 40(3/4). 287–297. DOI: 10.1515/tl-2014-0012.

Arsenijević, Boban & Joanna U.-S. Sio. 2009. Mediating merge: Evidence from the Cantonese *ge*. Paper presented at GLOW in Asia VII, February 24–27, 2009. Hyderabad. The English and Foreign Languages University.

Baker, Mark C. 2003. *Lexical categories: Verbs, nouns and adjectives*. Cambridge: Cambridge University Press.

Baker, Mark C. 2008. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.

Bauer, Laurie. 2009. Typology of compounds. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 343–357. Oxford: Oxford University Press.

Bauke, Leah S. 2014. *Symmetry breaking in syntax and the lexicon*. Amsterdam: John Benjamins. DOI: 10.1075/la.216.

Bauke, Leah S. 2016. Parametric variation in nominal root compounding. In Luis Eguren, Olga Fernández-Soriano & Amaya Mendikoetxea (eds.), *Rethinking parameters*, 203–235. Oxford: Oxford University Press.

Bauke, Leah S. & Thomas Roeper. 2017. How unlabelled nodes work: Morphological derivations and the subcomponents of UG operations. In Leah S. Bauke & Andreas Blümel (eds.), *Labels and roots*, 233–262. Berlin: De Gruyter. DOI: 10.1515/9781501502118-010.

Berwick, Robert C. & Noam Chomsky. 2016. *Why only us: Language and evolution*. Cambridge, MA: MIT Press.

Biberauer, Theresa. 2016a. Going beyond the input: 3 factors and syntactic innovation. Paper presented at the 2016 LAGB Annual Meeting (York).

Biberauer, Theresa. 2016b. Particles and the final-over-final constraint. Handout, talk given at *Linguistic Variation in the Interaction between Internal and External Syntax*, Utrecht, 8–9 February 2016.

Biberauer, Theresa. 2017. Factors 2 and 3: A principled approach. *Cambridge Occasional Papers in Linguistics* 10. 38–65.

Biberauer, Theresa & Ian Roberts. 2010. Subjects, tense and verb-movement. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 263–303. Cambridge: Cambridge University Press.

Boeckx, Cedric. 2014. *Elementary syntactic structures: Prospects of a feature-free syntax*. Cambridge: Cambridge University Press.

Booij, Geert. 1990. The boundary between morphology and syntax: Separable complex verbs in Dutch. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 1990*, 45–63. Foris: Dordrecht.

Borer, Hagit. 2005. *The normal course of events* (Structuring sense 2). Oxford: Oxford University Press.

Borer, Hagit. 2009. Roots and categories. Talk at the 19th Colloquium on Generative Grammar. University of the Basque Country.
Borer, Hagit. 2013. *Taking form* (Structuring sense 3). Oxford: Oxford University Press.
Borer, Hagit. 2014. Wherefore roots? *Theoretical Linguistics* (40). 343–359. DOI: 10.1515/tl-2014-0016.
Bošković, Željko. 2016. Introduction. *The Linguistic Review* 33. 1–16. DOI: 10.1515/tlr-2015-0012.
Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, MA: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.
Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and beyond* (The cartography of syntactic structures 3), 104–131. Oxford: Oxford University Press.
Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22. DOI: 10.1162/0024389052993655.
Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. DOI: 10.1016/j.lingua.2012.12.003.
Chomsky, Noam. 2015. Problems of projection: Extensions. In Elisa Di Domenico, Cornelia Hamann & Simona Matteini (eds.), *Structures, strategies and beyond: Studies in honour of Adriana Belletti*, 1–16. Amsterdam: John Benjamins. DOI: 10.1075/la.223.01cho.
Citko, Barbara. 2014. *Phase theory: An introduction*. Cambridge: Cambridge University Press.
Clark, Eve V. 1993. *The lexicon in acquisition*. Cambridge: Cambridge University Press.
Collins, Chris. 2002. Eliminating labels. In Samuel Epstein & T. Daniel Seely (eds.), *Derivation and explanation in the minimalist program*, 42–64. Oxford: Blackwell.
Collins, Chris. 2017. Merge(X,Y) = {X,Y}. In Leah S. Bauke & Andreas Blümel (eds.), *Labels and roots*, 47–68. Berlin: De Gruyter. DOI: 10.1515/9781501502118-003.
Cuervo, Ana Maria. 2014. Arguments for a root. *Theoretical Linguistics* 40(3/4). 375–387. DOI: 10.1515/tl-2014-0018.
de Belder, Marijke. 2011. *Roots and affixes: Eliminating lexical categories from syntax*. HU Brussel/Utrecht University. (Doctoral dissertation).
de Belder, Marijke. 2017. The root and nothing but the root: Primary compounds in Dutch. *Syntax* 20(2). 138–169. DOI: 10.1111/synt.12133.

de Belder, Marijke & Marjo van Koppen. 2016. One module, different levels of merge: AN(N) compounds in Dutch. *Studia Linguistica* 70(1). 1–33. DOI: 10.1111/stul.12040.

