Giorgio Francesco Arcodia*

Coordinating nominal compounds: Universal vs. areal tendencies

Abstract: Coordinating compounds, i.e. complex word forms in which the constituent lexemes are in a coordination relation, may be divided into two classes: hyperonymic, in which the referent of the whole compound is the "sum" of the meanings of the constituent lexemes (Korowai *yumdefól* '(her) husband-wife, couple'; van Enk, Gerrit J., & Lourens de Vries. 1997. *The Korowai of Irian Jaya: Their language in its cultural context*. Oxford: Oxford University Press: 66), and hyponymic, where the compound designates a single referent having features of all the constituents (English *actor-director*). It has been proposed that languages choose either type as the one with the "tightest" marking pattern; whereas the crosslinguistic tendency is to have tighter hyperonymic compounds, most languages of Europe rather have tighter hyponymic compounds (Arcodia, Giorgio Francesco, Nicola Grandi, & Bernhard Wälchli 2010. Coordination in compounding. In Sergio Scalise & Irene Vogel (eds.), *Cross-disciplinary issues in compounding*, 177–198. Amsterdam & Philadelphia: John Benjamins). In this paper, we will test this assumption on noun-noun compounds in a sample of 20 Standard Average European languages and in a balanced sample of 60 non-SAE languages, arguing that the preference for hyperonymic compounds is best explained by the default referential function of nouns; in hyponymic compounds, on the other hand, nouns are used to indicate properties. We will then compare nominal and adjectival coordinating compounds, showing that for the latter the hyponymic compounding pattern is the dominant one, as adjectives are prototypical property-denoting words.

Keywords: coordination, compounding, co-compound, standard average European, dvandva

---

*Corresponding author: Giorgio Francesco Arcodia,** Dipartimento di Scienze Umane per la Formazione "R. Massa", Università degli Studi di Milano-Bicocca, Edificio U6, Piazza dell'Ateneo Nuovo 1, 20126 Milano, Italy, E-mail: giorgio.arcodia@unimib.it

1 Introduction

Coordinating compounds may be loosely defined as complex word forms in which all of the constituent lexemes (typically, two) are in a coordination relation.[1] Despite the controversies on the definition and delimitation of the category of coordinating compounds, there appears to be some consensus on a separation between two (macro-)classes of coordinating compounds: those in which the referent of the whole compound is the sum of the meanings of the constituent lexemes, as Korowai *yumdefól* '(her) husband-wife, couple' (van Enk and de Vries 1997: 66), and those in which the compound designates a single referent having features of both (or all of the) constituents, as English *actor-director*. It has been proposed that the former be termed "hyperonymic coordinate compounds" (henceforth: HYPERs), as the referent of the compound is more general than the constituents themselves, and the latter "hyponymic coordinate compounds" (henceforth: HYPOs), as the meaning of the whole is more specific than the meaning of the parts (Arcodia et al. 2010); they roughly correspond to exocentric and endocentric coordinating compounds in Bisetto and Scalise's (2005) classification. Moreover, HYPERs have also been termed "co-compounds" in the literature (see Wälchli 2005), a label we shall be using also in this paper. It has been claimed that there is an important semantic correlate of this distinction: HYPERs typically express natural coordination, i.e. a coordination relation between two items that often occur together and form a "conceptual unit", whereas HYPOs tend to express accidental coordination, i.e. a relation between "items which are not expected to co-occur, and which do not have a close semantic relationship" (Wälchli 2005: 5).

The different nature of these two subtypes of coordination seems to be reflected in the structural coding of compounds: natural coordination is associated with "tighter" coding patterns, whereas accidental coordination is associated with "looser" patterns. According to Arcodia et al. (2010), if a given language has coordinating compounds, they are expected to belong to the hyperonymic type, whereas hyponymic coordination should be found in "looser" constructions, as binomials,[2] phrases, or phrase-like compounds. A case in point is Modern Greek:

---

1 On the definition of coordination, see Haspelmath (2004: 33–37) among others.
2 "Binomial constructions are generally defined as constructions that consist of two (or sometimes more) coordinated items that belong to the same lexical category, are linked by a conjunction and display a certain degree of conventionality and fixity" (Masini 2006: 2).

(1)   a. *maxeropíruna*
        knife-LNK-fork-N.NOM.PL[3]
        'cutlery'
        (Manolessou and Tsolakidis 2009: 27)
      b. *sinthétis*   *tragoudistís*
        songwriter-M.NOM.SG   singer-M.NOM.SG
        'songwriter-singer'
        (Manolessou and Tsolakidis 2009: 25)

Modern Greek HYPERs (1a) behave like "true" compounds, being a combination of stems with a linking vowel and inflection only on the righthand constituent, and constituting a phonological word, whereas hyponymic constructions as (1b) may receive internal inflectional markers and are two phonological words, but behave as syntactic atoms and cannot have an independent syntactic reference, thus lying somewhere between words and phrases (Ralli 1992; Gavrilidou 2013).

This marking pattern, which iconically reflects the semantic relation between the constituents (i.e. natural – tighter, accidental – looser), is apparently reversed in the languages of the Standard Average European (SAE) *Sprachbund*: for instance, Italian coordinating compounds mostly belong to the hyponymic type, as e.g. *attore cantante* 'singer actor', and they represent a tighter construction type than, say, binomials as *marito e moglie* 'husband and wife', by means of which hyperonymic coordination is often expressed. Moreover, the distribution of coordinating compounds productively formed by means of the hyponymic pattern within SAE appears to have many similarities with that of features which usually identify this linguistic area.[4] This pattern of compounding is the default option for coordinating compounding in languages placed in the core area of the SAE linguistic area (French, German, Italian, etc.), whereas in some peripheral languages as Basque or Greek (see above, Example (1)) HYPERs represent a tighter construction type. Hence, Arcodia et al. (2010) suggest that languages choose either the hyperonymic or the hyponymic relation

---

3 The glosses follow the general guidelines of the *Leipzig Glossing Rules*; the only additional gloss is LNK (linking vowel). For the sake of simplicity and convenience, we shall provide complete morpheme-by-morpheme glosses only when relevant; in all other cases, we will just provide a translation of the constituents/words. Also, in order to avoid confusion between orthographic hyphens and glossing hyphens, we did not segment words, and we used morpheme boundary hyphens only in the glosses. When no source is provided, the examples come from texts or from the author's own knowledge of the language at issue.
4 In this paper, we define a pattern as "productive" if it can be used to form new words when necessary; this is what Bauer (2001b) calls "availability" (but see *infra*, fn. 16).

as the one with the tightest marking; whereas the former appears to be the default choice and is iconically motivated, the latter could be yet another typologically unusual feature of Western European languages (see Dahl 1990 on the 'exoticness' of SAE). However, Arcodia et al. acknowledge that their hypothesis is based on limited data, and that "a proper typological sample should be investigated to verify this claim" (2010: 196). Moreover, no explanation for this (supposedly) skewed distribution of types of coordinating compounds is offered.

In this paper, we will first try to assess Arcodia et al.'s claim; to this end, in the first part of the paper, we will propose a survey of noun-noun coordinating compounds in a sample of European languages and in a balanced sample of 60 non-SAE languages. Nominal compounds seem most suited for our survey, as they arguably represent the most common pattern in the world's languages (see Bhat 1994, Bhat 2000), and they can be deemed as the most representative type of coordinating compounds (in terms of semantic subtypes; see Wälchli 2005).

In the second part of the paper, we will try to provide an explanation for this skewed distribution of marking patterns for coordination; specifically, we will focus on the referential function of nouns, and we will discuss the relationship between nominal and adjectival coordinating compounds. We shall argue that both areal diffusion and the semantics and (prototypical) function of nouns, as opposed to other word classes, can be used to explain the relative dominance of different construction types: HYPERs combine referent-denoting lexemes, whereas HYPOs combine property-denoting lexemes. Since the prototypical function of nouns is to denote a referent, rather than a property, the tendency is for noun-noun coordinating compounds to be more often HYPERs, and for adjective-adjective compounds to be more often HYPOs; the reverse is possible, but it is much less common. Also, we shall briefly introduce some aspects of the genesis of coordinating compounds which are relevant for our discussion, especially as far as the diffusion of the hyponymic compounding pattern is concerned; given that our focus is, essentially, synchronic, we will limit ourselves to some brief observations and suggestions here.

The paper is organized as follows. Firstly, in Section 2 we will define more precisely the object and scope of our research, discussing the classification we chose to use in the present work. In Section 3, we will introduce our sample and methodology for data collection. In Section 4, we will highlight the main findings and correlations which emerge from the analysis of our samples. In Section 5, we shall introduce our comparison between nominal and adjectival coordinating compounds, and we shall propose our explanation for the distribution of different construction types in the languages of our samples, providing also some tentative remarks concerning the possible pathways for the genesis and

diffusion of different types of coordinative compounding in Europe (and beyond). Section 6 will be devoted to a summary of the main points of the paper, with some hints for further research.

2 The classification of coordinating compounds

Coordinating compounds were first identified by Indian grammarians, specifically referring to Sanskrit *dvandva* compounds as *vātavarṣaḥ* 'wind and rain' (Kiparsky 2010: 302); the term has later been extended to include other types of coordinating compounds not occurring in Sanskrit (ten Hacken 2000). Currently, the exact definition of the scope of coordinating compounding and the division into subclasses is a matter of debate, as hinted at above (see the summary in Scalise and Bisetto 2009; Arcodia 2010).

For the purposes of the present study, a coordinating compound is defined as a word-like construction made of lexemes which have the same status, as opposed to compounds whose constituents have an asymmetrical relation (as, e.g. *swordfish*). This seemingly straightforward division between symmetry and asymmetry, however, is not always so clear-cut. For instance, basing (again) on the traditional terminology of Sanskrit grammar, Bauer (2001a: 698–699) separates *dvandva* from *karmadhāraya* compounds, seen as a subtype of "determinative" (*tatpuruṣa*) compounds; among *karmadhāraya* compounds, he includes both clearly asymmetrical compounds as *blackbird*, apparently symmetrical structures as *fighter-bomber* and more ambiguous cases as *woman doctor*, lying somewhere in between asymmetry ('a doctor belonging to the female gender') and symmetry ('someone who is both a woman and a doctor'), but much closer to the former than to the latter (see also Radimský 2015).

