To Appear in M. Becker, J. Grinstead, J. Rothman &B.D. Schwartz (eds.) *Generative Linguistics and Acquisition: Studies in Honor of Nina M. Hyams*. Philadelphia: Benjamins

The Syntactic Domain of Content[1]

Hagit Borer
Queen Mary University of London and USC

Abstract

A main motivation for relegating Word Formation to the lexicon is the fact that its output is often non-compositional. The present article, however, presentsa serious challenge to the presumed contradiction between non-compositionality and syntactic combinatorial processes. The investigation of N-N Constructs in Hebrew shows that equal syntactically complex expressions nonetheless interact differently with non-compositionality. Crucially, it is the *syntactic* differences between these expressions that give rise to distinct Content properties, with non-compositionality correlating not with syntactic structure as such, but with the presence of functional structure.

The emerging syntactic domain of 'word' Content in turn allows the language learner to make informed decisions on where to look for non-compositionality and to draw the appropriate structural conclusions from its presence.

1. Introduction

At the core of any lexicalist approach stands the notion of a 'word' as a listed item. More specifically, lexicalist approaches typically partition the domain of rule application to that which involves the syntax, and which displays canonical syntactic properties, and that which involves lexical information, specifically as associated with listed units. The motivation for such lexical operations tends to cluster into two rather conceptually distinct types. At one end of the spectrum there are operations which are presumed lexical because they are delimited by properties which are item-specific, i.e. properties that do not generalize and are hence "exceptions", for instance, English dative shift which affects *give* but not *donate*. Following a similar logic, the lexicon is the home of operations which give rise to item-specific properties, e.g. non-compositional Content, e.g. *transformation* in its technical linguistic sense or *transmission*.[2] At the other end of the spectrum we find a formal motivation based on syntactic restrictions. Thus, for instance, it is generally assumed that the syntax is prevented from eliminating argument positions otherwise lexically specified. The elimination of arguments, if needed, thus cannot be syntactic, but may be stated as an operation on a lexical entry. Similarly, one might assume that syntax trades exclusively in (maximal) phrases and that a separate component is thus required for handling the combinatorial properties of terminals (see, for instance, Ackema and Neeleman, 2004 for the latter motivation).

These different motivations notwithstanding, they are linked by one extremely important commonality. All are committed to the existence of listed (substantive) units, call them 'words', which

[1] This paper benefits from many useful comments made by audiences at the University of Toronto, at Utrecht University and at The University of Southern California. Special thanks to the editors and to two anonymous reviewers.

[2] The term Content, roughly correlating to (one interpretation of) Frege's *Sinn*, is in reference to the meaning of substantive vocabulary items, which I assume to be fundamentally based on world knowledge and the conceptual system. *Semantics*, when the term is used for terminals, is in reference to rigidly designating functions, such as those which typify functional vocabulary as well as, e.g. cardinals or quantifiers.

constitute individual, syntactically atomic packets of morphological, syntactic, and phonological instructions to the grammar. What, however, is a 'word', or more specifically, how can we determine what the reservoir of basic listed items consists of? From a syntactic or semantic perspective, we note, the issue is wide open. There is little a priori syntactic or semantic reason to assume that e.g. *the doctor* is two words, but *Mary* is one, or that *postman* is one word, but *postal worker* is two. A more coherent notion typically comes from phonology (e.g. a phonological word is a well-defined domain for the application of a specific phonological rule, e.g., primary stress assignment), but why should such a phonological domain constitute a privileged unit from the perspective of the syntax or the semantics?

In turn, and as is well known, syntactic, morphological and phonological properties do not always go hand in hand. Causative constructions may include two morpho-phonological separable units, (English, Romance), or one (Japanese, Turkish), without any syntactic or semantic difference resulting. The English verb *whiten*, morpho-phonologically derived, corresponds to two distinct syntactic structures of unequal complexity, one inchoative and the other causative. The verb *cool*, morpho-phonologically underived corresponds to those very same two syntactic structures. In all of these cases, what is a word, how complex it is, or how many of it there are in any give string seems orthogonal to syntactic structure, syntactic complexity or semantic interpretation.

The question is particularly salient from the perspective of language acquisition. According to some acquisition models (e.g. the Semantic Bootstrapping approach as in Grimshaw 1981 or Pinker 1984), the pairing of meaning and structure linked to listed items is a crucial milestone of acquisition. Specifically, and given the fundamental assumption that a word such as *kick* is a cluster of phonological, semantic and syntactic information; once the child has access to the meaning of *kick*, its syntax, by and large, comes for free. Crucially, then, the child is presupposed to have the knowledge that *kick* is a listed item, a 'word' in the relevant sense. But what is the English learning child to make of the existence, in English, of *make white*, consisting of two phonological words, alongside *whiten*, consisting of a single phonological word, with the very same meaning, but with the latter, but not the former, allowing an inchoative syntax? Even more tricky, what is the child to make of the listed nature of e.g. *fire*, encountering it in context as diverse as *the fire*, *firefly*, *fireplace* and *brushfire*, and with only some, but by no means all these expressions linked with some 'primary' notion of what *fire* might actually mean? Similarly, would the child be driven to assume that *kick* is a 'word' and hence may come with meaningful syntactic information related to its Content, but not so *kick the bucket*, which, surely enough, does have Content, but that Content cannot be translated into syntactically projecting argument structure? In other words, why is *kick* potentially a terminal, but not so *kick the bucket*? Differently put, it is hard to see how a theory of acquisition which is based on the coherence of the notion 'word' could possibly be successful absent any clear notion of what a 'word' actually is, or how the child might go about recognizing it.[3]

In the last decade the claim, prevalent in the 80's and the 90's, that 'words', however defined, are junctures of phonological, morphological and syntactic properties, did come under criticism (see especially Marantz, 1997 and subsequent work as well as Borer 1994, 2003, 2005 inter alia). The present article is a continuation of this research program, insofar as it presents a serious challenge to the claim that listed items, 'words', are syntactically atomic and hence, per force, when complex not syntactically constructed. Specifically, I will show that units with an identical morpho-phonological complexity and significantly shared syntactic properties nonetheless exhibit radically different properties which are

[3] The objection is, if anything, even more valid in theories which assume that the basic listed unit is not a 'word' but a lexeme (e.g. Beard, 1995 and much subsequent literature). Specifically, a lexeme need not correspond to an actual attested phonologically realized 'word', and commonalities between its occurrences are rather linked to common meaning, some shared phonological core and some, but not necessarily all, syntactic properties. As the over-riding factor defining a lexeme is some shared meaning and being a phonological word as such is not a relevant factor, the determination, on the part of the child, of which unit, exactly, is such that its syntax must project from its meaning becomes an even harder task.

contingent on their internal syntax. Even more crucially, I will show that the specific internal syntax of such units, such 'words', corresponds directly not only to syntactic and formal semantic computational properties, but is also an extremely accurate predictor of the availability of (conceptual-encyclopedic) Content composition. There will thus emerge a local domain that delimits the availability of non-compositional Content which can only be coherently described by appealing to syntactic structures. Concretely, and in a formally well-defined way, the more *syntactically* functional the "internal syntax" of otherwise morpho-phonologically identical words, the more compositional the Content. By 'functional' I refer here to the presence of segments of extended projections.

The empirical subject matter will concern a detailed comparison of the properties of three types of Construct nominal in Modern Hebrew – essentially N-N sequences which all constitute, identically, a single domain for the assignment of primary stress, which are morpho-phonologically of identical complexity, and which share some crucial syntactic properties. Insofar as some of them display listed, non-compositional Content, they would be expected, by traditional approaches, to be handled by a non-syntactic component of word formation. Insofar as some of them have entirely predictable syntactic phrasal properties which are not compatible with putative 'word' structure, as typically conceived, they would be expected, by traditional approaches, to be handled by the syntax. And yet, a formally distinct treatment of these distinct types would be clearly missing a generalization, ascribing the non-trivial syntactic similarities between them to a coincidence. Rather, as I will argue, all N-N Constructs are syntactically formed and involve extremely similar –albeit not identical – syntactic mergers. That some, but not others, allow for non-compositional Content would, in turn, serve to delineate the specific syntactic domain which delimits non-compositionality. Insofar as the relevant domain has formal properties which correspond to Content assignment but which cannot be derived from it (e.g. by projecting information from a lexical entry), this would result in lending strong support to a syntactic treatment of all complex 'words', including those which are associated with opaque Content.

Taking yet again acquisition into consideration, we note that from the perspective of such a system, Semantic Bootstrapping is altogether not viable, and rather, must be replaced with a system that allows the learner to compute meaning, semantic and Content-related, from syntactic structure, an approach best known as Syntactic Bootstrapping (see Gleitman, 1990 as well as the specific execution in Borer, 2004). Within such a system, what the child is attuned to is not the properties of terminal units in isolation, where such a 'unit' is altogether an ill-defined notion, but rather to the properties of syntactic constituent structure. As we shall see, it is, specifically, the syntactic structure that will tell the child that in e.g. *the water in the pond* and *a cold icy water*, *water* must have the same meaning, or Content (call it *WATER*), but on the other hand, that may not be the case for e.g. *watermark* or *watershed* (with its *TURNING POINT* Content) where 'real' *WATER* Content may be altogether absent. The syntax, of course, will not tell the child *what* the Content *WATER* or *WATERSHED* is. It stands to reason, however, that equipped with the knowledge that in *watermark*, but not so in *the water*, *water* may not correspond to its canonical Content, the child will be able to make significant headway in teasing out what *is* the canonical Content *WATER* such that it is shared in *the water in the pond* or *icy cold water*, and what, in anybody's system, must be an arbitrarily listed set of Contents, as, e.g. in the *TURNING POINT* Content for *watershed.*

The organization of this paper is as follows. In section 2, I review the similarities and differences between compositional Constructs and non-compositional Constructs (NC-Constructs), the latter, by assumption, akin to 'compounds' (although what, exactly, compounds are is not a theory-neutral question. See section 4 for some discussion). Section 3 is devoted to illustrating the fact that compositional Constructs in actuality come in two flavors, and while one of them is clearly very distinct from compounds, the other occupies an intermediate space, sharing some but not all syntactic properties with NC-Constructs. The structural picture in its entirety, as we shall see in section 4, resembles quite closely a continuum otherwise observed between 'incorporated' nominals, 'pseudo-incorporated' nominals and 'stand-alone' DPs. The domain for the emergence of Content is discussed in section 5, where I elaborate

on the nature of the Content interface and propose, specifically, that extended projection segments define, syntactically, the domain of non-compositional Content. Section 6 presents a brief conclusion.

2. Constructs, Compositional and Non-Compositional – a Review

As (1a-b) illustrate, Modern Hebrew clearly has N-N sequences which are non-compositional and hence compound-like. Here and throughout, NC-Constructs are in italics:[4]

1. a. ʻ*orex* *(ha.)din*
editor (the.)law
'(the) lawyer'

b. *bet* *(ha.)sefer*
house (the.)book
'(the) school'

What is trickier, however, is setting a precise boundary between the compounds in (1a-b) and noun concatenations such as those (2), which are syntactically productive and which have Content entirely predictable from their parts:

2. a. ʻorex (ha.)maʼamar
editor (the.)article
'(the) editor of (the) article'

b. bet (ha.)sar
house (the.)minister
'(the) house of (the) minister'

c. bet (ha.)ʻec
house (the.)wood
'wood house'

The structures in (2), Construct nominals, have been discussed extensively in the generative literature in the past decades, and there is a clear consensus that regardless of their specific analysis, they are clearly creatures of the syntax. In turn, the nominals in (11) and in (2) share some major structural properties, making it implausible that their respective derivations follow entirely distinct routes. A comparative study of compositional vs. NC-Constructs (compounds) was already undertaken in Borer (1989), establishing that the cases in (1) – but not in (2) – are associated with classical diagnostics of compounding. The relevant similarities and differences are summarized and illustrated in the remainder of this section.

2.1. Similarities

2.1.1.Phonological.