Dehé, Nicole, Ray Jackendoff, Andrew McIntyre & Silke Urban. 2002. Introduction. In Nicole Dehé, Ray Jackendoff, Andrew McIntyre & Silke Urban (eds.), *Verb-particle explorations*, 1–20. New York: De Gruyter. DOI: 10.1515/9783110902341.1.

É. Kiss, Katalin. 1987. *Configurationality in Hungarian*. Dordrecht: Reidel.

É. Kiss, Katalin. 2002. *The syntax of Hungarian*. Cambridge: Cambridge University Press.

É. Kiss, Katalin. 2008. The function and the syntax of the verbal particle. In Katalin É. Kiss (ed.), *Event structure and the left periphery*, 17–56. Dordrecht: Springer.

Embick, David. 2010. *Localism versus globalism in morphology and phonology*. Cambridge MA: MIT Press.

Embick, David & Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry* 39(1). 1–53. DOI: 10.1162/ling.2008.39.1.1.

Epstein, Samuel David, Hisatsugu Kitahara & T. Daniel Seely. 2012. Structure building that can't be. In Myriam Uribe-Etxebarria & Vidal Valmala (eds.), *Ways of structure building*, 253–270. Oxford: Oxford University Press.

Feng, Shengli. 1997. *Interactions between morphology, syntax and prosody in Chinese*. Beijing: Peking University Press.

Feng, Shengli. 2015. Light-verb syntax between English and Classical Chinese. In Audrey Li, Andrew Simpson & Wei-Tien Dylan Tsai (eds.), *Chinese syntax in a cross-linguistic perspective*, 229–250. Oxford: Oxford University Press.

Fowlie, Meaghan. 2013. Multiple multiple spellout. In Theresa Biberauer & Ian Roberts (eds.), *Challenges to linearization*, 129–169. Berlin: De Gruyter. DOI: 10.1515/9781614512431.129.

Frampton, John & Sam Gutmann. 2000. Agreement is feature sharing. Ms., Northeastern University.

Frampton, John & Sam Gutmann. 2006. How sentences grow in the mind: Agreement and selection in an efficient minimalist syntax. In Cedric Boeckx (ed.), *Agreement systems*, 121–158. Amsterdam: John Benjamins. DOI: 10.1075/la.92.08fra.

Fromkin, Victoria, Robert Rodman, Peter Collins & David Blair. 1996. *An introduction to language*. Third edition. Sydney: Harcourt Brace.

Gallego, Ángel J. 2014. Roots and phases. In Artemis Alexiadou, Hagit Borer & Florian Schäfer (eds.), *The syntax of roots and the roots of syntax*, 192–207. Oxford University Press.

Haiden, Martin. 2006. Verb particle constructions. In Martin Everaert & Henk C. van Riemsdijk (eds.), *The Blackwell companion to syntax*, vol. I, 344–375. Oxford: Blackwell. DOI: 10.1002/9780470996591.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Ken Hale & Samuel Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.

Harley, Heidi. 2009. Compounding in distributed morphology. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 129–145. OUP.

Harley, Heidi. 2014. On the identity of roots. *Theoretical Linguistics* 40(3/4). 225–276. DOI: 10.1515/tl-2014-0010.

Haug, Dag Trygve Truslew & Tatiana Nikitina. 2016. Feature sharing in agreement. *Natural Language & Linguistic Theory* 34(3). 865–910. DOI: 10.1007/s11049-015-9321-9.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298(5598). 1569–1579. DOI: 10.1126/science.298.5598.1569.

Hegedűs, Veronika. 2013. *Non-verbal predicates and predicate movement in Hungarian*. Utrecht: LOT.

Hornstein, Norbert. 2009. *A theory of syntax: Minimal operations and Universal Grammar*. Cambridge: Cambridge University Press.

Hu, Wei. 2013. *A distributed morphological study of compounding in English and Chinese*. Nankai University. (Doctoral dissertation).

Hu, Zengyi & Ke Chao. 1986. *A description of the Evenki language*. Beijing: Publishing House of Minority Nationalities.

Huang, C.-T. James. 2015. On syntactic analyticity and parametric theory. In Audrey Li, Andrew Simpson & Wei-Tien Dylan Tsai (eds.), *Chinese syntax in a cross-linguistic perspective*, 1–48. Oxford: Oxford University Press.

Huang, C.-T. James, Yen-hui Audrey Li & Yafei Li. 2009. *The syntax of Chinese*. Cambridge: Cambridge University Press.

Johnson, Kyle. 2003. Towards an etiology of adjunct islands. *Nordlyd* 31(1). 187–215. DOI: 10.7557/12.25.

Kageyama, Taro. 1993. *Bumpoo to gokeesee*. [Grammar and Word Formation]. Tokyo: Hituzi Syobo.

Kayne, Richard S. 1985. Principles of particle constructions. In Jacqueline Guéron, Hans-Georg Obenauer & Jean-Yves Pollock (eds.), *Grammatical representation*, 101–140. Dordrecht: Foris.