We thus admit that the borderline between symmetry and asymmetry may be blurry, and that not everybody might agree on the inclusion into (or exclusion from) the category of coordinating compounds for some items. Following Arcodia et al. (2010: 189, fn. 7; see also Bauer and Tarasova 2013), we use as a test to distinguish coordinating and attributive (i.e. with a head and a modifier) compounds reversibility of the constituents: if a compound is found with both orders of the constituents without any significant difference in meaning, then it means that it cannot be coordinative (and *woman doctor* fails the test; for more English examples, see Olsen 2014: 272–273). This is easily exemplified with the Italian compound *studente lavoratore* 'student-worker':

(2) [...] *la figura dello **studente lavoratore***[5]
ART.F.SG status-SG of.ART.M.SG student-SG worker-SG
'the status of student-worker'

(3) [...] *il **lavoratore studente** ha diritto*
ART.M.SG student-SG worker-SG have.PRES.3SG right
all' orario flessibile [...]
to.ART.M.SG schedule-SG flexible-SG
'[...] the worker-student has the right to a flexible working schedule [...]'
(Amoroso et al. 2009: 379)

However, the variation of the order of constituents is not necessarily as free as it may seem at first glance: while (2) is an excerpt taken from university regulations, in which the status of "student" is more important, in a sense, (3) is taken from a textbook on labor law (Amoroso et al. 2009), in which the status of 'worker' is more important; hence, the constituent with relative "pragmatic prominence" over the other is in the canonical (left-hand) position for the head in Italian compounds (and see e.g. Grossmann 2012: 151 for Romanian). Again, this shows that the distinction between "pure" symmetry and asymmetry may be blurred, even if only at the pragmatic level.

Moreover, it is important to stress that although reversibility may be used to distinguish coordinating from non-coordinating constructions, it is not a requisite for the definition of coordination; on the contrary, HYPERs are often non-reversible. See the following Cantonese examples[6]:

(4) a. *fu6mou5*
father-mother
'parents'
b. [??]*mou5fu6*
mother-father

The fixity of order in co-compounds is generally motivated by lexicalization/conventionalization, or by pragmatic reasons. There are often general preferences for a specific order of constituents (Wälchli 2005: 103–104); for instance, in Cantonese, as is generally the case for Sinitic languages, there is a strong (if not absolute) preference for male-female and elder-younger orders. Both

---

5 http://www.ec.unipi.it/test-di-accesso-di-servizio/item/406-studente-lavoratore.html (accessed 23 February 2016).
6 On the reversibility of co-compounds, see Wälchli (2005: 104, 218).

lexicalization and pragmatic preferences are anyway unrelated to the structural reasons which make the order of constituents irreversible in asymmetrical compounds (i.e. the order of head and modifier in a given language; see Wälchli 2005: 103). Generally speaking, it is often the case that HYPOs are less lexicalized/conventionalized than HYPERs, hence the "freer" order; for the very same reason, HYPOs in many languages may be recursively expanded (see below, Sections 4.1, 4.2).

As to the distinction between hyperonymic and hyponymic coordination, we use as a rough test the inclusion of the referents of the constituent lexemes in the set of entities denoted by the construction as a whole. Take, for instance, the Khmer (Cambodian) compound *tok tuu* 'table closet, furniture': both 'table' and 'closet' are obviously included in the set of entities denoted by furniture, and the compound *tok tuu* is thus an instance of HYPER. This applies, in our understanding, also to compounds made of closely related or (near-)synonymic constituents, as Burmese *ni: lan:* 'way-road, method' (Soe 1999: 23). However, in the case e.g. of English *singer-songwriter*, one cannot say that either 'singer' or 'songwriter' are included in the set of entities denoted by *singer-songwriter*: all tables are pieces of furniture, but not all singers are *singer-songwriters* (i.e. there are indeed singers which are not also songwriters). That is, *singer-songwriter* is not a superordinate concept with respect to *singer* or *songwriter*.

However, just as for the distinction between symmetry and asymmetry, here some unclear cases may be found too. Is a compound as Albanian *deledash* 'ewe-ram, hermaphrodite' (Newmark et al. 1982: 175) an instance of HYPER or of HYPO? Wälchli (2005: 162) labels those forms as "intermediate-denoting compounds", further exemplified by Eng. *southwest*, which denote an entity "intermediate or hybrid between A and B". Although *hermaphrodite* somehow denotes something that is, in a way, a sum of two entities, if we choose to rely on our test, it becomes obvious that *deledash* is a HYPO, since an 'ewe' is not part of the set of 'hermaphrodites'; compare the Japanese compound *shiyū* 'female.animal-male.animal, male and female animals', in which two constituents with a meaning comparable to those of *deledash* indicate a proper superordinate set. And what about compound names of countries or regions as *Austria-Hungary, Baden-Württemberg*, etc., which are commonly found in SAE languages? Wälchli (2005: 7–8) believes that they do not denote a superordinate entity (i.e. they are not HYPERs), but rather an entity on the same taxonomical level as its constituent parts. On the other hand, basing again on our test, we could either say that *Austria* and *Hungary* are both part of the entity denoted by the compound, or that *Austria-Hungary* denotes a unique referent, rather than a set of entities, and hence the test does not apply (on *Austria-Hungary*-type compounds, see also Shimada 2013). Hence,

when it comes to the interpretation of such cases in terms of the HYPER/HYPO distinction, some possibly questionable choices had to be made.

Lastly, despite the fact that compounds are commonly understood as lexemes, we use the (admittedly vague) term "word-like construction" (following Wälchli 2005: 1), as we are not implying that all the forms we label as "coordinating compounds" are words, at least in the prototypical sense (Ramat 2005). For instance, Romance compounds, including coordinating compounds, show internal inflection, violating the cohesion of the structure, as in the following Spanish example[7]:

(5) *poetas-pintores*
     poet-PL-painter-PL
     'poet-painters'
     (Rainer and Varela 1992: 125)

In his study of co-compounds, Wälchli (2005: 3) points out that "[t]here are very few languages where compounds are undoubtedly words". For instance, he provides examples of Erźa Mordvin co-compounds which are not really words; both constituents are inflected, just as in (5), and each bears word stress, thus holding "an intermediate position between words and phrases" (2005: 1–2):

(6) *t'et'a.t-ćora.t*
     father-PL-son-PL
     'father and son'
     (Wälchli 2005: 1)

Among the languages we considered, a similar situation (i.e. coordinating compounds which are not necessarily words) holds e.g. for Kobon (Davies 1981), Dom (Tida 2006) and Tzeltal (Polian 2006).

We thus adhere to a continuum view of the boundary between morphology and syntax, in which there may be more word-like (in this case, compound-like) and less word-like items, with a blurred boundary between words and phrases (Lieber and Štekauer 2009: 14). Given that our primary concern here are tighter vs. looser marking patterns, we will focus on that, rather than on wordhood *per se*, as a criterion: for instance, although the constituents of the Italian compound *deputato-conduttore* 'MP-tv host' are both inflected for (singular) number,

---

7 Booij (2009: 88) argues that this is not a case of a syntactic rule (namely, number agreement) manipulating a constituent of a word, but, rather, the output of a morphological rule. Since the issue is not directly relevant for our argumentation, we shall not discuss it further here.

they represent a tighter marking pattern than a binomial as e.g. *bianco e nero* 'black and white', having a conjunction *e* 'and' marking the coordination relation. Moreover, double marking of a feature, as in (5) and (6), is interpreted by Wälchli (2005: 54) as iconic for the expression of a symmetric relation (namely, coordination). Our "gradient" approach, hence, differs to some extent from that of Arcodia et al. (2010), who stress the distinction between morphological and syntactic strategies to express the two subtypes of coordination at issue here (see the critique in Olsen 2014: 273–274).

Thus, the constructions which we considered in our survey include, ranging from tighter to looser:

a. (more or less) clear cases of compounding, as the above mentioned (1) Modern Greek *maxeropíruna* 'knife-fork, cutlery', having inflection only on the righthand constituent, and constituting a single phonological word (with only one stress);

b. compound-like asyndetic[8] combinations of lexemes which have internal inflection, may be reversible and may not be phonological words, like HYPOs in many SAE languages, and co-compounds in some languages (e.g. Mordvin; see 6, above).

c. compound-like combinations with syndetic marking of coordination, as Japanese *moderukenhaiyū* 'model-and-actor', generally reversible, and phrase-like combinations with syndetic marking but defective morphosyntax (e.g. no articles even when required by general syntactic principles), like binomials, which normally do not constitute a single phonological word, and which may or may not be reversible.

This classification deserves further explanation: The criteria we employ refer to, basically, morphological and phonological cohesiveness, traditionally used to establish wordhood vs. phrasehood, but we also differentiate between relational and non-relational marking, and between symmetric and asymmetric marking, following Wälchli (2005). Group a. constructions represent the tightest ones, in our understanding, as they show lexical integrity (no internal inflection) and phonological integration (a stress pattern comparable to that of words), and they tend to have a fixed order of constituents; that is, they are both morphological and phonological words. Constructions in group b., on the other hand, are less cohesive, as they do allow internal inflection and may not be phonological words; an additional criterion is that of reversibility of the constituents, which points towards a lower degree of lexical integration between the

---

8 Note that by "asyndetic" I mean 'without an overt marker of coordination', whereas by "syndetic" I mean "having an overt marker of coordination," as e.g. binomials.

constituents, and, more generally, of lexicalization/conventionalization, as said above. The presence of overt morphology between the constituents, and the low degree of prosodic integration may all be interpreted as increasing the "looseness" of the construction. Constructions in group c., on the other hand, may be regarded as even looser, as they have an overt marker of coordination, which is not expected in tight coordination (Wälchli 2005: 45); also, the pattern is asymmetrical, as the marker appears only once, rather than on both constituents (compare *supra*, 5–6), and is thus counter-iconic for coordination (2005: 56). As to reversibility, whereas compound-like syndetic combinations tend to behave as HYPOs, i.e. they are generally reversible, binomials may be divided into subclasses, ranging from totally fixed to reversible (Masini 2006).