N-N combinations, both compositional and non-compositional, are single prosodic units. Specifically, combinations such as those in (1)-(2) have only one primary stress falling on the non-head. A variety of phonological operations, possibly sensitive to stress placement, affect both compositional and NC-Constructs in an identical fashion, e.g. non-final vowels of the head (in open syllables) are subject to deletion on a par with such pre-penultimate vowel deletion in the stress-suffixed forms (e.g. pluralization); (cf. (3-4). Further, feminine singular forms ending in *-á* when free, and masculine plural forms ending in *-ím*, when free, exhibit a distinct bound form for the head in both compositional and NC-

[4] As Sarah Ouwayda (p.c.) observes, all observations made regarding Construct types hold equally at the very least in Lebanese Arabic, and plausibly in all dialects of spoken Arabic.

Inflection, including the definite article, is separated from the stem with a dot (*ha.bayit*, 'the house', *pqid.im*, 'clerks'). Phonological liaison is marked as – (e.g. *le-xakira* 'for investigation). Feminine and masculine marking (f. m.) is only noted on stems if otherwise relevant for the argument.

Care was taken to ensure that all non-compositional Content is associated specifically with the N-N sequence and cannot be traced back to e.g. idiomatic VPs.

Constructs (cf. (**Error! Reference source not found.**-**Error! Reference source not found.**) (recall that all instances of non-compositional constructs, NC-Constructs, are italicized):[5]

3. a. camid cmid.ím
   bracelet.m bracelet.m.pl
   b. zahav zhav.ím
   gold gold.m.pl
   c. sagur sgur.ím
   closed.m closed.m.pl

4. a. cmid kesef
   bracelet silver
   'silver bracelet'
   b. *zhav parvaim*
   gold (????)
   'superior pure gold'
   c. cmid (ha.)mor.á
   bracelet (the.)teacher.f
   'the bracelet of the teacher'

5. a. šmira
   guarding
   b. *šmirat sáf*
   guarding threshold
   'gate keeping'
   c. šmirat yelad.ím
   guarding children
   'guarding children'

6. a. bat.ím
   'houses'
   b. *bat.ey midráš*
   houses Talmudic verse
   'Talmudic study center'
   c. bat.ey mor.ót
   houses teachers
   'teachers' houses'
   d. bat.ey ‘ec
   houses wood
   'wood houses'

2.1.2. Syntactic.

2.1.2.1 Regardless of compositionality, a modifier can never occur directly after the head, even when the head is modified separately from the non-head. Rather, such a modifier must follow the non-head, indeed, it must follow all Construct non-heads if there is more than one. Note that for NC-Constructs, it goes without saying that the head cannot be modified separately from the non-head, and the post-Construct placement might appear sensible. What is striking, however, is that the same placement is attested even in cases of full compositionality, in contrast with modification in free nominals, as in (7):[6]

7. *Free nominal*:
   (ha.)delet (ha.)xadaša šel (ha.)bayit (ha.)yašan
   (the.)door.f (the.)new.f of (the.)house.m (the.)old.m
   ’the/a new door of the/a old house’
8. *Construct, compositional*;
   a. delet (ha.)bayit (ha.)xadaša
   door.f the.house.m new.f
   ‘the/a new door of the house’
   b. *delet (ha.)xadaša (ha.)bayit

[5] *Parvaim*, as in (4b), is a veritable *cran* morph, in that it neither occurs elsewhere in the language, nor does it have any discernible meaning.

The phonological properties of the Construct have been derived in Tiberian Hebrew by Prince (1973) precisely from the absence of primary stress on the head. The extent to which such derivation could apply to Modern Hebrew, however, has been disputed, a matter that is by and large orthogonal to the subject matter of this study. Throughout, the transcription is intended as theory-neutral and correspond, roughly to the pronunciation of Modern Hebrew.

[6] If both head and non-head are modified, the order must be nested: $N_1\ N_2\ A_2\ A_1$.

door (the.)new (the.)house

9. *Construct, compositional:*
   a. ricpat (ha.)'ec (ha.)xadaša
      floor.f (the.)wood.m (the.)new.f
      'the/a floor of new wood'
   b. *ricpat (ha.)xadaša (ha.)'ec
      floor (the.)new (the.)wood
10. *Construct, non-compositional*:
   a. *bet* (ha.)*xol.im* (ha.)xadaš
      house.sg (the.)patient.pl (the.)new.sg
      'the/a new hospital'
   b. **bet* (ha.)xadaš (ha.)*xol.im*
      house (the.)new (the.)patients

2.1.2.2. The definite article, *ha*, cannot be realized on the head of the Construct regardless of compositionality. In turn, when it is realized on the (last) non-head, the entire expression, with the bare left-most N as its head, is syntactically definite. There are (at least) two environments in which such grammatical definiteness can be tested. One involves definite agreement on a modifying adjective in the form of a reiteration of the marker *ha* on the adjective (cf. 11a,b). The other involves the occurrence of the direct object marker *et*, obligatory for definite DPs and proper names and impossible otherwise (cf. 11c). That all Construct types regardless of compositionality are definite, in accordance with these tests, is illustrated by (12) for compositional cases, and by (13) for non-compositional ones:[7]

11. a. (**ha.)**bayit (**ha.)**xadaš (free nominal)
      (the.)house.m (the.)new.m
   b. ha.po'al.im cav'u *(et) ha.bayit
      the.workers painted OM the.house
   c. ha.po'al.im cav'u (*et) bayit
      the.workers painted (*OM) house
12. a. *ha.bet (ha.)mora (ha.)xadaš)
      the.house.m (the.)teacher.f (the.new.m)
   b. bet (ha.)mora (ha.)xadaš
      house.m the.teacher.f the.new.m
      '(the) new teacher's house'
   c. ha.po'al.im cav'u *(et) bet ha.mora
      the.workers painted OM house the.teacher
      'the workers painted the teacher's house'
   d. ha.po'al.im cav'u (*et) bet mora
      the.workers painted (*OM) house teacher
      'the workers painted a teacher's house'
13. a. *ha.*bet* (ha.)*xol.im* (ha.xadaš)
      the.house.sg (the.)patients.pl (the.new.sg)
   b. *bet* (*ha.)xol.im* (ha.)xadaš
      house.sg (the.)patients.pl (the.)new.sg
      '(the) new hospital'
   c. ha.po'al.im cav'u *(et) *bet* *ha.xol.im*

[7] *Definite* here and elsewhere in this article is 'syntactically definite', as borne out by definiteness agreement and by the distribution of the object marker *et*. On the potential difference between syntactic and semantic definiteness in the Construct, see Engelhardt (2000) and Danon (2001, 2008). This matter is largely orthogonal this study.

the.workers painted OM house the.patients
'the workers painted the hospital'

d. ha.poʻal.im cavʻu (*et) *bet* *xol.im*
the.workers painted (*OM) house patients
'the workers painted a hospital'

Anticipating somewhat, in section 3 I will argue that compositional Constructs are not a uniform class, and that there are further similarities between NC-Constructs and one sub-type of compositional Constructs.

2.2 Differences

2.2.1 Constituent Structure.

2.2.1.1. While compositional Constructs allow the modification of the non-head, such modification is altogether impossible for NC-Constructs without the loss of the non-compositional reading. Note that adjectives agree with the noun they modify in gender, number and definiteness, and hence e.g. *ha.xadaš.im* 'the.new.pl' in (14b) perforce agrees with 'patients ' and not with the entire expression, 'hospital' which is in this case singular:

14. a. bet ha.talmid.im ha.xadaš.im
house.sg the.students.pl the.new.pl
'the house of the new students'

b. (*)*bet* *ha.xol.im* ha.xadaš.im
house.sg the.patients.pl the.new.pl
'the new patients' house; *the new hospital;* the hospital for the new patients'

2.2.1.2. While the non-head in compositional Constructs may be coordinated (cf. 15), such coordination is excluded with NC-Constructs (cf. 16). Nor can two non-heads of a NC-Construct be coordinated, even when the head is identical (cf. 17):

15. a. bet talmid.im ve-talmid.ot
house students.m.pl and-students.f.pl

b. gan per.ot ve-yeraq.ot
garden fruits and vegetables
'a garden of fruit and vegetables'

16. (*)*bet* *xol.im* ve-xol.ot
house patients.m.pl and-patient.f.pl
'a house of male and female patients'; *hospital for male and female patients"

17. a. *gan* *yelad.im*
garden children
'a kindergarten

b. *gan* *xay.ot*
garden animals
'zoo'

c. gan yelad.im ve-xay.ot
garden children and-animals
'*a kindergarten and a zoo'
'a kindergarten and animals'
'a garden for children and animals'

2.2.2 Pronominal Reference

2.2.2.1. While a pronoun may refer to the head of a compositional Construct, excluding the non-head), (cf. 18a-b), such reference is impossible with a non-compositional reading (cf. 19):

18. a. hu bana li šney bat.ey ‘ec ve-*exad* mi-plastic
he built me two houses wood and-one of-plastic
'he built for me two wooden houses and one of plastic'

b. ‘amdu šam šney batey mor.ot mi-xul ve-*'exad*
stood there two houses teachers from-abroad and one
šel mora mi-be'er še
of teacher from Be'er Sheba
'there were two houses there of teachers from abroad, and one of a teacher from Be'er Sheba'

19. *hu bana lanu šney *bat.ey xol.im* ve-exad le-yetom.im
he built us two houses patients and-one for-orphans
'he built for us two hospitals and one orphanage'
cf. *bet xol.im;* *bet yetom.im*
house patients house orphans
‘hospital’ ‘orphanage’

2.2.2.2. A pronoun may refer to the non-head in (some) compositional Constructs (see section 3 for qualification). Such reference to the non-head in the NC-Construct, however, results in loss of the non-compositionality:

20. a. bikarti be-bet ha.talmid.im$_i$ ve-hem$_i$ higišu li te
visited in-house the.students and-they served me tea
'I visited the students' home and they served me tea'
a. bikarti be-*bet* *ha.txol.im*$_i$ ve-hem$_i$ higišu li te
visited in-house the.patients and-they served me tea
'I visited the patients' home and they served me tea'
'*I visited the hospital and the hospital's patients served me tea'

A particularly interesting illustration of the contrast comes from coordination. With a few prescriptively frowned upon exceptions, the head of the Construct may not be coordinated directly (cf. 21). The relevant interpretation can be gained, however, through the coordination of two Constructs, and with the second non-head realized as a pronoun referring to the first non-head, as illustrated in (21)-(22):

21. a. *bet ve-xacar ha.mora
house and yard the.teacher
'the teacher's house and yard'
b. *bet ve-xacar mora
house and yard teacher
'a teacher's house and yard'

22. a. bet ha.mora$_2$ ve-xacer-a$_2$
house the.teacher$_2$ and-yard-her$_2$
'the teacher's house and her yard'
b. bet mora$_2$ ve-xacer-a$_2$
house teacher$_2$ and-yard-her$_2$
'a teacher's house and her yard'

Such coordination for NC-Constructs results in the immediate loss of the non-compositional Content:

23. a. mitat (ha.)xol.im$_2$ ve-bet-am$_2$
bed (the.)patients$_2$ and house-theirs$_2$
'*(the) patients bed and their hospital'
'the/a hospital and their bed' (with pronominal reference vague)

cf. *bet* *(ha.)xol.im*
house (the.)patients
'(the) hospital'

b. *bet* *(ha.)xol.im*$_2$ ve-mitat-am$_2$
house (the.)patients$_2$ and bed-theirs$_2$
'(the) patients$_2$ home and their$_2$ bed'
'the/a hospital and their bed' (with pronominal reference vague)

24. a. *iš* *(ha.)sefer* *bet* *(ha.)sefer*
man (the.)book house (the.)book
'(the) scholar' '(the) school'

b. *'iš* *(ha.)sefer*$_4$ ve-bet-o$_4$
man (the)book$_4$ and-house-his$_4$
*the/a scholar and the school'
'the/a scholar and his house'

2.2.3. Definiteness Spreading:

As noted, the definite determiner in both compositional and NC-Constructs may only be realized on the last non-head member. In (some) compositional Constructs, the definiteness marked on that non-head is associated not only with the entire expression, but also with the non-head itself (see section 3 for qualifications). If the Construct has more than two members, such definiteness comes to be associated with every single noun in it, as can be illustrated through the obligatoriness of agreement on adjectives modifying such non-heads (and see fn. 6 for adjective ordering and agreement):

25. a. delet bet ha.mora ha.vatika ha.xadaš
door.f house.m the.teacher.f the.senior.f the.new.m
'the door of the new house of the senior teacher'

b. delet bet ha.mora ha.xadaš ha.lavana
door.f house.m the.teacher.f the.new.m ha.white.f
'the white door of the new house of the teacher'

Not so in NC-Constructs, where the non-head, although directly marked by a definite article as already noted, is not even coherently referring, let alone definite in any semantically meaningful sense. The 'orphans' under discussion in (26a) not only need not be specific, they need not exist altogether, nor does the 'king' (26b) need to be a specific one or exist. The expression is best translated as 'prince', rather than a 'king's son'.[8]

26. a. *bet* *ha.yetom.im*
house the.orphan.pl
'orphanage'

b. *ben* *ha.melex*
son the.king
'prince'

2.2.4 Semantic Headedness.

Adapting somewhat the 'IS A' condition of Allen (1978), we note that a compositional Construct IS A modified version of its head. Not so NC-Constructs, where such entailment need not apply:

27. a. bet mora is a bayit b. šomer mexoni.ot IS A šomer
house teacher IS A house guard cars IS A guard

[8] Hebrew Wiktionary exemplifies both *ben melex*, lit. 'son of a king' and *nasix*, 'prince' by referring to Prince Charles. With thanks to an anonymous reviewer for pointing out Prince Charles' relevance here.