Libben, Gary. 2006. Why study compound processing? An overview of the issues. In Gary Libben & Gonia Jarema (eds.), *The representation and processing of compound words*, 1–22. Oxford: Oxford University Press.

Lohndal, Terje. 2014. *Phrase structure and argument structure: A case study of the syntax–semantics interface*. Oxford: Oxford University Press.

Marantz, Alec. 1997. No escape from syntax: Don’t try morphological analysis in the privacy of your own lexicon. *UPenn Working Papers in Linguistics* 4. 201–225.

Marantz, Alec. 2001. Words. Paper presented at the 20th West Coast Conference on Formal Linguistics, 23–25 February 2001, University of Southern California, Los Angeles.

Marantz, Alec. 2013. Verbal argument structure: Events and participants. *Lingua* 130. 152–168. DOI: 10.1016/j.lingua.2012.10.012.

McIntyre, Andrew. 2002. Verb-second and backformations and scalar prefix verbs in German: The interaction between morphology, syntax and phonology. Ms., University of Leipzig.

Nedjalkov, Igor. 1997. *Evenki*. London: Routledge.

Nishiyama, Kunio & Yoshiki Ogawa. 2014. Auxiliation, atransitivity, and transitivity harmony in Japanese VV compounds. *Interdisciplinary Information Sciences* 20(2). 71–101. DOI: 10.4036/iis.2014.71.

Nunes, Jairo & Juan Uriagereka. 2000. Cyclicity and extraction domains. *Syntax* 3(1). 20–43. DOI: 10.1111/1467-9612.00023.

Oseki, Yohei. 2015. Eliminating pair-merge. In Ulrike Steindl, Thomas Borer, Huilin Fang, Alfredo García Pardo, Peter Guekguezian, Brian Hsu, Charlie O’Hara & Iris Chuoying Ouyang (eds.), *Proceedings of the 32nd West Coast Conference on Formal Linguistics*, 303–312. Somerville, MA: Cascadilla Proceedings Project.

Panagiotidis, Phoevos. 2015. *Categorial features: A generative theory of word class categories*. Cambridge: Cambridge University Press.

Panagiotidis, Phoevos. 2017. Categorial elements: Roots, adjectives, and mixed projections. Mini-seminar at University of Cambridge, 16-17 May 2017.

Plag, Ingo. 2006. Morphology in pidgins and creoles. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, Second edition, 304–308. Oxford: Elsevier.

Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge MA: MIT Press.

Ramchand, Gillian. 2008. *Verb meaning and the lexicon: A first phase syntax*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511486319.

Refsing, Kirsten. 1986. *The Ainu language*. Aarhus: Aarhus University Press.

Richards, Norvin. 2010. *Uttering trees*. Cambridge, MA: MIT Press.

Roberts, Ian. 2010. *Agreement and head movement: Clitics, incorporation, and defective goals*. Cambridge, MA: MIT Press.
Roberts, Ian. 2019. *Parameter hierarchies and universal grammar*. Oxford: Oxford University Press.
Roberts, Ian & Anna Roussou. 2003. *Syntactic change: A minimalist approach to grammaticalization*. Cambridge: Cambridge University Press.
Rooryck, Johan & Guido Vanden Wyngaerd. 2011. *Dissolving binding theory*. Oxford: Oxford University Press.
Rubin, Edward J. 2003. Determining pair-merge. *Linguistic Inquiry* 34(4). 660–668. DOI: 10.1162/ling.2003.34.4.660.
Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge: Cambridge University Press.
Song, Chenchen. 2016. Blocked particle verb movement in German and Hungarian: A unified analysis. Paper presented at the annual meeting of the Linguistics Association of Great Britain, York, 6–9 September 2016.
Song, Chenchen. 2017a. Revisiting Chinese P: Is it a necessary category? Presentation at the Morphosyntactic Variation in Adpositions Workshop, Cambridge, 8-9 May.
Song, Chenchen. 2017b. Prosodically-driven morphosyntactic change? Revisiting the history of Chinese disyllabic words. *Cambridge Occasional Papers in Linguistics* 10. 152–196.
Song, Chenchen. 2017c. No root-root merger in Chinese compounds. Ms. University of Cambridge.
Taraldsen, Tarald. 1983. *Parametric variation in phrase structure*. University of Tromsø. (Doctoral dissertation).
Tsujimura, Natsuko. 2013. *An introduction to Japanese linguistics*. Third edition. John Wiley & Sons.
Vikner, Sten. 2005. Immobile complex verbs in Germanic. *The Journal of Comparative Germanic Linguistics* 8(1). 85–117. DOI: 10.1007/s10828-004-0726-9.
Zeijlstra, Hedde. 2008. On the syntactic flexibility of formal features. In Theresa Biberauer (ed.), *The limits of syntactic variation*, 143–174. Amsterdam: John Benjamins. DOI: 10.1075/la.132.06zei.
Zhang, Niina Ning. 2007. Root merger in Chinese compounds. *Studia Linguistica* 61(2). 170–184. DOI: 10.1111/j.1467-9582.2007.00131.x.