The fact that at least a subclass of binomials have rigid constituent order, but we still consider them as a looser type than the constructions in group b., which are mostly reversible, might be perceived as an inconsistency in the application of our criteria. However, note that there do not appear to be hard-and-fast principles at work here; reversibility of an item depends, among other things, from the degree of lexicalization/conventionalization, which may vary even for items in the same class. Hence, we assigned more relative weight to the presence of an overt coordinator, which is, in our opinion, a stronger indicator of looseness, as it increases the distance between the coordinands. Reversibility will be used as an auxiliary criterion, *ceteris paribus*, as we shall see below (Section 4.1).

3 Sample and methodology

As stated in the introduction, one of the main aims of this paper is to test the tentative claims put forth by Arcodia et al. (2010) on the areally skewed distribution of tighter *vs*. looser marking patterns for natural and accidental coordination, and on the formal distinction between hyperonymic and hyponymic coordination. To this end, we created two samples:
a. a sample consisting of 20 languages of Europe
b. a balanced sample of 60 non-SAE languages.

The structure and motivations behind the selection of languages deserve some explanation. As to the first sample, 17 of the 20 languages chosen belong to the Romance, Germanic, Balto-Slavic groups or the Uralic family, or are located in the Balkans (here, Albanian and Greek); these are the core SAE languages, as per Haspelmath's (2004) definition; the remaining three, namely

Basque, Maltese and Turkish, do not belong to any of these genetic/areal groupings,[9] but are generally considered marginal members of the broadly defined European linguistic area, as they anyway share some SAE properties (see, e.g. the samples in Bernini and Ramat 1996; Mauri 2008). Being peripheral members of the area, their behavior with respect to the coding of coordination is not necessarily expected to be consistent with the core European languages. Also, we excluded languages from the Caucasus region, which have been considered in the non-SAE sample (see below).

Note that the European sample is not meant to be balanced as to the representation of branches and sub-branches of each family/group; it has been designed primarily with representativeness (i.e. coverage of branches/areas) and convenience in mind. That is, we included at least one language for each major branch (Germanic, Celtic, etc.), but we did not try to choose a more or less equal number of languages from each branch (see Table 1 below).

Needless to say, the abundance of published analyses of word formation in the languages of Europe, as well as the wealth of data and our relative familiarity with most of these languages allowed us to perform a rather thorough survey of coordinating structures. Unfortunately, this is not (always) true for the languages of the second sample, as we shall see below.

The 60-language sample was built following the principles set out by Dryer (1989, 1992), which have been designed specifically to avoid not only genetic, but also areal biases. He proposes that languages be divided into "genera", i.e. "genetic groups roughly comparable to the subfamilies of Indo-European, like Germanic or Romance" (Dryer 1989: 267). These genera should be then divided into five large continental areas, increased to six in Dryer (1992), which are assumed to be independent of one another: namely, Africa, Eurasia, Southeast Asia-Oceania, Australia-New Guinea, North America and South America. However, we depart significantly from his first premise: we did adopt his areal subdivision, but we chose individual languages *from different genera*,[10] rather than genera themselves (see Table 2 below). So, for instance, we drew data from Modern Hebrew or Indonesian, rather than considering the whole genera "Semitic" and "Sundic". This is because, while his method works quite well (despite the uncertainties as to what constitutes a "genus"; see Dryer 1992: 84, fn. 2) for the investigation of phenomena as word order, when it comes to

---

9 Needless to say, Turkish might also be included in the Balkan region.

10 We used as a basis the list of genera in the appendix of Dryer (1992), but we had to expand it to include a few other groupings of comparable level of genetic depth. For instance, *Tsezic* was not included in Dryer's sample, but is on the same level as his genus *Nax* (i.e. Nakh, a branch of the Northeast Caucasian family).

compounding the differences among individual languages within a genus may be too significant to simply lump them together (the same sampling technique has been used in Bauer's 2001a typological study of compounding).

Moreover, we did not pick freely among Eurasian languages: in order to have a proper control sample, we excluded European languages (see Table 2 below). This, however, meant that we had to make choices as to "borderline cases," like Armenian and Georgian. Although Armenia and Georgia are sometimes included into Europe, and both languages have been considered in European typological samples, e.g. for the EUROTYP and MEDTYP research projects (see Mauri 2008: 14; the same goes for Nakh), we decided to include them in our Eurasian (non-SAE) sample, while Turkish, another borderline case, was included in the European sample. Although these choices could be questioned, we believe that no significant distortion of the data may derive from this, as, again, we expect that the preference for HYPOs be limited to the core of the SAE area.

Within the principles described above, the choice of languages has been motivated primarily by convenience. We gave preference, first and foremost, to languages for which we have personal expertise (e.g. Mandarin, Japanese), and hence for which the analysis of firsthand data was easier; we then chose the languages for which we could find the best data in terms of quantity and quality. Nevertheless, the nature and quality of data are far from being uniform: we simply have better data for some languages than for others (see the appendix). For some languages, we found specific descriptions of morphology and/or compounding (e.g. Mapudungun; Baker and Fasola 2009; Zúñiga 2014), whereas for many others we had only "ordinary" grammars (e.g. Pipil; Campbell 1985). Some grammars had sample texts appended (e.g. Lele; Frajzyngier 2001), some had none; also, the number, length and variety of sample texts varied considerably. Lastly, for some languages, dictionaries or glossaries were available (e.g. Somali; Puglielli 1998), whereas for other languages nothing of the sort was available.

The quality and coverage of descriptions is especially crucial for the topic at issue: compounds are sometimes overlooked in grammatical descriptions, and coordinating compounds are even more prone to be neglected. Furthermore, HYPOs are often less lexicalized/conventionalized than HYPERs, and grammar and dictionaries may fail to record them. Take, for instance an Italian HYPO as *aperitivista-cocktailista* 'aperitifist-cocktailist': it has only one Google hit,[11] i.e. it is a *hapax legomenon*, and as such it is unlikely to be ever included in a

---

11 http://www.corriere.it/cronache/12_agosto_13/rito-happy-hour_ccacb8ec-e513-11e1-97d9-de28e70d5d31.shtml (accessed 29 February 2016).

dictionary, lexicon or grammar; there are many more HYPOs which have very limited attestation. On the other hand, HYPERs may be attested/frequent only in specific registers or styles, and hence go unnoticed (Wälchli 2005: 9, 18).

Thus, a *caveat* to be kept in mind is that, as usual, presence is easier to assess than absence: while we may say that a specific language has HYPERs, HYPOs, etc., for many languages we cannot claim with certainty that a given type of compound does not exist; it may just be the case that our sources do not mention them, or that they are rare and hence not easy to find in texts. On the other hand, we also believe that it is unlikely for a fully fledged class of compounds to go unnoticed; hence, when they are neither mentioned nor found, they should be, at best, very rare and unproductive.

Note, also, that (to the best of our knowledge) there is no established methodology available for the crosslinguistic study of HYPOs. Whereas HYPERs may be expected to be found in central lexical domains such as 'people', 'parents', 'siblings', etc. (see Wälchli 2005), as they express coordination between items which are indeed *expected* to cooccur, there appear to be no comparable domains for (nominal) HYPOs.[12] For the languages for which first-hand data was available and accessible (e.g. Hungarian, Turkish, Vietnamese, etc.), we tried to look up pairs which frequently occur in European languages, as e.g. *actor-singer* and *deaf-mute*; needless to say, this method is far from ideal, and it was not possible to apply it to all the languages of our sample. Hence, as usual, the results should be taken with a grain of salt.[13]

The languages in both samples have been grouped according to the distribution and marking patterns of coordinating compounds. If we combine the variables discussed earlier, the logical possibilities for classification are the following (we assigned a short name to each group for the sake of convenience):

> HYPER-languages. These are languages which have at least some examples of HYPERs; if HYPOs are also present, they are the instantiation of a looser construction type.
> HYPO-languages. These are which have at least some examples of HYPOs; if HYPERs are also present, they are the instantiation of a looser construction type.

---

12 I would like to thank an anonymous reviewer for this suggestion.

13 An anonymous reviewer pointed out that HYPOs might be a feature of written language, and hence be more prominent in the European sample because the languages included all have a long written history. However, in our non-SAE sample a number of languages with an established literary tradition (e.g. Hebrew, Malayalam, Khmer, etc.) are included, and yet almost no HYPOs have been detected. It is true that the few instances of HYPOs that we found come from languages with a written tradition, but it is unclear whether this is a significant correlation or just a coincidence.

HYPER/HYPO-languages. These are languages which have both HYPERs and HYPOs which are not formally distinguished (i.e. they are not the instantiation of different construction types).

UNCLEAR-languages. These are languages in which there are no (or but a couple of) clear cases of coordinating compounds, meaning that the pattern either does not exist or is extremely rare and unproductive.

NEITHER-languages. These are languages in which no coordinating compounds are attested.

Our expectations are: a. no (or almost no) HYPO/HYPER-languages anywhere; HYPO-languages dominating the core European area; c. the rest of the world split between HYPER-languages and the remaining two groups (i.e. UNCLEAR and NEITHER).

4 Analysis of the data

Arcodia et al.'s (2010) main claim, as said above, is that SAE languages have an "unusual" behavior when it comes to the coding of coordination relations: whereas the crosslinguistic tendency is that of using a tighter construction for hyperonymic coordinating structures, European languages tend to use a tighter construction for hyponymic coordinating structures. Let us firstly examine our European sample to assess this claim.