28. a. *bet* *sefer* IS NOT A bayit (necessarily)
house book IS NOT A house (necessarily)
'school (lit. book house)'
b. *yošev* *roš* IS NOT A yošev
sitter head IS NOT A sitter
'chairman (lit. head sitter)'

Table 1 is a summary of similarities and differences observed thus far between compositional Constructs and NC-Constructs:

Table 1
Compositional Constructs (CC) and NC-Constructs (NCC): Similarities and Differences

			NCC	***CC***
Similarities	a	Stress-assignment domain	Yes	Yes
	b	Definite Article Placement	Yes	Yes
	c	Modifier placement	Yes	Yes
Differences	a.	Semantic headedness (compositionality)	No	Yes
	b.	Coordination	No	Yes
	c.	Pronominal reference to the head	No	Yes
	d	Non-head modification	No	Yes
	e.	Pronominal reference to non-head	No	Yes (some)
	f	Definiteness Spreading	No	Yes (some)

Observing the properties in Table 1 now, we note that the lack of compositional Content for some Constructs correlates directly with syntactic opacity. In turn, cases of compositionality correlate with syntactic transparency. The label *compound* for NC-Constructs thus appears well deserved. This said, a closer scrutiny reveals that at least some compositional Constructs are not as syntactically distinct from our *compounds* as may be presumed on the basis of the discussion thus far, a matter we turn to directly.

3. Modification Constructs

3.1. M-*Constructs vs.* I-*Constructs – the syntax*

We take as our starting point the distinction between Individual Genitives (I-Genitives) and Modificational Genitives (M-Genitives) as motivated in Munn (1995). Compositional constructs can be clearly divided into modificational (M-Construct) and individual (I-Construct) types as illustrated in (29)-(30):[9]

[9]In a careful study of English Saxon genitives, Munn (1995) distinguishes between Individual Genitives, such as those in (i) and Modificational Genitives, such as those in (ii). As Munn shows, the non-head in the former is an individual and a full DP while the non-head in the latter is a modifier and not a full DP. At least some of the relevant properties of the English construction should emerge from the text discussion:

i. Mary's hat and her bag
Mary's three hats
*three [Mary's hats] (unless *Mary's hat* is hat-design type)

29. bet (ha.)'ec; kos (ha.)mic; mad.ey (ha.)cava
    house (the.)wood; glass (the.)juice; uniform.pl (the)army
    '(the)wooden house; (the) juice glass; (the) army uniform'

30. bet (ha.)mora; na'al.ey (ha.)yalda; mano'a (ha)-mexonit
    house (the.)teacher; shoes (the.)girl; engine (the)-car
    '(the) teacher's house' '(the) girl's shoes' '(the) car's engine'

The intuitive meaning difference is clear, but considerably more crucial is the fact that the types are syntactically and semantically distinct. In (31), I list the syntactic differences between these two types of Constructs:[10]

31. *The non-head in* M-*Constructs -*
    a. when modified, the modifier is interpreted as a defining a sub-kind (32a) vs. (32b).[11]
    b. can only be modified when indefinite (33a) vs. (33b) (no definiteness spreading)[12]
    c. cannot be quantified (34a) vs. (34b)
    d. does not allow pronominal reference(35a) vs. (35b)
    e. does not allow determiners or adjectives that entail reference (36a) vs. (36b)

32. a. *I-Construct*:
    bet mora ce'ira
    house teacher young
    'a young teacher's house'
    b. *M-Construct*

---

ii men's coats and (*their) shoes
(as in *coats and shoes typically worn by men*)
many [women's jackets]
*women's many jackets
(as in *jackets typically worn by women*)

Overwhelmingly, as we shall see, M-Constructs correspond to compounds in English and are illicit as M-Genitives in Munn's sense:

i. a. wood house; tea cup; composition style, army uniform
b. *wood's house; tea's cup; composition's style; army's uniform

The differing distribution of labor between M-Constructs/M-Genitives and compounds/NC-Constructs in English and in Hebrew is discussed in Borer (2012b).

[10] Dobrovie-Sorin (2003) likewise draws a syntactic distinction between (our) I-Constructs and M-Constructs, suggesting that compounds are related to the latter. See also a brief discussion in Hazout (1991).

[11] That the constraint is structural emerges directly from the high degree of 'coercibility' of the modification. Insofar as *transparent glass* could be considered a *sub-type* of glass, as opposed to an individual modification of a particular sheet of glass, this would suffice to render (32b) licit. Similar effects are observed by Munn (op. cit.) for English M-Genitives (tall man's coat, but #pleasant man's coat), as well as by Dobrovie-Sorin, Espinal and Bleam (2006) for Spanish and Catalan pseudo incorporated, number neutral nominals.

[12] Note that the non-head in M-Constructs can host a definite article, as in (29). Furthermore, the M-Construct as a whole can be modified when definite (cf. (i)):

i. ricpat ha.ec ha.yafa
floor.f the.wood.m the.beautiful.f
'the beautiful wood floor'

What, specifically, is *not* possible is the modification of the non-head with an adjective which has a definite agreement marker on it, as in (33b), showing that even when the non-head hosts a definite article, it is not, itself, definite. Diffrerently put, in M--Constructs Definiteness Spreading does not apply. See Borer (2012b) for some additional discussion of this point.

xalon zxuxit {#škufa; venezianit}
window glass {#transparent; Venetian}
'#a [transparent glass] window
'a [Venetian glass] window

33. a. *I-Construct*:
bet (ha.)mora (ha.)vatika house (the.)teacher (the.)senior
'the house of a/the senior teacher'
b. *M-Construct*:
xalon ha.zxuxit {#ha.škufa; ??ha.venezianit}
window the.glass {#the.transparent; ??the.Venetian}
'the [{#transparent; ??Venetian glass}] window'

34. a. *I-Construct*:
bet šaloš/harbe mor.ot; / bet kol mora
house three/many teachers house every teacher
'a house of three/many teachers' 'every teacher's house'
b. *M-Construct*:
*qir me'a/harbe leven.im; *bet kol 'ec
wall hundred/many bricks house every wood
'*a wall of hundred/many bricks' '*a house from every (type of) wood'

35. a. *I-Construct*:
bet (ha.)mora$_3$ ve-rahit.e-ha$_3$
house (the) teacher and furnitures-her
'the/a teacher's house and her furniture'
b. *M-Construct*:
i. *xalon (ha.)zxuxit$_1$ ve-dalt-a$_1$
window (the.)glass.f and door-her
'*the/a glass$_3$ window and it$_3$s door'
ii. *mad.ey (ha.)cava$_i$ ve-kumt.ot-av$_i$
uniform.pl (the.)army and hats-his
'*the army$_i$ uniform and its$_i$ hats'

36. *I-Construct*:
a. bet {'eyze} mora {kolšehi/mesuyemet}
house {some} teacher {some/specific}
'a house of some/specific teacher'
b. *xalon {'eyze} zxuxit {kolšehi/msuyemet} M-Construct
window {some} glass {some/specific}
'a window of some/specific glass'
(All excluded readings under relevant interpretation).

3.2. M-*Constructs, Compounds and Pre-N-N Determiners*

Consider now again the respective properties of I-Constructs, M-Constructs and NC-Constructs (and setting aside across-the-board similarities), as in Table 2:

Table 2
I-Constructs, M-Constructs and NC-Constructs

	NCC	***MC***	***IC***

a.	Semantic compositionality	No	Yes	Yes
b.	Coordination	No	Yes	Yes
c.	Pronominal reference to the head	No	Yes	Yes
d.	Non-head modification	No	Sub-kind modification only	Yes
e.	Pronominal reference to the non-head	No	No	Yes
f.	Cardinals or quantifiers w/the non-head	No	No	Yes
g.	definiteness spreading	No	No	Yes
h.	Determiners and reference denoting adjectives w/non-head	No	No	Yes

Clearly, M-Constructs share many of the properties of NC-Constructs. The similarity is further supported by a development in spoken Modern Hebrew which affects M-Constructs and NC-Constructs, but not I-Constructs, marking the latter as clearly distinct from both M-Constructs and NC-Constructs. Specifically, the placement of the definite article in Constructs is shifting in spoken Modern Hebrew from a realization on the non-head to a realization on the head itself, which is to say, to the left periphery of the nominal expression. In such cases, the entire Construct is definite. Importantly, in such cases the non-head cannot be independently marked with a definite article, nor can it be interpreted as definite, as illustrated by (37)-(38):[13]

37. *NC-Constructs (italicized):*
    a. ha.*yom huledet* šel-i
       the.day birth mine
       'my birthday'
    b. ha.*bet* *sefer* ha.ze
       the.house book the.this
       'this school'

    c. ha.‘*orex* *din* ha.ca‘ir
       the editor law the.young
38. *M-Constructs* (underlined)
    a. ha.<u>kos mic</u> ha.zot
       the.glass juice the.this
       'this glass of juice'
    b. ha.<u>magevet mitbax</u> ha.meluxlexet ha.zot
       the.towel kitchen the.dirty the.this
       'this dirty kitchen towel'

[13] The generalization has a number of apparently listed exceptions affecting specifically NC-Constructs but never M-Constructs, and thus, contrasted with (39a) we have the cases in (i) (and where *kneset* is otherwise only attested as the name for the Israeli parliament):

i. *ha.bet* *ha.kneset* ha.ze; *ha.bet* *ha.sefer* ha.ze
   the.house the.*kneset* the.this the.house the.book the.this
   'this synagogue' 'this school'

Insofar as such 'doubling' is only attested with NC-Constructs, we note, it serves to bolster the structural distinctions between NC-Constructs and M-Constructs otherwise argued for in the next subsection.

c. ha.mad.ey cava ha.'ele
the uniform.pl army the.these

39. a. **ha.yom ha.huledet;* **ha.gan ha.yelad.im;*
the.day the.birth **the.garden the.children
(the birthday) (the kindergarten)
**ha.ʻorex ha.din*
*the.editor the.law
(the lawyer)

b. *ha.kos ha.mic; *ha.magevet ha.mitbax;
the.glass the.juice; the.towel the.kitchen;
*ha.mad.ey ha.cava
the.uniform the.army

When applied to Constructs with a (contextually plausible) individual non-head, such placement of the definite article has the effect of converting them, however implausibly, to M-Constructs (cf. 40): the non-head acquires the syntax and the interpretation associated with a modifier defining a sub-kind, disallowing definite and non-sub-kind modification. It can no longer be pluralized or quantified, and pronominal reference to it becomes impossible (cf. 41-42):