4.1 The European sample

The analysis of the European sample yielded both expected and unexpected results. On the one hand, the distribution of HYPOs is highly consistent with Arcodia et al.'s proposal. HYPOs are found in all of the core languages of the area, and hyperonymic coordination in these languages is usually expressed by means of binomials, as in the following German examples (and see above for Italian, Section 1):

(7) a. *Dichter-Maler-Komponist*
        poet-painter-composer
        'poet-painter-composer'
        (Neef 2009: 396)
     b. *Berg   und Tal*
        mountain and valley
        'mountain and valley'
        (Olsen 2014: 272)

According to Neef (2009: 396; see also Becker 1992: 27–29), compounds as *Dichter-Maler-Komponist* may be easily separated from asymmetric compounds as they have stress always on the last constituent, they lack linking elements and are typically written with hyphens. The core European languages, thus, belong to the HYPO-type.

And, again as expected, HYPERs tend to be the tightest construction type in the periphery of SAE; that is, we find mostly HYPER-languages. We have already seen the case of Greek above in (1), which has "true" HYPERs, but in which HYPOs have an intermediate status between words and phrases, being thus an instantiation of a "looser" construction type. For Basque, a non-Indo-European peripherical language within Europe, almost only HYPERs are mentioned in our source, the exceptions being *kafesnea* 'coffee-milk, coffee with milk' and *ur-ardoa* 'water-wine, water with wine' (Hualde 2003: 352). Albanian (8) and Turkish (9), two more languages on the edge of the SAE area, seem to have similar classes of HYPOs and HYPERs, in apparent violation of Arcodia et al.'s (2010) proposed generalization that hyperonymic and hyponymic coordination be distinguished:

(8) a. *kafe-restorant*
        coffee.shop-restaurant
        'coffee shop-restaurant'
        (Newmark et al. 1982: 175)
     b. *deledash*
        ewe-ram
        'hermaphrodite'
        (Newmark et al. 1982: 175)
     c. *veshmbathje*
        dress-pants
        'clothing'
        (Newmark et al. 1982: 175)

(9) a. *oyuncu-yönetmen*
        actor-director
        'actor-director'
     b. *ana(-)kız*
        mother-daughter
        'mother and daughter'
        (Göksel 2009: 221)
     c. *ana(-)baba*
        mother-father
        'parents'

As to Albanian, it may be the case that some HYPOs are calques from SAE languages: there are but a couple of such forms in our source (the other being *post-telegraf-telefon* 'post-telegraph-telephone'), and they both appear to be made of loanwords. With a cursory Google search we found two more examples of the sort, namely *bar-bufe* 'bar-buffet' and *divan-krevat* 'sofa-bed'. On the other hand, most examples of HYPERs in the literature are the product of the combination of two verbs, rather than nouns, and they are seemingly indistinguishable from HYPOs as (8b). Unfortunately, we could not find any information as to the stress patterns of compounds, and hence we don't know if HYPERs have a tighter prosody (e.g. a single stress) than HYPOs.[14]

In Turkish, HYPERs appear to be the productive model as well (Göksel 2009; on the productivity of co-compounding in Turkish, see also Wälchli 2005), but HYPOs as (9a) may indeed be found. Other examples include e.g. *sağır-dilsiz* 'deaf-mute' and *prens piskopos* 'prince-bishop', which look like calques from a SAE language (the latter almost for sure is); one should not forget the strong influence of French on Modern Turkish.[15] As to the relative tightness of the structures, there are conflicting accounts: Kasar (2004: 273) suggests that all hyphenated and separated word pairs are not compounds; Göksel (2009), on the other hand, states that pairs as *ana(-)kız* 'mother and daughter' (9b) are indeed compounds, as they follow the righthand stress pattern of Turkish words. However, *ana(-)kız* is significantly different from *ana(-)baba* 'parents' in (9c): whereas the latter actually means 'parents' and can be used as a general term, *ana(-)kız* always indicates a specific pair of mother and daughter, not a superordinate category. That is, while you can say *ana-baba olmak kolay değil* 'it's not easy to be parents', a sentence as [??]*ana(-)kız olmak kolay değil* 'it's not easy to be mother and daughter' is odd at best (İsa Sarı, p.c.); in other words, the constituents of *ana(-)kız* retain their separate reference. In point of fact, the plural form of *ana(-)kız* is *analar(-)kızlar*, with separate number marking on each constituent; this is never the case for Turkish compounds (compare *ana(-)baba**lar***). Hence, only forms as in (9c) appear to be HYPERs, whereas constructions as in (9b) are, at best, reduced syntactic structures.

Number marking patterns cannot be used to distinguish between HYPOs and HYPERs, as they both get the plural marker on the rightmost constituent

14 The fact that all "SAE-type" HYPOs are hyphenated, whereas "native" HYPOs and HYPERs are mostly (but not always) written as a single word might be interpreted as suggestive of greater phonological integration for the latter. However, as is known, spelling patterns for compounds are often inconsistent and quirky, so no real inference may be drawn in the absence of prosodic data.

15 I would like to thank an anonymous reviewer for pointing this out to me.

only (*oyuncu-yönetmenler, sağır-dilsizler, prens piskoposlar*). However, a HYPO as *oyuncu-yönetmen* is not conventionalized and fixed (the constituents may be reversed, and, also, other constituents may be added between the two members, as, e.g. *oyuncu şair yönetmen* 'actor-poet-director'); *ana*(-)*baba*, on the other hand, has a fixed structure (İsa Sarı, p.c.). Hence, HYPOs may be said to be the instantiation of a looser construction than "typical" HYPERs; in other words, Turkish does have European-style HYPOs, but the distribution of marking patterns follows the non-SAE model, with tighter HYPERs.

Another interesting situation is that of the Uralic languages of our sample. In these languages, HYPERs appear to be the "older" pattern with few attestations, mostly lexicalized, whereas HYPOs resembling those of Western European languages seem to be productive to some extent. See the following Finnish and Hungarian examples:

(10) a. *ohjaaja-näyttelijä*
          director-actor
          'director-actor'
          (Niemi 2009: 239)
      b. *maailma*
          land-air
          'world'
          (Wälchli 2005: 191)

(11) a. *színész-rendező*
          actor-director
          'actor-director'
      b. *hírnév*
          news-name
          'reputation, glory'
          (Wälchli 2005: 191)
      c. *nadrágszoknya*
          trousers-skirt
          'culottes'
          (Gouesse 2004: 137)

HYPOs as in (10a) appear to be productive in Finnish (Niemi 2009), and are identical to the Hungarian type exemplified in (11a). On the other hand, HYPERs as in (10b) and (11b) are considered to be "isolated lexicalizations [...] that suggest that co-compounding was more common at earlier stages of Finnic and Hungarian" (Wälchli 2005: 207). In point of fact, the productivity of

HYPERs in Old Hungarian has been highlighted in studies on historical word formation (Forgács 2015), and HYPERs become more common in Uralic languages located more to the East of Eurasia (as the above mentioned Erźa Mordvin; see Wälchli 2005: 206–207). It thus appears that HYPERs represent the older pattern for Uralic, whereas HYPOs might have been introduced due to contact with Western European languages; note that calquing of SAE (specifically, German) compounds has been quite common in Hungarian since the late eighteenth century (Forgács 2015; see below, Section 5.2). Also, consistently with Arcodia et al.'s (2010) prediction, HYPOs as (10a) and (11a) appear to be the instantiation of a looser pattern than HYPERs: both in Finnish and in Hungarian, HYPOs have two main (word-level) stresses, and the constituents inflect separately (compare e.g. *hírnevek* 'reputations').

However, borders become blurred when it comes to forms as (11c): *nadrágszoknya* has only one main accent, and inflection appears only on the rightmost constituent. A hint towards its looser nature is the reversibility of its constituents (*szoknyanadrág*), which may be interpreted also as a sign of lack of cohesion of the construction. Note, also, that this is the only example of this kind which we found, and we believe it may be interpreted as an isolated case of lexicalization.

A situation similar to that of Uralic is found also in the two Baltic languages Latvian (12) and Lithuanian (13):

(12) a. *lekcija-koncerts*
         lecture-concert
         'lecture-concert'
      b. *diennakts*
         day-night
         'day, 24-hour period'
         (Wälchli 2005: 205)
      c. *tẹ̀vs   mãte*
         father mother
         'parents'
         (Larsson 2002: 222)

(13) a. *dažytojas-tinkuotojas*
         painter-plasterer
         'painter-plasterer'
      b. *vištgaid is*
         chicken-rooster-M.SG.NOM
         'capon; homosexual'
         (Dabašinskienė 2010: 52)

HYPERs as (12b) are lexicalized forms; they are rare and seemingly unproductive (Larsson 2002; Navickaitė-Klišauskienė 2016; Stundžia 2016). Baltic HYPOs (12a, 13a), on the other hand, closely resemble their counterparts in SAE languages to the West: they are generally hyphenated and both members of the compound are inflected (e.g. for number). On the other hand, an example as (12c) is also attested, a hyperonymic construction made of two independently inflected words; according to Larsson (2002), this type of asyndetic coordinating construction is common in Latvian. Given that asyndetic coordination (and subordination) is possible in Latvian syntax (Mathiassen 1997: 208, 214), this is arguably the construction type which is closest to syntax, although they are not "looser" than HYPOs as (12a). Since compound-like constructions of the hyperonymic type in Baltic may have different degrees of cohesiveness (see Larsson 2002: 221–222), no general conclusion may be drawn on HYPERs as a class; we may however conclude that HYPOs, apparently, do form a fairly consistent class, with the same features as HYPOs in other SAE languages, and at least a subset of HYPERs represent a tighter construction type in these languages. An item as (13b), a HYPO with strong integration of the constituents (*višta* 'chicken' and *gaidys* 'rooster') and a conventionalized, non-transparent meaning should be a very lexicalized form, clearly tighter than (13a) (and compare 11c).

Modern Greek, as said above (1a-b), appears to have a very clear distinction between HYPERs and HYPOs which fits perfectly with the generalizations being tested here: HYPERs represent a tighter construction type, and HYPOs represent a looser construction type (i.e. it follows the "non-SAE" pattern, as a peripheral language in the area). Moreover, "loose" compounds of the HYPO class are said to have appeared only after the mid-twentieth century, and to have been coined under the influence of French (Manolessou and Tsolakidis 2009: 24). However, the actual picture is not as neat, since we also have cases like:

(14)   a. *psigío-katapsíktis*
           fridge-N.NOM.SING-freezer-M.NOM.SING
           'fridge-freezer'
           (Gavrilidou 2013: 23)
       b. *psigiokatapsíktis*
           fridge-LNK-freezer-M.NOM.SING
           'fridge-freezer'
           (Gavrilidou 2013: 23)
       c. *isagogés-eksagogés*
           importation-N.NOM.PLUR-exportation-M.NOM.PLUR
           'importations and exportations'
           (Gavrilidou 2013: 13)

Thus, a HYPO as 'fridge-freezer' in (14a) also lexicalized as (14b), thus becomes as tight as a HYPER; moreover, an instantiation of the looser construction type as (14c) actually indicates a superordinate concept. Hence, even a "well-behaved" language as Greek has counterexamples to the generalizations proposed by Arcodia et al. (2010).

The last peripheral language in our sample is Maltese, an Afro-Asiatic language whose vocabulary has been heavily influenced by Italo-Romance languages (and English). Noun-noun compounding has limited productivity in Maltese (as is generally the case in Semitic), and there appear to be no HYPERs at all; the examples of HYPOs we could find are:

(15) a. *student-ħaddiem*
         student-worker
         'student-worker'
      b. *mara   raġel*
         woman   man
         'hermaphrodite'
         (Borg and Azzopardi-Alexander 1997: 298)

However, Fabri (2009) provides the following examples:

(16) a. *mara   raġel*
         woman   man
         'woman with male features'
         (Fabri 2009: 214)
      b. *raġel mara*
         man   woman
         'effeminate man'
         (Fabri 2009: 214)

The examples in (16a)–(16b) clearly belong to the same type as Eng. *woman doctor*, i.e. they are not true coordinating structures. Pluralization for those constructions requires inflection of both constituents, whereas, generally speaking, attributive compounds receive plural marking only on the (left-hand) head constituent; on the other hand, gender agreement is controlled by the left-hand constituent for all of those forms (Fabri 2009: 215). The same appears to be true also for a clear HYPO as *student-ħaddiem* 'student-worker' (15a), for which English and/or Italian influence is highly likely; we could not find data on the inflection pattern of (15b).

The results for our European sample are summarized in Table 1.

Table 1: Our European sample.

Language	Family/group	Type
Albanian	Indo-European, Albanian	HYPER
Basque	Vasconic (isolate)	HYPER
Czech	Indo-European, Slavic	HYPO
English	Indo-European, Germanic	HYPO
Finnish	Uralic, Finnic	HYPER
French	Indo-European, Romance	HYPO
German	Indo-European, Germanic	HYPO
Greek	Indo-European, Greek	HYPER
Hungarian	Uralic, Hungarian	HYPO
Italian	Indo-European, Romance	HYPO
Latvian	Indo-European, Baltic	HYPER
Lithuanian	Indo-European, Baltic	HYPER
Maltese	Afro-Asiatic, Semitic	HYPO
Polish	Indo-European, Slavic	HYPO
Romanian	Indo-European, Romance	HYPO
Russian	Indo-European, Slavic	HYPO
Spanish	Indo-European, Romance	HYPO
Swedish	Indo-European, Germanic	HYPO
Turkish	Turkic, Southern	HYPER
Welsh	Indo-European, Celtic	UNCLEAR

Thus, what emerges from our survey is that HYPO-languages actually dominate the core area of SAE.[16] HYPERs are quite common in peripheral languages as Basque, Greek, Turkish and (perhaps to a lesser extent) Albanian, and they are present as relics in Uralic and Baltic; in most of these languages, HYPOs are also present, and it appears that at least some of them are modelled after Western European constructions. In the peripheral languages which have both HYPERs and HYPOs, the formal distinction between the two is mostly neat; however, we did find a number of cases which represent potential counterexamples to Arcodia et al.'s generalization, although some of them are but isolated and lexicalized items.

However, note that, again in partial contradiction with Arcodia et al.'s proposed generalization, the fact that HYPOs dominate Western Europe does not entail that no HYPERs are attested, and they may be formally undistinguishable from HYPOs. See the following Italian examples:

---

16 Note, however, that this does not entail that HYPOs are equally productive (here, *profitable*; see Bauer 2001b) in all languages in the area: for instance, Štichauer (2009) suggests that HYPOs have very limited productivity in Czech.

(17) a. *radio-televisione*
         radio-television
         'radio-television'
         (Radimský 2015: 122)
      b. *paternità-maternità*
         fatherhood-motherhood
         'fatherhood and motherhood, parenthood'
         (Radimský 2015: 337)
      c. *teoria-prassi*
         theory-practice
         'theory and practice'
         (Radimský 2015: 338)

While some of those may be said to be variant forms of binomials, for many there is no corresponding looser form (as e.g. 17a; Radimský 2015: 122). Radimský analyzed a corpus of 372,361 lemmatized Noun-Noun combinations drawn from the ItWac corpus,[17] and found only about 80 of them; as he points out, Italian HYPERs are "peripheral to other compounding schemas" (2015: 123).

Another case in point is Russian: although HYPERs are almost absent from Standard Russian, they have been more common in some dialects and specific text genres (Wälchli 2005). Russian is particularly interesting because HYPOs and HYPERs tend to be in complementary distribution, as they appear in different varieties of the language.[18] Comparable (and comparably uncommon) examples of HYPERs may be found also e.g. in English, as *love-hate* (as in [...] *the* love-hate *between Mozart and his father*; Schonberg 1987: 38), in French, as *histoire-geographie* 'history and geography' (as a school subject; Villoing 2012: 45), etc.

All in all, what emerges from our survey of the European data is that the generalizations at issue here are largely correct: HYPOs are a productive model in the core area of SAE, and they are mostly formally distinguished from HYPERs; the relative tightness of the marking patterns is often reversed in the peripheral languages of the area. Interestingly, tighter HYPERs may be found also in varieties of European languages spoken outside Europe, as Indian English *wife-children* and *father-mother* ('parents'; Wälchli 2005: 1), as well as in European-lexified creoles and pidgins, as Macao Creole Portuguese *pai-mai* 'father-mother, parents'; moreover, they are also attested in sign languages of Europe, as e.g. Czech Sign Language BROTHER^SISTER 'siblings' (Richteróva

---

17 http://wacky.sslmit.unibo.it/doku.php (accessed 8 March 2016).
18 I would like to thank an anonymous reviewer for pointing this out to me.

et al. 2016: 189). Nevertheless, there are some counterexamples which deserve our consideration. We shall get back to this in 5.2.

4.2 The non-SAE sample

If the generalizations at issue here are correct, we expect to find in our non-SAE sample only HYPER-, UNCLEAR- and NEITHER-languages, as said above. In point of fact, these expectations are mostly met. 30 out of the 60 languages in our sample do not appear to have coordinating compounds (the NEITHER-type), there are three UNCLEAR-languages and 26 HYPER-languages, as e.g. Mandarin:

(18) a. *dāoqiāng*
          sword-spear
          'sword and spear, weapons'
      b. *yǎnyuánjiāndǎoyǎn*
          actor-and-director
          'actor-director'

HYPERs as (18a) are abundant in Mandarin, and are produced by the juxtaposition of lexical morphemes; on the other hand, HYPOs as (18b) generally require an explicit coordinator, *jiān*.

On the other hand, one language in the sample, namely Armenian, shows features of the HYPER/HYPO-type:

(19) a. *goγavazak*
          thief-bandit
          'thief-bandit'
          'the category of thieves and bandits'
          (Donabédian 2004: 12)
      b. *ʸeresp'oxan-k'aγak'apet*
          MP-mayor
          'MP-mayor'
          (Donabédian 2004: 13)
      c. *ayruji*
          man-and-horse
          'chivalry'
          (Donabédian 2004: 15)

    d. *gišer-c'erek*
      night-day
      'day and night'
      (Dum-Tragut 2009: 671)
    e. *hayr*   *u*   *mayr*
      mother and father
      'mother and father'
      (Dum-Tragut 2009: 671)

According to Donabédian, *goɣavazak* (19a) could mean either 'thief-bandit' or 'the category of thieves and bandits'; in other words, it could be interpreted either as a HYPO or as a HYPER. Moreover, it could also be read as an asymmetric compound ("un bandit qui est comme un voleur"; Donabédian 2004: 12). On the other hand, compounds as (19b) are seemingly taken to be just HYPOs ("composés cumulatifs", in Donabédian's words), and they are said to be calques, rather than "autochtonous" compounding patterns (< French *député-maire*? Donabédian 2004: 13). A hyperonymic coordinating constructions with univerbation of the conjunction as (19c) may also be found; however, it is claimed to be a rarity. Besides, asyndetic HYPERs as (19d) and (seeming) binomials as (19e) are attested in the language as well;[19] note that 'mother and father' is a typical lexical domain for HYPERs crosslinguistically (Wälchli 2005: 9), but it is dominated by binomials or looser constructions in SAE languages. As said above, Armenian is a language that is often included in Europe for the purposes of areal sampling, but it is very peripheral with respect to SAE; hence, it is perhaps not so surprising that this is about the only language which clearly defies all generalizations, given its transitional nature between Europe and Asia. Incidentally, the fact that "proper" HYPOs are said to be calques provides support for this analysis; a genetic (Indo-European) origin for this pattern appears very unlikely, especially since, as we shall say below (Section 5.2), the hyponymic pattern appears to be fairly young in the family. Further evidence could come from diachronic data on Armenian word formation.

    Some other cases deserve further discussion. For Hebrew, no coordinating compounds are mentioned in our sources, but we did find a couple of blends and acronyms which might be interpreted as coordinating structures:

---

19 Whereas the commonly used coordinating conjunction *ew* "refers to a connection of coordinate objects or qualities (...)", *u* "expresses a closer range between the connected objects, qualities, etc." (Dum-Tragut 2009: 290).