40. a. ha.tmun.ot muzeon ha.'ele
the.pictures museum the.these
(*'these pictures of the museum')
(ok: 'these museum-type pictures')

b. ha.naʻal.ey yalda ha.xum.ot
the.shoes girl the.brown.pl
(*the brown shoes of the girl)
(ok: the brown ʻgirl-type shoesʼ)

41. a. ha.bet mora ha.ze/ha.zot
the.house.m teacher.f the.this.m/*the.this.f
'this [teacher house]'
*'[this teacher] house'

b. ha.siml.ot rofa ha.'ele
the.dresses doctor the.these
'these [physician dresses]'

42. a. *ha.simlat rofa (ha.)vatika
the.dress physician (the.)senior
'*the dress of (the) senior physician'

b. *ha.simlat kol rofa
the.dress every physician
'the dress of every physician'

c. *ha.simlat $rofa_3$ ve-kova-a_3
the.dress physician and hat-her
'the physician's dress and her hat'

d. *ha.simlat {'eyze} rofa {kolšehi/mesuyemet}
the.dress {some} physician {some/specific}
'the dress of some/specific physician'

Siloni (2001) notes that when semantically definite (non-affixal) prenominal determiners such as *oto*, 'the same' and the postnominal demonstratives *ze/'ele* (this/these/that/those) are used with (compositional) Constructs, the entire expression is definite, but not so the non-head. This, Siloni reasons, suggests that while definiteness does spread from the non-head to the head, indefiniteness does not, and is rather

associated independently with each N member of the Construct. However, the properties of *oto* and *ze/'ele*, as it turns out, are identical to those just outlined for the reanalyzed definite article *ha.* when it occurs at the left periphery – they are only compatible with M-Constructs and NC-Constructs, and are strictly barred in the context of individual non-heads:

43. a. 'oto ‘*orex din* ca‘ir / ‘*orex din* ca‘ir ze
same editor law young editor law young this.m
'the same lawyer' 'that young lawyer'

b. ota kos mic / kos mic zot
same glass juice glass juice this.f
'the same glass of juice' 'this glass of juice'

44. a. oto bet mora (*vatika) /
same.m house.m teacher.f (*senior.f) /
'the same [(*senior) teacher's house]'
/ bet mora (*vatika) ze
/ house.m teacher.f (*senior.f) this.m
'this [(*senior) teacher's house]'

b. *ota simlat kol rofa /
same dress every physician /
'the same [dress of every physician]'
/ *simlat kol rofa zot
/ dress every physician this
'this [dress of every physician]'

c. *otan siml.ot $rofa_3$ ve-kova.e-a_3 /
same dresses physician and hats-her /
'the same [dresses and hats of a physician] '
/ *kova $rofa_3$ ze ve-simlat-a_3
/ hat physician this and-dress-her
'this [hat of a physician] and her dress'

d. *ota simlat {'ezye} rofa {kolšehi/mesuyemet} /
same dress {some} physician {some/certain} /
'the same [some/specific physician's dress] '
/ *simlat {'eyze} rofa {kolšehi/mesuyemet} zot
/ dress {some) physician {some/specific} this
'this [some/specific physician's dress]

We note now that the emergence of a reanalyzed determiner placement for NC-Constructs and for M-Constructs, but not for I-Constructs would follow directly if we assume that in I-Constructs, the non-head is a DP, but in M-Constructs and in NC-Constructs it is not. Specifically, in I-Constructs both the head and the non-head constitute full DPs, each with its own definiteness feature. *Definiteness Spreading*, as well as *Indefinite Spreading* for that matter, is in turn an operation which effectively copies the (in)definiteness value of the embedded DP onto the one dominating it, much as occurs in Saxon Genitives (e.g. *the dog's tail* vs. *a dog's tail*). We already observed that *Definiteness Spreading* does not occur in M-Constructs and NC-Constructs, a natural result of the fact that they are *not* DPs or individuals, but rather they are predicates and hence neither sensibly definite nor sensibly indefinite. It is precisely in these cases that the placement of the definite article on the *non-head* creates both a syntactic and semantic anomaly – it divorces the placement of the article from the presence of any D node, as well as from the presence of an expression that can sensibly be definite or indefinite. The reanalyzed position of the article summarily does away with both of these anomalies. Not so the non-head in I-Constructs, which has individual reference.

3.3 M-Constructs vs. Compounds

The similarities between M-Constructs and NC-Constructs now give rise to a legitimate question: Is it possible that M-Constructs are simply cases of compositional compounds? Even more specifically, are all DET-N-N combinations, including those in (40) and (43b) compounds? In terms of their interpretation, they most certainly come very close to the interpretation of typical compositional English primary N-N compounds. Nonetheless, the answer must be 'no'. Setting aside compositionality in and of itself, three of these properties were already reviewed and are summarized in Table 2 as (b-d), and concern classical compounding diagnostics. They concern, specifically, the availability of coordination for the non-head in M-Constructs but not in NC-Constructs (and see specifically the contrast between (15b) and (17c)), the availability of pronominal reference for the head in M-Constructs (cf. 18a) vs. its impossibility in (19a), and the availability of non-head modification for M-Constructs (albeit in a restricted form) vs. its absolute absence in NC-Constructs (cf. the licit M-Construct in (32b) vs. the illicit NC-Construct in (14b). As it turns out, in fact, the non-head in M-Construct may not only be modified by an adjective (providing it defines a sub-kind) but also by a PP and through the non-head itself heading a Construct (cf. 45). All these modification possibilities are summarily excluded in the presence of non-compositional Content (cf. 46). Note in particular the exclusion of (46c) consisting of the embedding of an NC-Construct within another NC-Construct (all examples are provided a DET-first alternate, to ensure the exclusion of I-Constructs):

45. a. mitkan energiya tiv'it /
facility energy natural /

/ ha.mitkan energiya tiv'it ha.ze
/ the.facility energy natural the.this
'a/this natural energy facility'

b. na'al.ey yaldat rexov / ha.na'al.ey yaldat rexov ha.'ele
shoes girl street / the.shoes girl street the.these
'(these) street girl shoes'

c. mic [tapuz.im mi-sfarad] /
juice [oranges from Spain] /
/ ha.mic [tapuz.im mi-sfarad] ha.ze
/ the.juice [oranges from-Spain] the.this
'(this) juice from Spanish oranges'

46. a. **bet xol.im* xroniy.im /
house-patients chronic.pl /
/ **ha.bet* *xol.im* xroniy.im ha.ze
/ the.house patients chronic.pl the.this
'a/this hospital for chronic patients'

b. **orex* *din* piturim / **ha.'orex* *din* piturim ha.ze
editor law dismissal / the.editor law dismissal the.this
'*the/this dismissal lawyer'

c. **bet* *xol.ey* *nefeš* / **ha.bet* *xol.ey* *nefeš* ha.ze
house patients soul / the.house patients soul the.this
'*the/this mental hospital'
cf: *xole* *nefeš*
patient/sick soul
'mentally ill individual'

d. **bet* *xol.im* me-'ayar.ot pituax /
house patients from-towns development/

/ *ha.*bet* *xol.im* me-ʻayar.ot pituax ha.ze
/ the.house patients from-towns development the.this
'a/this hospital for patients from underdeveloped towns'

In yet one more difference that divides NC-Constructs from compositional Constructs, whatever Content is conveyed by I-Constructs and M-Constructs can be conveyed with a free nominal as well:[14]

47. a. bayit šel mora
house of teacher
b. ha.yad šel ha.yeled
the.hand of the.boy
c. ha.harisa šel ha.ʻir
the.destruction of the.city

48. a. magevet šel mitbax
towel of kitchen
'kitchen towel'

b. gag šel reʻaf.im
roof of slates
'slate roof'
c. mic šel limon
juice of lemon
'lemon juice'

No such parallel form is ever possible for NC-Constructs, where an attempt at a free nominal gives rise not only to loss of non-compositional Content, but oftentimes to incoherence:

49. a. bayit šel sefer cf. *bet sefer*
house of book 'school'
'a house of a book'
b. xole šel nefeš cf. *xole nefeš*
patient/sick of soul 'mentally ill'
(incoherent)

Possibly most strikingly, and clearly indicative of a structural difference, we note that a NC-Construct, *as a whole*, may head a Construct, creating, effectively, a left-branching structure. This situation is strictly excluded for both M-Constructs and I-Constructs:

50. [*bet-sefer*] sade; [*bet-xol.im*] sade; [*beged-yam*] meši;
house book field house patients field suit sea silk
'field school' 'field hospital' 'silk bathing suit
[ʻ*orex din*] xuc; [*yošev roš*] moʻaca; [*bet-mišpat*] ʻal[15]
editor law out; sitter head council ; house trial top
'lawyer' 'council chairman' 'higher court'

[*gvinat brinza*] con; [*tabaʻt nisuim*] zahav

[14] As in M-Constructs, the free forms which correspond to them do not allow the non-head to be definite or quantified. The properties of the non-head in M-Constructs, then, do not emerge specifically from the Construct, but are those associated in general with modification.

[15] An anonymous reviewer wonders why *'al*, otherwise used in the language as a preposition meaning 'about' or 'on top of', is included here. Hebrew, however, has no intransitive prepositions, but on the other hand, has many that have nouns at their core (indeed, on a par with English *on <u>top</u> of*), of which *'al*, 'top', clearly is one.

cheese brynza sheep; ring wedding gold
'sheep [brynza cheese]' 'gold wedding ring'

51. a. *[gag bayit] re'af.im; *[na'al.ey yalda] 'or
roof house] slates [shoes girl] leather
'a house with slate roof' 'leather girl's shoes'
b. *[magevet mitbax] bad; *[livn.ey bniya] xemar
towel kitchen cloth bricks construction clay
'cloth kitchen towel' 'clay construction bricks'

4. Structural Considerations

4.1. Heads Up

Table 3 is a summary of all the properties which distinguish our three Construct types. While properties (a-e) distinguish NC-Constructs from compositional constructs, be they I-Constructs of M-Constructs, properties (h-l) group together NC-Constructs and M-Constructs as contrasting with I-Constructs. Finally, properties (f,g) exhibit a three-way distinction:[16]

[16] Plural marking on the non-head is, in actuality, attested in all Construct types, including NC-Constructs. As already noted in Borer (1989), however, in NC-Constructs the presence of a plural non-head fails to mark true plurality of any sort, and minimal pairs are attested where the choice of plural vs. singular marking for the non-head is clearly listed as such in conjunction with the relevant Content (and see section 5.2. below for more discussion):

i. a. *avodat yad* *avodat yada.im* *ozlat yad*
work hand work hand.pl scarcity hand
'handmade' 'manual labor' 'helplessness'

b. *štuax regel* *štuax raglay.im* *holex regel*
flat foot/leg flat feet/legs walker foot
'flat footed' phyllopodus (type of crab) 'pedestrian'

On the other hand, in I-Constructs plural marking on the non-head behaves, predictably, as plural marking would in standard DPs, thus distinguishing between the truth conditions of (iia) and (iib) in standard ways:

ii. a. bet ha.mora b. bet ha.mor.ot
house the teacher house the teacher.pl
'the house of the teachers' 'the house of the teachers'

M-Constructs, however, present a mixed picture. In the absence of individual reference, plural marking does not entail multiple objects, but rather, in most cases it appears to correspond to COUNT, and is missing elsewhere, as (iii) illustrates:

iii. a. gag re'af.im; 'aron bgad.im; madaf sfar.im
roof slate.pl closet garment.pl shelf book.pl
'slate roof' 'wardrobe' 'bookshelf'
b. kir beton; gag 'ec ; 'aron lexem; madaf muzika
wall concrete; roof wood closet bread shelf music
'concrete wall''wooden roof' 'bread pantry' 'mudiv shelf'

This, however, is not the entire picture. Alongside (iii) we have cases where both plural marked and non-plural marked non-heads are used with no change in construal, as in (iva) , and finally, cases where plural marking creates a coercive sub-kind effect, as in (ivb) although the non-head is ontologically COUNT (modifiers for non-heads in (iva) provided to exclude NC-Constructs; plural heads in (ivb) to highlight the non-plural non-head):

Table 3
Construct Types: Summary

		NCC	***MC***	***IC***
a.	Semantic compositionality	No	Yes	Yes
b.	Paraphrasability as free nominal	No	Yes	Yes
c.	Coordination	No	Yes	Yes
d.	Pronominal reference to the head	No	Yes	Yes
e.	[N-N]-N structures	Yes	No	No
f.	Non-head modification	No	Sub-kind modification only	Yes
g.	Pluralization of non-head	Listed	(restricted)	Free
h.	Pronominal reference to the non-head	No	No	Yes
i.	Cardinals or quantifiers w/the non-head	No	No	Yes
j.	(In)definiteness spreading	No	No	Yes
k.	Reanalyzed DEF placement	Yes	Yes	No
l.	Determiners and reference denoting adjectives w/non-head	No	No	Yes

4.2. I-Constructs and M-Constructs

Effectively, the picture in Table 3 suggests that there are (at least) three types of N-N Constructs in Hebrew. Only one of them allows for an individual non-head (*I*-Constructs), and only one of them is syntactically opaque and (could) have opaque Content. Insofar as there exists a non-head which on the one hand gives rise to Content compositionality, but on the other hand does not allow for individual reference, the three-way distinction observed here is fundamentally identical to that which characterizes the typological continuum frequently described as involving incorporated nominals, pseudo-incorporated

iv. a. mic tapuz /mic tapuz.im; na'aley yalda/na'aley yelad.ot;
juice orange /juice orange.pl; shoes girl /shoes girl.pl;
'orange juice' 'girls' shoes'
ke'ev regel/ke'ev ragl.ayim
pain foot /leg/pain foot.pl/leg.pl
'foot/leg pain'

b. magav.ot mitbax (moderni); civ.ey kir (xiconi)
towel.pl kitchen.m modern.m paint.pl wall.m external.m
'[modern kitchen] towel ' paints for external walls'
klip.ot limon tari
rind.pl lemon.m fresh.m
'rinds of fresh lemon'

This picture is incorporated into the distinct diagnostics of different Construct types as property (g) in Table 3 . A fuller account, however, is not attempted.

nominals (in the sense of Massam, 2001) and full individual status. In fact, it is entirely clear that the properties of the non-head in M-Constructs, or for that matter, English M-Genitives, tally point by point with those of nominal expressions labeled as pseudo-incorporated (or 'bare singulars' in the sense of de Swart and Zwart 2009 or Dobrovie-Sorin, Espinal and Bleam 2006) in a broad range of languages. In line with many syntactic approaches to this continuum, suppose we assume that the individually referring non-head is a full DP thereby accounting at the very least for the occurrence of quantification, for (in)definiteness spreading and for pronominal reference. Not so the non-head in both M- and NC-Constructs. In fact, given the impossibility of quantifiers or cardinals for the non-head of M- and NC-Constructs, there is little reason to assume that either D or QuantityP (#P) project. In turn, it is equally clear that N must project in the non-head position of both M-Constructs and NC-Constructs. It is equally clear that CL must be allowed to project as well, at times, giving rise to plural marking, where attested, yet again in both M-Constructs and NC-Constructs.[17]

Seeking to now hold constant the remaining properties of all N-N Construct types, suppose we assume the schematic structure in (52), where instances of F stand for segments of the nominal extended projection (see fn. 17). Within that structure, suppose we assume that the non-head, regardless of its internal structure, merges, across the board, in *some* functional specifier, and that the head N moves over that specifier to some higher functional head, much along the lines originally suggested in Ritter (1988) and much subsequent work. Depending, now, on the specific properties of the non-head, and whether it is a full DP, or alternatively an instance of either CLP or NP, the configuration that emerges corresponds either to an I-Construct or alternatively to an M-Construct. As we shall see shortly, NC-Construct is to be derived from the configuration of the latter:[18]

[17] The structure of nominal expressions assumed here is that of Borer (2005), where I assume that at the very minimum of three segments (i.e. functional heads), as in (i):

i. $[_D\ [_{\#}\ [_{CL}\ [_N\ ]]]]$

D is the locus of reference, as is standardly assumed, and is required for the semantic type <e> to emerge. # stands for Quantity andcorresponds roughly, but not entirely, to what is at times labeled NumP. #P is the home of cardinals as well as quantifiers such as *much* and *many*. In sharp contrast with many analyses of NumP, however, plural marking is *not* a property of #, but is rather a marker of the COUNT/MASS distinction, and is hence a classifier, a dividing function(marked CL in the text structures) which is realized in English as *–s* (typically) and in Hebrew as *–im* or *-ot*. MASS structures, in turn, do not involve the projection of CL.

[18] I remain silent here on the specific specifier which hosts the non-head in any of these types. In English, the head of I-Genitives must merge above the head of M-Genitives, a fact supported by the availability of a numeral before the latter but only following the former (cf. ii), as well as the possibility of having both, with the Individual Genitive preceding (cf. iii):

i. Mary's three hats
   *three Mary's hats (with IG reading for Mary)
   three men's hats
   *men's three hats (with MG reading)

ii. a. John's women's underwear
    b. *women's John's underwear

These tests, however,are moot in Hebrew, where cardinals occur preceding all construct types including I-Constructs as in (iii) and where an M-Construct can never be embedded under an I-Construct (cf. iv):

52. $[_D\ [_{\#}\ \ [_{CL}\ [_N\ \ ]]]]$

53. $[_{F2}\ \ N\ \ [_{F1}\ \ldots\ [_{specifier}$ *Non-Head* $]$ ([[[ )….. ~~N~~ …$[_{NP}$ ~~N~~ …….

DP I-Construct

CLP/NP M-Construct

CLP as well as NP are unsaturated predicates, thereby excluding individual reference for the non-head of M- and NC-Constructs. Once such predicates merge with the head of the Construct, N, itself, by assumption a predicate, they are interpreted as predicate modifiers.

Turning to the uniform phonological effects attested in all Construct types, we must reject the suggestion originally in Shlonsky (1990) (but see also Borer, 1999) according to which these emerge as a result of the syntactic incorporation of the (head of) the non-head into the head. As Siloni, (1996, 2001) points out, such an incorporation account flies in the face of the availability of coordinated non-heads in (15) (and see also Benmamoun, 2000 for Arabic). Rather crucially, the coordination of non-heads *is* possible for M-Constructs, as it is, indeed, for pseudo-incorporated nominals, making it untenable that the M-Construct, but not the I-Construct, is derived through incorporation, or, in other words, is a compound. Instead, suppose we assume, in line with a recent proposal in Ouhalla (2009), that the Construct constitutes a Prosodic Phrase, in the sense of Ackema and Neeleman (2004), and that the phonological effects attested across Construct types are the result of such prosodic phrasing.[19]

4.3. Deriving Compounds

Having concluded that the phonological effects in the Construct need not involve movement or the formation of a 'word', at least one approach to Hebrew NC-Constructs would be to suggest that they are not compounds altogether but rather, a species of phrasal idioms, e.g. on a par with *a coat of arms* or similar such expressions in English, where it is typically not assumed that a special word-formation operation is implicated in the emergence of the listed Content (and see Snyder 2001 for the explicit claim that the relevant constructs are *not* compounds).

Suppose we digress briefly to investigate what, exactly, the difference might be between 'idioms' and 'compounds', in the relevant sense. Specifically, let us assume that 'compounds', in English or otherwise, are not lexical, but are rather *syntactic* formations emerging as a result of merging two Ns. Under such a

---

i. a. šloša batey morot ce'irot
three house.pl teacher.pl young.pl
'three houses of young teachers'
b. šloša batey 'ec
three house.pl wood
'three wooden houses'
c. šloša gan.ey yelad.im
three garden.pl children
'three kindergardens'

iv. *'anivat yuqra(t) (ha.)menahel
tie prestige (the.)manager
'*the/a manager's prestige tie'

The head raising picture outlined in the text is simplified for expositional reasons. For the specifics of deriving the HEAD>NON-HEAD order and adjacency, see in particular Ritter (1991) as well as Siloni (1996). See also Borer, (1999) for a full review of the issues involved.

[19] And where the term *Prosodic Phrase* is in reference to the domain of primary stress assignment, thereby including perforce all 'phonological words'. In a departure from Benmamoun (2000) and Siloni (2001) and contra Halle and Marantz (1993), the approach under consideration here and in Borer (to appear) does not allow either an external or internal Merge of syntactic constituents in PF.

scenario, however, *all* non-compositional Content, for compounds such as *chicken wire* or for 'idioms' such as *kick the bucket* or *coat of arms* would be associated with syntactically constructed phrases, and the claim that NC-Constructs are 'idioms' but not 'compounds' would become rather vacuous.

Nonetheless, and as I will suggest (and see Borer, to appear for extensive motivation), *chicken wire*, a syntactic constituent, is assigned Content in a different way from *kick the bucket* or *coat of arms*, likewise a syntactic constituents, but with different properties. In that context, Snyder's (op. cit.) statement that MC-Constructs are 'idioms', but not so English compounds, would amount to the claim that Content is assigned to NC-Constructs in a way that is more akin to how it is assigned to *kick the bucket* or *coat of arms*, and that such Content assignment is accomplished in a distinct way for *chicken wire*, although *chicken wire*, just like NC-Constructs and *coat of arms*, is syntactically constructed.

It is fair to say, however, that this is not what Snyder has in mind. Rather, his underlying assumption is that compounds, such as in English, are constructed by means of a distinct, non-syntactic formal system, call it Word Formation, and that their output is listed. It is that listing, in turn, which allows them to interface with non-compositional Content. In turn, or so it would appear, NC-Constructs are neither lexically constructed nor lexically listed in the same way. If they do end up with non-compositional Content, it is assigned to them through some separate mechanism which is otherwise available e.g. for idiomatic syntactic constituent such as *kick the bucket* or *coat of arms*.

Considering now the evidence reviewed for the structural syntactic similarities between NC-Constructs and other N-N Construct types, it emerges that Snyder (op. cit.) certainly has a point in arguing that NC-Constructs are syntactically constructed. Any attempt to relegate them to some other formal component, call it Word Formation, would require a wholesale duplication of statements on combinatorial possibilities, on determiner placement and interpretation, on modifier placement, and so on, and at the end of the day, would not even account satisfactorily for the phonological domain effects, as they are clearly not restricted to listed forms. The question to be asked, then, is not whether NC-Constructs are syntactically derived – it is clear that they must be, and that a system must be designed to allow them to receive non-compositional Content although they are syntactically derived. The shoe, rather, is on the other foot (to use an idiom) – given the overwhelming similarities already observed between NC-Constructs and English compounds, and given our conclusion that all these properties must be available for the *syntactically* derived NC-Constructs, we must ask what evidence remains to compel us to ban English compounds from the syntax and relegate them to some parallel combinatorial component, given the fact that UG clearly *is* capable of putting together and assigning Content to structures that have all the relevant properties of English compounds, but which must be *syntactically* derived.

Once this conclusion is reached, it becomes entirely clear that even if *kick the bucket* might be assigned Content in a different way from *chicken wire*, there is little evidence from this fact alone to suggest that *kick the bucket* is syntactic but *chicken wire* is morphological, and even less evidence to substantiate the claim that the operation that puts together *chicken wire* is formally distinct from that which puts together the NC-Construct *bet sefer*, 'school' in Hebrew. Rather, and assuming that there is every reason to endorse a syntactic compositional system that can put together *kick the bucket, chicken wire* and *bet sefer*, the question, or rather questions, must be as in (54):

54. A. How is Content assigned to *chicken wire*?
    How is Content assigned to *bet sefer*?
    How is Content assigned to *kick the bucket*?
    B. How many Content assigning systems are we dealing with here,and if more than one, why should that be so and how do we know which system corresponds with each item?