(20) a. *midraxá + rexóv = midrexóv*
        pavement   street
        'promenade'
        (Schwarzwald 2011: 528)
    b. *sakiním + kapót + u-mazlegót = sakúm*
        knives   spoons   and-forks
        'cutlery'
        (Schwarzwald 2011: 528)

The item in (20a) is a blend of two words, and could be interpreted either as a HYPO ('something that is both a pavement and a street') or as an attributive compounds ('a pavement that functions as a street'). (20b), on the other hand, clearly has hyperonymic semantics, but it is the product of the reduction of a structure with an overt coordinator (*u*). We thus classified Hebrew as an UNCLEAR-language.

    The existence of a clear-cut structural difference between HYPERs and HYPOs has been called into question for Japanese too. Compare (21a) and (21b):

(21) a. *fū-fu*
        husband-wife
        'husband and wife'
    b. *haiyūkenkantoku*
        actor-and-director
        'actor-director'

Just as seen for Mandarin above (18a), morphemes are simply juxtaposed to produce HYPERs; on the other hand, HYPOs as (21b) are said to require the coordinator *ken* (Kageyama 2009: 514; see (18b)). However, Shimada provides examples of what he deems could be HYPOs which are the product of juxtaposition, just as HYPERs (2013; transliteration adapted):

(22) a. *sakka-tomodachi*
        writer-friend
        'writer friend'
        (Shimada 2013: 87)
    b. *shufu-gakusei*
        housewife-student
        'housewife-student'
        (Shimada 2013: 87)

Not all of these examples are equally convincing: *sakka-tomodachi* 'writer-friend' (22a), for instance, seems odd in that the two identities can hardly be seen as providing a label for a single person, as they belong to different levels of categorization; it is probably better analyzed as an attributive compound (a possibility that Shimada himself acknowledges). However, a case as, e.g. (22b) does indeed look like a genuine HYPO. Whether these are true HYPOs, and whether they represent a productive class, is unclear; the fact that they are hardly ever mentioned in the literature on Japanese word formation suggests that this is but a marginal pattern of compounding.

As Japanese has no productive noun inflection, we cannot use the *locus* of marking as a test to distinguish between the patterns exemplified in (21a) and (22b). It is worth pointing out again that, as said above, in most languages HYPERs appear to be more lexicalized than HYPOs, and only the former are normally found in dictionaries; moreover, HYPERS are mostly made of bound roots, whereas HYPOs are made of words. Kageyama (2009: 515) remarks that HYPERs (or, better, a subclass of HYPERs, which includes items as (21a)) have a characteristic pitch pattern; this is yet another factor which may help us differentiate the two classes.

Another HYPER-language, which however also admits apparent counter-examples, is Vietnamese:

(23)   a.   *chim-muông*
            bird-beast
            'animals'
            (Nguyen 1997: 66)
       b.   *hoàng tử giám mục*
            prince   bishop
            'prince   bishop'

HYPERs like (23a) are commonly found in Vietnamese, and they are the only type of coordinating compound mentioned in our sources. An asyndetic hyponymic combination as *hoàng tử giám mục* 'prince bishop' (23b), while attested, is about the only clear example of HYPO we could find, and it is undoubtedly a calque from a European language. Moreover, although the notion of "stress" may not easily apply to Vietnamese word-level prosody, there appears to be a perceived "juncture" between *hoàng tử* 'prince' and *giám mục* 'bishop', whereas HYPERs as *chim-muông* are prosodically tighter (Alexis Michaud, p.c.). Thus, we believe that it should not be taken as a challenge to the classification of Vietnamese as a language in which HYPERs are the basic and tighter type of coordinating compound.

The results for our non-SAE sample are summarized in Table 2.

Table 2: Our non-SAE sample.

Area	Language	Genus	Type
Africa	Modern Hebrew	Semitic	UNCLEAR
	Lele	East Chadic	NEITHER
	Tamazight	Berber	NEITHER
	Degema	Edoid	NEITHER
	Kanuri	Saharan	NEITHER
	Somali	Eastern Cushitic	NEITHER
	Turkana	Nilotic	NEITHER
	Akan	Kwa	NEITHER
	Koyra Chiini	Songhai	NEITHER
	Bambara	Mande	NEITHER
Eurasia	Hunzib	Tsezic	HYPER
	Armenian (M. East.)	Armenian	HYPER/HYPO
	Marathi	Indic	HYPER
	Khalkha Mongolian	Mongolian	HYPER
	Malayalam	Dravidian proper	HYPER
	Japanese	Japanese	HYPER
	Korean	Korean	HYPER
	Georgian	Kartvelian	HYPER
	Persian	Iranian	HYPER
	Ingush	Nakh	HYPER
Southeast Asia-Oceania	Mandarin	Chinese	HYPER
	Burmese	Burmic	HYPER
	Vietnamese	Viet-Muong	HYPER
	Cambodian	Khmer	HYPER
	Tagalog	Philippine Austronesian	UNCLEAR
	Indonesian	Sundic	HYPER
	Tetun	C.-E. Malayo-Polynesian	HYPER
	Meithei	Meithei	HYPER
	Lao	Kam-Tai	HYPER
	Highland Yao	Miao-Yao	HYPER
Australia-New Guinea	Warlpiri	Pama-Nyungan	NEITHER
	Kobon	E. New Guinea Highlands	HYPER
	Mangarayi	Mangarayi	NEITHER
	Korowai	Trans N. Guinea Ok Awyu	HYPER
	West Kewa	Trans N. Guinea Engan	HYPER
	Dom	Trans N. Guinea Chimbu Wahgi	HYPER
	Baruga (Tafota)	Trans N. Guinea Binanderean	HYPER
	Tauya	Madang	NEITHER
	Kamasau	Torricelli	NEITHER

Table 2: (*continued*)

Area	Language	Genus	Type
	Wardaman	Gunwingguan	HYPER
N. America	Mohawk	Iroquoian	NEITHER
	Koasati	Muskogean	NEITHER
	Sochiapan Chinantec	Chinantecan	HYPER
	Chimalapa Zoque	Mixe-Zoque	NEITHER
	Purapecha	Tarascan	NEITHER
	Tzeltal	Mayan	HYPER
	West Greenlandic	Eskimo-Aleut	NEITHER
	Jamul Tiipay	Yuman	NEITHER
	Slave	Athapaskan	NEITHER
	Pipil	Aztecan	NEITHER
S. America	Maká	Maká	NEITHER
	Imbabura Quechua	Quechua	UNCLEAR
	Hup	Nadahup	NEITHER
	Wari/Pacaas Novos	Chapacura-Wanham	NEITHER
	Baure	S. Arawakan	NEITHER
	Jaqaru	Jaqaru	NEITHER
	Maipure	N. Arawakan	NEITHER
	Tsimané	Mosetenan	NEITHER
	Jarawara	Arawan	NEITHER
	Mapudungun	Mapudungu	NEITHER

As to the areal tendencies concerning the distribution of the five types identified above within our samples, Table 3 provides a region-by-region overview:

Table 3: The distribution of types in our samples.

Area	HYPER	HYPO	H/H	UNCLEAR	NEITHER
Africa	0	0	0	1	9
Eurasia (excl. SAE)	9	0	1	0	0
Southeast Asia-Oceania	9	0	0	1	0
Australia-New Guinea	6	0	0	0	4
North America	2	0	0	0	8
South America	0	0	0	1	9
Europe	**7**	**12**	**0**	**1**	**0**

The distribution of HYPER-type languages is largely consistent with Wälchli's remarks on the skewed areal distribution of co-compounding (2005: 235–236):

> [T]here is reason to believe that co-compounds may emerge more or less everywhere, even if high areal restrictions to Eurasia, New Guinea, and a few other places suggests that it is extremely rare for areally unconditioned evolution to give rise to a fully-fledged lexical class of co-compound.

Wälchi (2005: 237–240) further adds that New Guinea is "the next largest co-compound area", whereas co-compounds are not characteristic of the languages of Australia and Africa; in the Americas, co-compounds may be expected in Quechua and in some Meso-American and Mayan languages. And, as a matter of fact, this is the situation we found in our sample: Southeast Asia - Oceania and (Eastern) Eurasia clearly emerge as the areas with the highest concentration of HYPERs, followed by the Australia-New Guinea area; on the other hand, HYPERs (and, more generally, coordinating compounds) are virtually absent from the African and South American languages of our sample, and are at best marginal in North America (and see also the data in Bauer 2001a: 699).

To sum up, the data from our sample proves that Arcodia et al.'s (2010) generalizations on the distribution of marking patterns of compounding and, incidentally, Wälchli's (2005) remarks on the areal distribution of co-compounding are mostly correct. The most serious counterexample which we found, i.e. Armenian, could be argued to be the product of a "double" areal influence, being at the border between Europe and Asia.

5 A possible account for the distribution of patterns of coordination in compounding

As said above, our investigation yielded both expected and unexpected results: while there appears to be a very strong divide between languages with tighter HYPOs and languages with tighter HYPERs, and the division is clearly areal, there are also some cases which defy the generalizations being tested here. In this section, we shall address the question of the motivations both for this skewed distribution of marking patterns for coordination in morphology and for some of the apparent exceptions; also, we shall briefly discuss the diachronic side of the development of the patterns at issue.

5.1 On the preference for HYPERs in the world's languages

The fact that having HYPOs as the tightest construction type appears to be a Western European singularity, whereas in the overwhelming majority of

languages with coordinating compounds the combination of two nouns triggers a hyperonymic interpretation, is indeed striking. The explanation we propose for this strong crosslinguistic preference, as hinted at above (Section 1), is based on the prototypical function for nouns, i.e. reference, as opposed to modification. Let us start from a comparison with adjective-adjective coordinating compounds, made of prototypically property-denoting words.

As said in the introduction, we restricted our survey to nominal compounds; when adjectival compounds come into the picture, the proposed areal division between HYPERs and HYPOs gets blurred. Although we did not carry out a thorough survey of adjective-adjective coordinating compounds, a stunning fact emerges even from a cursory examination of the languages in our sample; namely, adjectival HYPOs, besides HYPERs, may be easily found in a typical co-compounding language, as e.g. Mandarin:

(24) a. *chángduǎn*
          long-short
          'length'
      b. *suānlà*
          sour-spicy
          'hot and sour'

Example (24a) is a typical example of the so-called "scalar" type of HYPERs, which denote "an abstract scale with opposite qualities A and B as extreme poles" (Wälchli 2005: 152); (24b), on the other hand, is just the predication of two properties attributed to the same referent (as in *suānlà tāng* 'hot and sour soup'). Hence, *suānlà* is more specific than either *suān* 'sour' and *là* 'spicy'; it is a hyponymic compound. On the formal level, (24a) and (24b) are indistinguishable; ambiguity does not generally arise, we suggest, because HYPERs made of adjectives typically involve properties which are incompatible (as 'long' and 'short'), and thus are not expected to be predicated of the same referent, whereas adjectival HYPOs involve adjectives which can be compatible (food can be sweet and sour at the same time). We have two homonymous constructions, but only one, i.e. the hyponymic one, seems to have no special restrictions on its input. Thus, HYPOs can be the "default" type for coordinate adjectival compounds even in typical co-compounding languages.

In SAE languages, adjectival coordinating compounds are typically HYPOs, unsurprisingly, although the coding of combination of adjectives as (tight) compounds, as binomials or just as phrases varies much, both intra- and cross-linguistically (compare German *schwarzweiß* 'black and white', meaning both 'being part black and part white' and 'grayscale' and *bittersüß* 'bittersweet',

compounds with inflectional markers appearing only on the righthand member, to French *noir et blanc*, Italian *bianco e nero*, but *aigre-doux, dolceamaro*). In Modern Greek, a language with tighter nominal HYPERs (see Example 1), adjectival compounds are HYPOs, and they are structurally as tight as nominal HYPERs (as, e.g. *glikópikros* 'sweet-bitter, bitter-sweet'; Ralli 2013: 158). The "division of labor" between compounds and binomials which we observe for nouns in SAE, hence, does not apply to adjectives.

On the other hand, if we go back to Radimský's list of Italian HYPERs (see above, Section 4.1), we may remark that nearly all the (genuine) HYPERs are the product of the combination of process (or psych) nouns: *ricerca-sviluppo* 'research and development', *scarico-carico* 'unloading and loading', *odio-amore* 'hate-love, love and hate'.[20] This is not the case for typical examples of HYPOs, both in Italian and in other SAE languages: they normally combine two concrete nouns, often professions/roles (see the several examples discussed in 4.1), tools/machines (French *canne-épée* 'stick-sword, swordstick'; Fradin 2009: 430) or places/businesses (see Fradin 2009; Grossmann 2012).

Just as in the Mandarin case discussed above, it appears that in SAE we may have two different constructions for noun-noun coordinating compounds, without any formal distinction between HYPOs and HYPERs. However, whereas in the *singer-poet* case only the hyponymic interpretation is available, when process nouns are involved, as *ricerca-sviluppo* 'research and development', the hyperonymic interpretation is perhaps the only acceptable one.[21] We suggest that the semantic features of the constituents, here roughly defined as "concrete" *vs.* "process", guide the interpretation as hyponymic or hyperonymic: this may be expected, since we typically conceive a process as something that develops in time, and hence as the combination of two processes is normally seen as either a succession (*ricerca-sviluppo*) or as something that occurs

---

20 Note, however, that in SAE languages one also finds HYPERs belonging to the class termed "alternative and approximative", like 'July-August' (meaning 'July or August'; Wälchli 2005: 151–152), and HYPERs indicating periods of time as 'January-June' (meaning 'the period from January to June'). Here, we suggest, semantics constrains again the interpretation: since a hyponymic interpretation would not make much sense here (as in: 'something which is both January and June'), it is ruled out in favor of a hyperonymic reading.

21 An anonymous reviewer pointed out that in many languages (notably Germanic) compounds may have a number of different interpretations, and a speaker may not be sure as to how a novel compound is to be interpreted; even context plays a role. However, I believe this is not (always) true for coordinating compounds: for instance, English *actors-directors* is always understood as indicating 'people who do both acting and singing as a profession' (i.e. each one of them is an *actor-singer*). No one, I believe, could ever think that it means 'the community of actors and singers'.

alternatively (*odio-amore*), or as both (*scarico-carico*). For HYPOs, on the other hand, what normally happens is that, in essence, two properties are predicated: an *actor-singer* is someone who can both act and sing (professionally). Since properties are not so strongly connected with the temporal dimension (see Langacker 1987), the default interpretation is the simultaneous attribution of two properties to the same referent, as in *actor-singer, singer-poet*, etc.

However, as said above, HYPERs are rare in SAE languages: a binomial is the "default" construction for this type of coordination. Binomials are not equivalent to "European" HYPERs, in that the input of the former is not so restricted; in this case, however, the overt marker of coordination triggers a non-hyponymic reading, and hence no ambiguity arises. Thus, the input of HYPERs appears to be restricted mostly to process nouns (but see Note 20), for which anyway a competing construction exists, namely binomials (compare It. *ricerca-sviluppo* with *ricerca **e** sviluppo*, both 'research and development'); HYPOs and binomials, on the other hand, have less restrictions on their input. In the light of the above, the relatively low number of HYPERs in SAE should not come as a surprise. Their absence from most discussions on word formation, which (correctly) take HYPOs to be the productive model of coordinating compounding in SAE languages, further attests to their secondariness.

On the other hand, somewhat predictably, the combination of property-denoting words as adjectives is often understood as hyponymic, i.e. as simultaneously attributing two properties to the same referent. In this case, it is the hyperonymic interpretation that can be argued to be, in a sense, more marked: it is actually available only for adjectives which, as said above, denote the extreme poles of a scale and, as such, are generally semantically (and pragmatically) unsuitable for simultaneous predication. The typological markedness of scalar co-compounds has been highlighted by Wälchli (2005: 153–155), who treats them as a non-prototypical co-compounds, which "can be part of a lexical class of co-compounds only if there are also prototypical co-compounds"; in point of fact, they are "characteristic of East and South East Asian languages with a high level of co-compounding" (and, we may add, Basque; see Hualde 2003). Moreover, they are, in a sense, a "derivative" subclass of co-compounds, since they mostly derive, historically, from alternative/approximate co-compounds (see fn. 20) in questions (e.g. Classical Tibetan *mgyogs-bul* 'fast-slow, fast or slow > speed'; Wälchli 2005: 153).

Thus, to sum up, in HYPOs, what is being combined are, in essence, properties, which can be simultaneously attributed to the same referent; in HYPERs, two (or more) referents are combined to indicate one that "contains" them both (/all). Since the prototypical (unmarked) function for adjectives is "modification by a property" (Croft 2000, 2001), and, as said above, two (or more) properties

may be commonly (by default?) understood as being simultaneously attributed to the same referent, adjectival HYPOs are expected to be typologically unmarked (i.e. more common than HYPERs). Hence, HYPOs can be the default type for coordinate adjectival compounds even in "typical" co-compounding languages. Nouns, however, do not prototypically encode properties, but referents: this is arguably the basis for the most common pattern of nominal coordinating compounding, i.e. HYPERs, in which referents are joined together to indicate a "set". The hyponymic type, on the other hand, requires that the two constituents be used, in a sense, as attributes; this constitutes marked usage for nouns, a point which may help to explain both the skewed distribution of nominal HYPOs in the world's languages and the restriction to specific types of nouns, i.e. concrete nouns which may be interpreted as indicating, among others, professional or human attributes. In this respect, HYPOs are not fundamentally different from attributive compounds as *frogman* or *snail mail*: here *frog* and *snail* are used as attributes ('amphibious' and 'slow', respectively; see Scalise et al. 2005); in HYPOs, however, the relation is symmetrical, i.e. both constituents equally contribute to identifying the referent of the whole compound. In other words, whereas a *frogman* is, indeed, a man, but not a frog, an *actor-singer* is both.

Moreover, the relative markedness of nouns and adjectives in referential and modifying function may also explain the distribution of tighter *vs.* looser marking patterns for coordination. Being a composite notion, typological markedness includes the notion of structural markedness (Croft 2001: 90): that is, when a noun is used to refer to an entity, it should not require *more* overt morphemes than when used to indicate a property (its non-prototypical function); the opposite should be true for adjectives. Hence, having tighter nominal HYPERs and tighter adjectival HYPOs means that the coordination of referent-denoting nouns, on the one hand, and of property-denoting adjectives, on the other hand, are structurally less marked than the coordination of property-denoting nouns, as in nominal HYPOs; the preference for HYPER-languages is hence also consistent with the predictions of structural markedness. In contrast, having a tighter structure for hyponymic coordination of nouns, as opposed to hyperonymic coordination of nouns, as in SAE, is not consistent with the predictions of structural markedness, and hence is expected to be dispreferred. Of course, this applies only to specific functional/relational morphology, i.e. to overt markers of coordination, whereas our parameters for tightness include much more (see above, Section 2); however, equal marking (i.e. no overt coordinator either in hyponymic or in hyperonymic coordination) is anyway consistent with the predictions of structural markedness.

A residual issue concerning the motivations behind the preference for HYPERs is that of the relation between the nature of the coordination relation, in terms of natural *vs.* accidental coordination (see the introduction), and the other semantic/pragmatic distinctions employed in our analysis. In other words, we may want to ask whether the association between hyperonymy, referent-denoting words, tightness and natural coordination, on the one hand, and hyponymy, property-denoting nouns, looseness and accidental coordination, on the other hand, is indeed significant, or just epiphenomenal.