Turning to an answer to these questions, it emerges that there are quite a few reasons to assume that Hebrew NC-Constructs pattern with English compounds and not with phrasal idioms such as *kick the bucket* or *coat of arms*. First, unlike phrasal idioms, and with the exception of plural marking NC-

Constructs never involve any functional material otherwise attested in phrasal idioms – no adjectives, no pronouns, no (non-affixal) articles etc., all potentially possible in idiomatic expressions as well as in compositional Constructs. Second, languages do not typically exhibit a 'phrasal idiom strategy', with non-compositional Content systematically favoring specific syntactic structure, and yet Constructs, just like English compounds, are, by far, the language's predictable source for compound formation. While English certainly does have VP idioms as well as complex nominal idioms such as *coat of arms*, the latter are positively rare, and both types exist alongside phrasal idioms that span full sentences including at times embedded clauses (*we will cross that bridge when we come to it*) as well as discontinuity and open values (*the cat got X's tongue*) and 'turns of phrase' which are altogether of unclear syntactic origin, such as *as good as it gets* or *by and large*. Nor is there any evidence of productivity in that domain. By contrast, almost every English A-N compound suffers from some degree of non-compositionality, and many an N-N compound can only be analyzed as 'compositional' in a rather vague sense (e.g. *firefly* vs. *fireman*). The situation in Hebrew is the same. While phrasal idioms certainly do exist, one would be hard-pressed to think of any non-compositional expressions based on a free nominal (i.e. no non-compositional nominal on a par with *coat of arms*). There is little reason to assume, on these grounds, any distinction between the properties of English compounds and Hebrew NC-Constructs.

Most crucial, however, is the structural fact, associated with both English compounds and Hebrew NC-Construct, which allows them, but no other complex constituents in similar configurations, to function as heads. Recall that NC-Constructs, but neither M-Constructs nor I-Constructs can themselves function as heads of other Constructs, thereby giving rise to a configuration in which the head of a Construct is itself branching (cf. Table 3 (e) and examples in (50) . An identical situation holds in English, as (55a) illustrates. In turn, (55a) contrasts directly with (55b) where such a 'branching' head is directly ungrammatical. For completeness sake, note that English non-heads may branch, in compounds as well as otherwise:[20]

55. a. paper [towel rack] (towel rack made of paper)
    b. *Mary's [computer's screen] (and compare with the relative grammaticality of the computer's screen of Mary's)
56. a. [paper towel] rack (a rack for paper towels)
    b. [Mary's computer's] screen

It thus emerges that for what it's worth, the claim of Hebrew NC-Constructs to 'compoundhood' equals that of English N-N combinations. We must now turn to the question of how they are derived, and what, if anything, singles them out from phrasal idioms.

I suggested above that the non-head for both M- and NC-Constructs is not a DP, but rather an NP or an instance of CLP, both unsaturated predicates. Turning to the distinction between these two instances of the construct, suppose we reconsider the possibility that incorporation is involved in the derivation of some, but not all N-N constructs. Recall that incorporation could not possibly be implicated in the generation of either I-Constructs and M-Constructs, given the fact that in both, non-heads may be coordinated and modified. However, as should become immediately clear, the objection is simply irrelevant for the case of NC-Constructs, because as we already showed, any attempt to coordinate or modify the non-head results directly in compositional Content. But if neither coordination nor modification are attested in NC-Constructs, the incorporation of the non-head into the head should be, in principle, possible. Suppose, then, that the N head of the specifier in (52) is free to incorporate into the raised head, providing no other grammatical principles are violated. Suppose further that incorporation is an operation which merges predicates (<et>), and that both NP and CLP are of type <et>, but # and D are of type <e>. Suppose we assume further that incorporation from within extended projections stranding

[20] As already noted in fn. 18, M-Constructs cannot be embedded under I-Constructs in Hebrew. Recursive, right branching M-Constructs *are*, however possible, as e.g. (45b) shows.

some of its segments is not licit, thereby disallowing the incorporation of NP or CLP from within a DP or an #P (contra Baker, 1988). It now follows that if the non-head in the Construct is #P or DP, incorporation cannot take place, thereby excluding incorporation across the board in I-Constructs. M-Constructs correspond to a structure in which the non-head is a predicate which does not incorporate. Compounds, i.e. NC-Constructs, finally, is what emerges as a result of the incorporation of the non-head, by assumption a predicate, into the head, as in (57). We note, before proceeding, that such an incorporation analysis accounts directly and indeed is strongly supported by the availability of NC-Constructs and NC-Constructs alone as branching heads.

57. $[_{F2}$ N- $[_{NP/CLP}$*Non-Head*] $[_{F1}$ … $[_{specifier}$ $[_{NP/CLP}$ ~~*Non-Head* ~~]
([[[ )….. ~~N~~ $[_{NP}$ ~~N~~ …..

At first sight, the analysis seems puzzling, as it forces, or so it appears, an incorporation solely for cases which are non-compositional. Given the structure in 53) why is it, one may legitimately wonder, that incorporation is obligatory for NC-Constructs but barred for M-Constructs? The puzzle, however, is only apparent, as the entailment is in fact goes the other way around. Rather than force incorporation for non-compositional Construct cases, it is incorporation which is a pre-condition for the emergence of non-compositional Content. Absent incorporation, then, Content must remain compositional. It therefore follows that insofar as NC-Constructs are by definition non-compositional, their derivation must have involved incorporation, for without such incorporation, non-compositional Content cannot emerge.

Our puzzlement, however, is now replaced with a formal query - what is it about incorporation which allows a non-compositional Content, and why is such non-compositional Content otherwise barred?

5. The Syntactic Domain of Content

5.1. Non-Compositionality in Syntactic Word Formation.

Questions concerning the (non-)compositionality of complex words extend well beyond Hebrew Constructs or, for that matter, compounds, whatever their structure may turn out to be. Rather, such questions are at the core of any attempt to combine word formation and phrasal syntax into a single computational system. Insofar as the listed, non-compositional Content of complex words is one of the primary motivations for banning them from the syntax in Chomsky (1970) and much subsequent work, any attempt to reintegrate complex words into the syntax without addressing this matter is, at best, incomplete.

The few attempts at a syntactic account, to date, have focused, and correctly in my view, on identifying a well-defined syntactic domain within which such non-compositionality might be available. Thus Arad (2003) proposes that the domain under consideration is that of (first) categorization – and specifically within her approach, the point at which the root merges with a category label (i.e. ***n, a, v***) (and see also Embick, 2010). A different, larger domain is proposed in Borer (to appear), based on the observation that non-compositional Content clearly *can* be associated with complex words beyond the domain of first categorization (clear non-compositional forms underlined):

58. a. *reactionary* (*ACT, REACT, REACTION, <u>REACTIONARY</u>*)
b. *naturalize* (*NATURE, NATURAL, <u>NATURALIZE</u>*)
c. *editorialize* (*EDIT, EDITOR, <u>EDITORIAL</u>, <u>EDITORIALIZE</u>*)

In turn, and while the structural domain which allows for non-compositional Content is clearly bigger than that of first categorization, it is nonetheless delimited quite severely. Specifically, Argument Structure Nominals (Grimshaw's 1990 Complex Event Nominals; henceforth AS-nominals) may never have non-compositional Content (cf. Marantz, 2000; Borer, 2012a; to appear). As a clear illustration of

this fact, consider the impossibility of the AS-Nominals in (59) when contrasted with (60). Note further that the deverbal nominals in (59) could be integrated into an event structure entirely felicitously, provided they are embedded within a light verb construction (cf. 61)). The anomaly of (59a-b), then, cannot be semantic:

59. a. *the *transformation* of the structure by the linguist
(jargon reading of *transformation*)
b. *the *reading* of the world by Aristotle
(compare *Aristotle's reading of the world*, authorship interpretation; (contrast also with licit *the interpretation/ understanding of the world by Aristotle*)

60. a. the *transformation* of our department by the administration
b. the *reading* of course summaries by undergraduates

61. a. the linguist did/performed a transformation on the structure
b. the reading of the world cannot be the reading *made* by academicians (Google search)

The effect is, from any possible perspective, very surprising. Both compositional and non-compositional nominals are derived from the same verb with an identical suffix, and thus there is little about their morpho-phonology that could account for this contrast. Presumably, in anybody's account, the non-compositional sense associated with (linguistic) *transformation* or *reading* must be listed *somewhere*. It is not clear, however, why such listing should correspond to the inability to take arguments, or why the ability to take arguments should correspond to the impossibility of listing, all the more so as listedness, arguments included, is precisely the hallmark of the lexicon, as typically assumed. We note further that the arguments of e.g. *transformation*, had it been allowed any, would be identical to those that are otherwise assigned by the compositional AS-Nominal or by the source verb as is clear from the interpretation of (61a-b) and from the light verb case in (61a).

Intuitively, it appears that what keeps the derived nominals in (60) compositionally "honest" so to speak, is not their relationship with a source verb as such, but rather, the actual existence of a full argumental complex. When the full argument complex, presumably including a V as well, is nominalized, the deverbal nominal itself must be compositional. When it is the verb alone that nominalizes and without any arguments in presence, non-compositionality may (but need not) emerge for the deverbal nominal. The simplest, most direct way to capture this generalization would thus be syntactic: in AS-Nominals, the nominal head scopes over the verbal/argumental complex, the latter including whatever functional structure is implicated in the presence of arguments. In the absence of such a functional argumental complex, the nominal head scopes over the verb alone, and excludes, specifically, any functional structure that may be implicated in the merger of arguments.

Building on this intuition, suppose we assume now that the syntactic domain of non-compositionality is restricted by functional structure, where by 'functional structure' I refer here to the reservoir of nodes which are non-terminal segments of extended projections (e.g. T, Asp, D, Deg etc.). Before elaborating, however, it is worthwhile to review in greater detail our system of Content assignment.

Suppose we assume the existence of a reservoir of atomic, indivisible Content units, call it the *Encyclopedia*. While there certainly are constraints on what may or may not be an atomic Content unit, I assume that such constraints do not come from the grammar, nor are Content units specifically linguistic or language-determined units. Rather they are conceptual and are constrained as such. Interfacing between the Encyclopedia, a non-linguistic module, and the linguistic system we find a 'reading device' that is capable of recognizing syntactic strings of particular size and matching them with individual Content units. A successful single encyclopedic search – *en-search* – thus returns a single atomic Content unit for a qualifying linguistic domain. We note now that the term *non-compositionality* is but another name for an *atomic Content unit*. Insofar as *transformation* in its linguistic technical sense has Content

that cannot be predictable from its parts, it is as much an atomic Content unit as *cat*. Insofar as *transformation* in its compositional sense does have a predictable relationship with its parts, it is not a single atomic Content unit any more than e.g. *eating apples*, although, of course, *transform* in all likelihood is an atomic Content unit. In computing the meaning of compositional *transformation*, then, a single *en-search* could return Content for *transform*, call it *TRANSFORM*, which would then be composed with whatever (fixed) function is associated with *–ation* to give rise to the composed interpretation of *transformation*. Importantly, in this system roots or stems do not have inherent 'basic' Content as such, and all Content, of both complex and simple constituents is assigned at the same stage, and through the *en-search* reader.[21]

The most straightforward way of capturing the obligatorily compositional Content of AS nominals would be to propose that the functional structure implicated giving rise to event structure blocks *en-searching*. Specifically, suppose a single *en-search* cannot extend past a functional bracket. Differently put, a phrase that contains a functional bracket may not return a single atomic Content unit, and thus must be compositional. Suppose we assume further that Content, once assigned may not be overridden, and that *en-searching* may target any qualifying domain. The representations that would now emerge for deverbal nominals without argument structure - R-nominals – would thus be as in (62), where in the absence of any functional brackets, two domains may be defined, giving rise to a compositional assignment, as in (62a) (domain boxed) and to a non-compositional one, as in (62b):

62. *transformation, R-Nominal*:

a. $[_N\boxed{[_V \quad V \quad]} \quad N \quad [_V \quad \cancel{V} \quad]]$

(trans)form *ation*

TRANSFORM *ation* → *the transformation is complete*

b. $\boxed{[_N[_V \quad V \quad] \quad N}\quad [_V \quad \cancel{V} \quad]]$

(trans)form *ation*

TRANSFORMATION → (linguistic jargon)

Consider now AS-nominals. Here, the movement of the verb to adjoin to N must proceed through a number of intervening functional heads which in turn license the relevant arguments (event-related functional nodes labeled as F2/F1 for ease of exposition and structure simplified). The result is the configuration in (63), where, specifically, at least two functional brackets (or one, for intransitives) separate *ation* from *transform*. As functional brackets stop *en-searches*, the sole *en-searchable* domain for the AS-nominal *transmission* cannot extend beyond the constituent that includes *transform*, and which returns *TRANSFORM*. Compositionality in AS-nominals is thus fully enforced, contingent, indeed, precisely on the very presence of the functional structure which makes them Argument Structure Nominals:[22]

63. *transformation, AS-Nominal*

a. $[_N \quad N \quad [_{F2}$ subj F2 $\quad [_{F1}$ obj F1 $\quad [\ \cancel{V}\ ]]]]$

[21] And see Borer (to appear) on the formal status of derivational categorizers such as *–ation*. From the perspective of the present presentation what is crucial is the fact that they do not constitute segments of extended projections (i.e. they are not instances of F).

En-searches, by assumption, operate on bracketed *phonologically realized* representations, so as to enable them to assign atomic Content to e.g. *transmission*, but not to the structurally identical *transmittance* or *transmittal*. See reference for an elucidation of the division of labor between Content and Formal Semantics.

[22] The treatment in Borer (to appear) is rather crucially phase-based, and involves the assignment of Content to *transform* at the point at which it first merges with a functional head, and not as based on the output string in (63). As the matter is largely orthogonal to our main point here, a simpler exposition was opted for. The reader is, however, asked to bear in mind this important caveat and consult the reference for a more accurate picture.

b. $[_N[[_{F2}[_{F1}$ V]] N] $[_{F2}$ subj $[_{F2}[_{F1}$ ~~V~~]]$[_{F1}$ obj $[_{F1}$ ~~V~~] [~~V~~]]]]
transform ation *(of) the-town*

[[$_{F2}$[$_{F1}$ V]] N]
TRANSFORM *ation*

*[[$_{F2}$[$_{F1}$ V]] N]

Armed with these conclusions on the domain of Content, let us return now to the three types of Constructs discussed in sections 2-4. A re-examination of the structure in 53) reveals directly that regardless of the nominal projection of the non-head, the head and the non-head may never be part of a single *en-search*, as at the very least, one (if not more) functional bracket must separate them:

64. $[_{F1}$ N_1 $[_{F1}$ $[_{Specifier}$ *non-head*] ~~N_1~~ … $[_{NP}$ ~~N_1~~]]]

If NC-Constructs emerge as a result of a single *en*-search, and thus correspond to a single atomic Content unit, then it now emerges that even when the non-head is NP or CLP, the structure in 53) cannot, in and of itself, be that of NC-Constructs. Rather, the structure in 53), must always be compositional and computed on the basis of the assignment of distinct Content to the head and to the non-head.

Consider, however, the possibility that the non-head may incorporate into the head. By assumption, such incorporation cannot affect the non-head of I-Genitives, as DPs may not incorporate, nor, by assumption, can functional structure be stranded, as already noted. If, however, the non-head is N or CLP, such incorporation is licit. The incorporation would result directly in eliminating the offending F1 bracket intervening between the head and the non-head, and with the emerging structures in (65) (note that F1 is part of the extended projection dominating N1, the head):[23,24]

65. a. $[_{F1}$ N_1-$[_N$ N_2] $[_{F1}$ $[_N$ ~~N_2~~] ~~N_1~~ … $[_N$ ~~N_1~~]]]]
bet-sefer *sefer* *bet* *bet*
SCHOOL (book) (house)

b. $[_{F1}$ N_1-$[_{CL}N_2$-CL] $[_{F1}$ $[_{CL}$ ~~N_2-CL~~ ~~N_1~~ $[_N$ ~~N_1~~]]]]
bet- xol.im *xol.im* *bet* *bet*
HOSPITAL (patient.pl) (house)

[23] While non-compositionality is only available under incorporation, the converse is not the case, and incorporated constituents should, in principle, allow compositional reading, on a par with, e.g. the optionality of compositionality for *transmission* or, for that matter, for English compounds. Nonetheless, and as already observed, the only cases of incorporation in Hebrew give rise to non-compositional Content, a conclusion that is inevitable from the fact that M-Constructs may never head a construct (cf.Table 3(e) and examples in (51b)). The reason, I believe, is to be sought in the factors which severely restrict M-Genitives in English, resulting in the availability of e.g. (i), but in the obligatory compounding of the cases in (ii). These conditions are discussed in some detail in Borer (2012b):

i. man's coat; boy's room; baby's toy; women's voices

ii. a. *wood's table; *bird's call; *table's top; *fish's pond(s)
b. wood table; birdcall; table top; fish pond(s);

The incorporation in (ii) is thus obligatory due to English-specific constraints which exclude (iia). No such constraints exist in Hebrew, however. We may now assume that incorporation for compositional cases in Hebrew is unavailable because of economy: such incorporation would give rise to a configuration with properties that are identically available without such incorporation.

[24] Glosses are provided for e.g. *sefer* and for *bet* ('book' and 'house' respectively) for expositional purposes. The reader should bear in mind, however, that such distinct Content is *never* assigned to these strings in the representation in (65a). Rather, Content is assigned exactly once, and solely to the boxed representation.

In the absence of any intervening brackets in (65a) the assignment of atomic Content is now straightforward. Equally straightforward is the fact that the resulting constituent, post-incorporation, may itself serve as a head of a Construct. This last conclusion, we note, holds for (65b) as well.

Less straightforward, however, is the emergence of non-compositional Content for (65b). Here, it appears, $[_{CL}$ does separate N_1 from N_2, raising the question of how atomic Content can be assigned to the boxed domain. Specifically, we do not expect the emergence of atomic Content in the presence of plural marking on the non-head. Why, then, is plural marking possible within NC-Constructs?[25]

5.2. Why Plural Marking is Different

A perusal of the relevant properties of plural marking reveals it to have some important properties that distinguish it from other segments of extended projections, but are, on the other hand, shared by other classifiers, or markers of COUNT structure merging in CL. Rarely, if ever, do numbers or quantifiers, or indeed articles, give rise to non-compositional Content.[26] Plural marking, however, is regularly implicated in the emergence of exactly such Content in a broad number of languages (including Hebrew and English), in the form of *pluralia tantum*. One would be rather hard-pressed to claim that e.g. *glasses* is compositionally derived from *glass*, or *briefs* from *brief*. Rather, these are clearly cases where *en-searching* should be allowed to pick the entire string, *glasses* or *briefs*, while attempting to match it with Content. A similar effect is attested with classifiers in classifier languages, where an identical stem may acquire different Content as depending on the specific classifier that is associated with it. The stem *tienwoe* in Cantonese may combine with the classifier *ki* typically used in the context of long objects, to give rise to the Content of either a *TELEPHONE WIRE* or the *TELEHONE* itself. However with the Classifier *tung*, literally 'through' Content that emerges is that of a *TELEPHONE CONVERSATION*.

Nor is the phenomenon restricted to the nominal domain. Slavic perfective prefixes notoriously impact Content. Thus a Polish stem such as *czytała* would have the Content *READ* when occurring in isolation as well as in the presence of the perfective prefix *prze(-czytała*). When occurring with other perfective prefixes, however, it may mean *PRESENT* (*od-czytała*), *DECODE* (*roz-czytała*) or *UPLOAD* (*w-czytała*). Importantly, it is rather difficult to claim that classifiers or perfective prefixes have independent Content on their own, or that the Content of the complex forms is compositional in any sense, precisely because the very same prefix or classifier could make a very different contribution to Content with different stems.[27]

---

[25] Plural marking within compounds is typologically very common (see Booij, 1996), and is certainly attested in English as well. Most commonly, as observed in Kiparsky (1982) for irregular plurals such as *lice infested*, but also, note, for the *pluralia tantum* cases in (i) (and note that as such, they contrast with the well-known cases in (ii):

i. a. *(eye) glass store ; *rapid boat; *brief design
   b. (eye) glasses store; rapids boat; briefs design

ii. scissor edge; trouser leg

[26] This generalization cannot be reduced to the affixal nature of plural marking. Tense marking is equally affixal, and yet it is never implicated in the emergence of an atomic Content unit.

Cases such as *the Bronx* or *La France* may appear at first sight to be counter-examples to the exclusion of determiners from non-compositional atomic Content units. However, to the extent that either *France* or *Bronx* can occur without an article, as in *three Bronx men arrested; propriété industrielle en France* 'industrial property in France', they have an identical Content, casting serious doubt on a claim to non-compositionality for the definite expression.

[27] With Special thanks to Andrew Simpson and Zoe Wu for the Cantonese facts, and to Agnieszka Lazorczyk for the Polish cases.

But if both instances of stems when combined with classifiers or perfective prefixes fail to give rise to compositional Content, what reasons are there to assume that these are instances of functional structure altogether? Possibly, these are but 'bound roots' of sorts, and the non-compositionality of e.g. *glasses* is thus to be viewed on a par with that of *chicken wire*.

Crucially, however, both Slavic perfective prefixes and classifiers/plural marking do differ from e.g. *chicken* in *chicken wire*, insofar as regardless of their ability to constitute a single Content unit with the stem they are attached to, they nonetheless retain their grammatically rigid functions in broader syntactic contexts. While *briefs* may not be compositionally derived from *brief*, it is nonetheless clearly a COUNT noun, triggering count, and specifically plural agreement, as is true for all cases of *pluralia tantum*. While *ki tienwoe* ('telephone wire' 'telephone device') in Cantonese may not be compositionally derived from combining *ki* (typically 'long') with *tienwoe* (typically 'telephone'), it nonetheless continues to behave like a COUNT noun, obligatorily occurring in the presence of cardinals. While the Content of *roz-czytała* ('decode') is not compositional, the emerging verbal expression must be telic, and so on. No such effects are attested for e.g. English compounds, where the non-head never has a syntactic function. It thus emerges that excluding classifiers or perfective prefixes from our functional lexicon altogether would be an error.

This said, it is clear that the plural marking attested on non-heads in NC-Constructs or, for that matter, in English or Dutch compounds is *not* syntactically or semantically active in any way, a point already noted and discussed in some detail by Booij (1996) as well as by Acquaviva (2008) (and see fn. 16 for a brief review of the Hebrew NC-Construct picture). Following specifically on the insight in Booij (1996), we note that whether plural marking is syntactically active or not depends on its syntactic context (and hence 'contextual' plural). More concretely, we note that a classifier in the form of plural marking is syntactically – and semantically – active if, and only if – it is a segment of a (nominal) extended projection. When marking a non-head of a compound or an NC-Construct, or, for that matter, the non-head of an M-Construct, this is not the case. That the structure is nonetheless licit, indicates that classifiers, and by extension the functional node hosting Slavic prefixes, are only 'syntactically' active if embedded within an extended projection. That e.g. T-marking or D-marking are not licit for the non-head in compounds or in NC-Constructs thus goes hand in hand with the fact that they may not be syntactically or semantically inert, a property that directly distinguishes them from classifiers and perfective markers.

Suppose, then, we assume precisely that, namely that the nodes that we are labeling here as CL and PERF (a presumed functional node for Slavic perfective prefixes) are functional, in the required sense, only if they are themselves selected by some F. Differently put, CL and PERF are functional insofar as they are segments of an extended projection, but not otherwise. It thus emerges that in the structures in (69) they are functional (F), but not so in the English compound in (70a) or in the NC-Construct in (70b):

66. a. $[_{D}$ *the* $[_{\#}$ *three* $[_{CL}$*cat-s* [~~*cat*~~]]]];
F F F
$[_{D}$ $[_{\#}$*many* $[_{CL}$*factor-s* [~~*factor*~~]]
F F F

b. $[_{D}$ *the* $[_{\#}$ *three* $[_{CL}$ *rapid-s* [~~*rapid*~~]]];
F F F
$[_{D}$ $[_{\#}$*many* $[_{CL}$*scissor-s* [~~*scissor*~~]]
F F F

67. a. $[_{N}[_{CL}$ *rapid-s*[~~*rapid*~~]] *view*]; $[_{N}$ $[_{CL}$ *pant-s* [~~*pant*~~]] *pocket*]
NF NF

b. $[_{N1}$*bet-* $[_{CL2}$ *xol.im*] $[_{F1}$ $[_{CL2}$ ~~*xol.im*~~] ~~*bet*~~… $[_{N1}$ ~~*bet*~~]]]]
NF F NF
HOSPITAL (patient.pl) (house)

Returning to the non-compositionality of the boxed constituent in (65b), we note that its availability to *en-searching* now follows directly from the fact that in the absence of a dominating segment of a nominal extended projection, CL is no longer functional in the required sense. Atomic Content can thus be associated, and indeed is, with the relevant constituents

A final brief note is in order concerning phrasal idioms and the queries posed in (54). I did propose a system here that is capable of assigning Content to *chicken wire* as well as to *bet sefer* 'school', and in an identical manner. Insofar as the claim here is that Content must be contained within the first functional bracket, however, it is clear that whatever Content is assigned to e.g. *kick the bucket* or to *by and large* cannot possibly proceed that way, as both *kick the bucket* and *by and large* contain functional brackets, and in the case of *by and large* it is altogether not clear that it contains anything but. The inevitable conclusion, then, is that Content assignment to phrasal idioms must be distinct, as no single *en-search* could possibly return an atomic Content here.

The conclusion, in turn, fits extremely comfortably into the compelling semantic and syntactic arguments put forth, in particular, in Nunberg, Sag and Wasow (1994) but also by others, for the (partial) compositionality of idioms. While some measure of non-compositionality certainly would need to be accomodated, by *some* means, treating idioms as cases of e.g. *watershed* or *transmission* appear altogehter on the wrong track. The matter is discussed in some detail in Borer (to appear), and is set aside here for reasons of space.

6. Conclusion

The main aim of this article is to compare the strings in (68a-c), all of which define an identical phonological domain, and specifically all are prosodic phrases which fall within the jurisdiction of a single primary stress assignment. Even more importantly, they all share important syntactic characteristics. Nonetheless, they exhibit interpretational differences, and one of them, but not the others, allows for atomic Content, typically assumed to be within the prerogative of lexically listed 'words'. However, and precisely because neither the prosodic nor the syntactic characteristics appear formally radically different, there appears to be little motivation to assume that one of these, and specifically the one in (68c) should be relegated to a non-syntactic component while continuing to construct the others syntactically. Perhaps most striking, finally, is the convergence of specific syntactic properties with the establishment of a Content domain. The clear conclusion, then, is the correlation historically assumed to hold between primary stress domains, typically words, and Content matching simply cannot be maintained: within an identical primary stress domain, Content at times *must* be compositional, and at others may not be, and with an equally complex and clearly syntactic structure, Content, likewise, sometimes *must* be atomic, and at other times cannot be. It thus emerges that relegating any of these strings to some formally distinct component, call it the lexicon, is neither motivated nor warranted:

68.	a.	bet	ha.yalda	b.	bet	ha.'ec	c.	*bet* *sefer*
		house	the.girl		house	the wood		house book
		'the girl's house'			'wooden house'			'school'

In the introduction to this work, I presented a number of logical problems which face a language learner in the absence of well-defined expectations concerning the syntactic domain of Content. We note now that if the domain of Content is, as suggested here, determined, universally, by segments of extended projections, then the expectation for a single *en-search* and the emergence of atomic Content are extremely well-defined. As such, the presence of atomic Content may instruct the child on the specific syntactic structure at hand, and in a similar fashion, knowledge of the syntax or semantics of expressions will delimit the range of interpretations potentially assigned to them. If the child is cognizant, for instance, of the presence of individual reference for *ha.yalda* 'the girl' in (68a), s/he would never be tempted to assign the expression either atomic Content, or the structure and interpretation associated with

M-Constructs. Conversely, if the child knows the actual Content of (68c) then s/he is able to recognize it as atomic, and surmise that neither functional structure nor modification should be available. By extension, the presence of a single Content for at least some English compounds (*paper tiger, chicken wire*) should instruct the learner to seek a different structure for compounds from that which would be assigned to Saxon genitives, be they I-Genitive or M-Genitive. In both languages, we note, there is little need for a combinatorial component of the grammar which is distinct from the syntax, and which is housed in the lexicon, nor is the learner forced to assume that a radical structural difference exists between e.g. (68a) and (68c), or, for that matter, between English Saxon Genitives and compounds. Rather, the system allows the learner to assign structure to all these within the very same formal component, call it syntax, and to capitalize on her knowledge of UG-determined Content domains to do the rest.

References

Ackema, P. and A. Neeleman. 2004. *Beyond Morphology*. *:Interface Conditions on Word Formation*. Oxford: Oxford University Press.

Acquaviva, P. 2008. *Lexical Plurals*. Oxford: Oxford University Press.

Allan, M. 1978. *Morphological Investigations*. Ph.D. dissertation, University of Connecticut, Storrs.

Arad, M. 2003. Locality constraints on the interpretation of roots: the case of Hebrew denominal verbs. *Natural Language and Linguistic Theory*, 21: 737-78.

Baker, M. 1988. *Incorporation. A Theory of Grammatical Function Changing*. Chicago: The University of Chicago Press.

Benmamoun, E. 2000. Agreement Asymmetries and the PF Interface, in J. Lecarme, J. Loewenstamm & U. Shlonsky. eds., *Studies in Afroasiatic Grammar 2*. Amsterdam: John Benjamins.

Booij, G. 1996 Inherent versus contextual inflection and the Split Morphology hypothesis, *Yearbook of Morphology 1995*. Dordrecht: Kluwer 1-16.

Borer, H. 1989. On the Morphological Parallelism between Compounds and Constructs, in G. Booij et al. eds., *Morphology Yearbook 1*. Dordrecht: Foris.

Borer, H. 1994. The projection of arguments, in E. Benedicto and J. Runner. eds., *Functional Projections. University of Massachusetts Occasional Papers in Linguistics 17*. University of Massachusetts, Amherst: GLSA.

Borer, H. 1999. Deconstructing the Construct, in K. Johnson & I. Roberts. eds. *Beyond Principles and Parameters*. Dordrecht: Kluwer Academic Publishers.

Borer, H. 2003. Exo-skeletal vs. endo-skeletal explanation: Syntactic projection and the lexicon', in J. Moor and M. Polinsky. eds.,*The Nature of Explanation in Linguistic Theory*. Stanford, CA: CSLI.

Borer, H. 2004. The Grammar Machine, in A. Alexiadou, E. Anagnostopoulou, and M. Everaert. eds. *The Unaccusativity Puzzle*. Oxford: Oxford University Press

Borer, H. 2005. *In Name Only, Structuring Sense Vol. I*. Oxford: Oxford University Press.

Borer, H. 2012a In the event of a nominal, in Everaert, M., M. Marelj and T. Siloni eds. *The Theta System: Argument Structure at the Interface*. Oxford: Oxford University Press.

Borer, H. 2012b. The Projection of N, Handout. Queen Mary, University of London.

Borer, H. to appear. *Taking Form, Structuring Sense Vol. II*. Oxford: Oxford University Press.

Chomsky, N. 1970. Remarks on Nominalization, in R.A. Jacobs and P.S. Rosenbaum. eds., *Readings in English Transformational Grammar*. Waltham, MA: Ginn.

Danon, G. 2001. Syntactic Definiteness in the Grammar of Modern Hebrew, *Linguistics* 396, 1071-1116.

Danon, Gabi. 2008. Definiteness Spreading in the Hebrew Construct State, *Lingua*.

Dobrovie-Sorin, Carmen. 2003. From DPs to NPs: A bare phrase structure account of genitives, in M. Coene and Y. D'Hulst. eds. *From NP to DP, Vol. 2: The Expression of Possession in Noun Phrases*. Amsterdam: John Benjamins. 75-120.

Dobrovie-Sorin, C., T. Espinal and T. Bleam. 2006 Bare nouns, number and types of incorporation, In S. Vogeleer & L. Tasmowski (eds.), *Non-definiteness and plurality*. Amsterdam: John Benjamins, Linguistics Today 95, 51-79.

Embick, D. 2010 *Localism versus Globalism in Morphology and Phonology*. Cambridge, Mass: MIT Press.

Engelhardt, M. 2000. The Projection of Argument-taking Nominals, *Natural Language and Linguistic Theory 18*. 1. 41-88.

Frege, G. 1892/1960. On Sense and Reference, in *Translations From the Philosophical Writings of Gottlob Frege*. Edited and translated by P. Geach and M. Black. 56-78. Oxford: Basil Blackwell.

Gleitman, L. 1990. The structural source of verb meanings, *Language Acquisition*, 1, 3-56.

Grimshaw, J. 1981. Form, function and the language acquisition device. In *The logical problem of language acquisition*, C.L Baker and J.J. McCarthy. (eds.) Cambridge, Mass: MIT Press.

Grimshaw, J. 1990. *Argument Structure*. Cambridge: MIT Press.

Halle, M. and A. Marantz. 1993. Distributed morphology and the pieces of inflection, in K. Hale and S.J. Keyser. eds. *The View from Building 20*. Cambridge, MA: MIT Press, 111-176

Hazout, I. 1991. Action Nominalizations and the Lexicalist Hypothesis, Ph.D dissertation, University of Massachusetts, Amherst.

Kiparsky, P. 1982 Lexical Phonology and Morphology, In In-Seok Yang. ed., *Linguistics in the Morning Calm*, Seoul, 1982.

Marantz, A. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon., in Dimitriadis, A., L. Siegel, C. Surek-Clark and A. Williams. eds., *Proceedings of the 21st Annual Penn Linguistics Colloquium. University of Pennsylvania Working Papers in Linguistics 4.2*, 201– 225. Philadelphia, PA: Department of Linguistics, University of Pennsylvania.

Marantz, A. 2000. Roots: The universality of roots and pattern morphology, Paper presented at the conference on Afro-Asiatic languages, University of Paris VII.

Massam, D. 2001. Pseudo Noun Incorporation in Niuean. *Natural Language and Linguistic Theory* 19.

Munn, A. 1995. The possessor that stayed close to home. In: Samiian, Vida and Jeanette Schaeffer. eds. *Proceedings of the Western Conference on Linguistics. WECOL 24*. 181-195.

Nunberg, G., I. Sag and T. Wasow. 1994. Idioms. Language 70: 3, 491-593

Ouhalla, J. 2009 Variation and Change in Possessive noun Phrases: The Evolution of the Analytic Type and Loss of the Synthetic Type, *Baall,. Brill's Annual of Afro-Asiatic* pp. 311-337

Pinker, S. 1984. *Language Learnability and Language Development*. Cambridge, Mass: MIT Press.

Prince, A. 1973. *The Phonology and Morphology of Tiberian Hebrew*. Ph.D. dissertation, MIT.

Ritter, E. 1988. A Head-Movement Approach to Construct-State NPs. *Linguistics* 26.6. 909-929.

Ritter, E. 1991 Two functional categories in noun phrases: evidence from Modern Hebrew, *Syntax and Semantics* 25 37-62

Shlonsky. U. 1990. Construct State Nominals and Universal Grammar. Ms, University of Haifa.

Siloni, T. 1996. Hebrew Noun Phrases: Generalized Noun Raising. In A. Belletti and L. Rizzi. (eds.) Parameters and Functional Heads: Essays in Comparative Syntax. Oxford: Oxford University Press. 239-267.

Siloni, T. 2001. Construct states at the PF interface, in Pica, P. and J. Rooryck (eds.) *Linguistic Variations Yearbook. Vol. 1*. Amsterdam: John Benjamins. 321-333.

Snyder, W. 2001 On the nature of syntactic variation: Evidence from complex predicates and complex word-formation, *Language* 77:324-342.

Swart, de H. and J. Zwarts. 2009. Less form – more meaning: Why bare singular nouns are special. Lingua 119.2, 280-295.

May 2012

h.borer@qmul.ac.uk