As said above, natural coordination is a relationship between items which often occur together and can form a conceptual unit, whereas in accidental coordination the items involved lack this close semantic relationship, and do not often co-occur. Olsen (2014, 2015) proposes that HYPERs follow a syntactic pattern of coordination, in which the meaning of the conjoined structure "is not obtained by simply adding the meanings of the conjuncts together, but rather by subsuming the co-conjuncts under a common conceptual frame", i.e. the so-called "semantic integrator" (Olsen 2014: 280; see Lang 1991). This common conceptual frame is what we termed the hyperonym, as 'furniture' for the above mentioned Khmer HYPER *tok tuu* lit. 'table closet'. If we equate the idea of a "conceptual unit" with that of a "common semantic integrator", then the preference for natural coordination in hyperonymic compounding can be explained straightforwardly: since in HYPERs the constituents refer to a "set" or, at least, to an overarching notion including the constituent themselves (see e.g. Khalkha Mongolian *gal togoo* 'fire pan, kitchen'), it is more "natural", in a sense, for a set to be formed by frequently co-occurring items which have a rather close semantic relationship. In fact, as pointed out correctly by Olsen (2014: 280, quoting Lang 1991: 607), when the coordinands in set-forming coordination are not expected to form a set because of their weak semantic relationship, odd interpretations arise, as in the infamous example:

(25)   *No entry for dogs and Chinese*

Hence, we may expect that "natural" set-forming coordinands are more likely to be recruited and lexicalized/conventionalized as compounds. Also, items that often co-occur tend to be phonologically and prosodically more integrated (Bybee 2003), and hence more tightly connected. This does not seem to apply to HYPOs, since two unrelated or distantly related properties may well be attributed to the same entity, as English *poet-doctor* or Italian *deputato conduttore* 'MP-tv host'; we already remarked above that HYPOs tend to be less entrenched (for lack of a better term) in the lexicon of languages, and many *hapax legomena* and words with exceedingly low frequency may be found

among them. Nevertheless, it appears that more predictable combinations are in fact more common, as in the various *actor-director* examples discussed above (and see the methodological note in 3). Thus, the semantics of different subtypes of coordination, again, may be fruitfully used to explain the association between natural coordination and set-forming hyperonymic structures.

5.2 Remarks on the diachrony of coordinating compounds

Given the distribution of HYPERs and HYPOs in the languages of our samples, and considering the different semantic relations between the constituents and the compounds themselves, one may want to ask whether they may be also explained by the way these patterns were created and diffused.

As hinted at in the preceding section, Olsen (2014, 2015) believes that HYPERs follow a syntactic pattern of coordination, and resort to a syntactic interpretive strategy, and hence could originate from the lexicalization of asyndetic phrasal coordination. HYPOs, on the other hand, are words with a clearly identifiable head, and "in the case of a headed word structure, the head will pick out a category to be modified by the nonhead constituent", triggering a hyponymic interpretation as for *caricaturist-cartoonist* ('a cartoonist who is also a caricaturist'; Olsen 2014: 281); this is consistent with Bisetto and Scalise's classification, in which, as said above (1.), HYPOs are said to be endocentric.[22] Also, Olsen points out that when a SAE coordinating compound is used as a modifier, i.e. it is not in the head position of a word, one gets hyperonymic readings, as in *husband-wife team* (but cf. Wälchli 2005: 7) or *grocery-meat market*. Hence, in Olsen's opinion an opposition between syntax-like coordination and morphology-like coordination could explain the difference between SAE and the rest of the world.

The fact that HYPERs often originate from asyndetic coordination has already been pointed out before in the literature (Wälchli 2005: 245–251): in our sample, this is true e.g. for Mandarin. Olsen, quoting Uotila (1980), remarks that asyndetic coordination could represent the oldest construction in Indo-European languages, but it was replaced quite early on by syndetic coordination; this process is argued to have been stronger in the Western part of the Indo-European speaking world, and hence relics may still be found in languages to the East, as e.g. Baltic (see Example (12c) above). On the other hand,

---

22 Note, e.g. that in those (rare) cases in which the gender of the constituents in an Italian HYPO is not the same, the compound takes its gender from the left-hand constituent, as in *bar pasticceria* '[bar$_M$-pastry shop$_F$]$_M$' (Scalise et al. 2005: 138).

asyndetic coordination is more common in Uralic languages; as seen above (4.1), in this family HYPOs appear to be found only in the westernmost languages, whereas HYPERs are common in the Uralic languages spoken to the East of the SAE area, and used to be more common e.g. in Old Hungarian.

However, as pointed out by Wälchli (2005: 249–250), there is also evidence that HYPERs do not necessarily derive from phrasal coordination. For instance, he suggests that the Modern Greek co-compound *andrógyno* 'man-woman, married couple' did not develop from coordination, but rather from a *bahuvrīhi* compound (Ancient Greek *andrógynos* 'coward, hermaphrodite'). Furthermore, it cannot be excluded that the juxtaposition of e.g. two nouns occurs spontaneously, without an underlying syntactic (or morphological) model; this is the case, for instance, of some HYPERs that may be found in child language even in the SAE area (Wälchli 2005: 250). Also, as pointed out above, HYPERs can emerge even in European-lexified pidgins and creoles, attesting to their possible origin in language contact (with substrate/adstrate languages; Arcodia et al. 2010: 182–183). Lastly, the fact that HYPERs are often zero-marked means that diachronic investigation proves to be extremely difficult (Arcodia et al. 2010: 182); we may add that it becomes close to impossible when working with languages without historical documentation, as is the case e.g. for many indigenous languages of Australia-New Guinea and of the Americas. Thus, while it appears to be the case that an origin in syntactic coordination for HYPERs is likely in many cases, it is by no means universal, and often we simply lack the kind of historical evidence we would need to prove or disprove an evolution from a syntactic coordinate construction. Lastly, note that syntactic coordination may trigger a hyponymic (intersective) interpretation for property-denoting adjectives, and even nouns: compare *my friend and relative*, attributing two identities to the same referent, and *my friends and relatives*, with an additive (overlapping) interpretation (*scil.* 'all the people that are my relatives or my friends'; Wälchli 2005: 77–78). Hence, the association between syntax and hyperonymic interpretive strategies may not be as stable as it has been claimed.

Going back to the European scenario, could the lack (or, better, scarcity) of HYPERs be explained by the early loss of asyndetic coordination? Needless to say, this would require an in-depth study of the history of the marking patterns for coordination in individual Indo-European languages, which is obviously beyond the scope of the present article. Hence, as stated in the introduction, we shall limit ourselves to a few tentative remarks on the diachrony of HYPERs and HYPOs in this family.

Firstly, even a peripheral modern European language with tight HYPERs as Greek rather had productive HYPOs in the past, as e.g. *iatrómantis* 'healer-diviner' or *hippaletryôn* 'horse-rooster', whereas the hyperonymic compounding

pattern was created in the Late Medieval Period (Manolessou and Tsolakidis 2009: 28); HYPOs, hence, clearly represent the older pattern in Greek.[23] As to Latin, we do find a few examples of nominal HYPOs as *tragicomoedia* 'tragedy-comedy', but they are said to be very rare and unproductive, even at the stage of Late Latin (Brucale 2012: 107–108, 115), and hence unlikely to have provided a model for Romance; for instance, compounding was marginal in Old French, and HYPOs seem to be a relatively recent innovation (Rainer and Buridant 2015: 1977–1978). As to Germanic languages, coordinating compounds (both nominal and adjectival) were virtually unknown in Old and Middle High German, and HYPOs seem to have appeared in Modern German (Müller 2015: 1881). Old English seemingly had but a handful of HYPERs, involving pairs of relatives, as *suhtor-fædran* 'nephew and uncle' (Dietz 2015: 1918), Lass (1994: 196) cites *were-wulf* 'man-wolf, werewolf' as a possible case of HYPO, but its interpretation is unclear (coordinative or appositive?). Coordinative compounding was not productive in Middle English either (Dietz 2015).

On the other hand, Old Irish, at the western periphery of the present-day SAE area, did have HYPERs, as *cennainim* 'head and soul' (Stifter 2010: 109), whereas a modern Celtic language as Welsh uses an overt conjunction in analogous constructions (*mellt a tharannau* 'lightning and thunder'; Awbery 2004: 318). As to Sanskrit, for which we have a wealth of historical data on compounding, we know that *dvandva* compounds (i.e. HYPERs) originate from asyndetic coordination (see Kiparsky 2010; Olsen 2014); on the other hand, compounds as *rājarṣiḥ* 'king-sage, a royal sage' or *narasiṃhaḥ* 'man-lion' are generally considered to be attributive (*karmadhāraya*) compounds, with the left-hand constituent acting as a modifier, just as e.g. Eng. *frogman* (Goldman and Sutherland-Goldman 1999: 212–213; see above, 5.1).

Hence, with the notable exception of Greek, it does seem true that, in the past, peripheral languages of Europe (as Old Irish or Baltic) and Indo-European languages outside Europe (as Sanskrit) did have HYPERs, possibly as remnants of an earlier strategy of asyndetic coordination; on the other hand, for the hyponymic compounding pattern all the (limited) evidence available points towards a more recent origin in Western European languages, again with the exception of Greek, which went the opposite way (i.e. from HYPOs to HYPERs). Then, the SAE pattern apparently spread eastwards to the periphery of Europe (Baltic, Uralic and Turkish), and even beyond (Armenian).

---

23 A possible early example of HYPER in Ancient Greek is the word *nychthêmeron* 'night-day, 24-hour period'. It seems unlikely, however, that it represents some relic of an even earlier productive pattern (although this possibility may not be ruled out).

Interestingly, according to Rainer and Buridant (2015: 1978), Modern French nominal HYPOs derive from "appositions" (e.g. Molière's *bourgeois gentilhomme* 'bourgeois gentleman'), i.e. from attributive compounds. It might then be the case that HYPOs in SAE languages are actually the product of a reanalysis of morphological, headed structures as symmetric, double-headed structures; the fact that the borderline between the two types of compounds, i.e. attributive and hyponymic coordinative, is still sometimes blurred even in contemporary SAE languages, lends further support to this hypothesis.

6 Conclusions and hints for further research

In the present paper, we first tried to assess Arcodia et al.'s hypotheses on the distribution of marking patterns for hyponymic and hyperonymic coordination in the lexicon. Our data supports an areal division between SAE and the rest of the world, in that only the former area seems to have HYPOs as the tightest construction type, and the tendency is stronger in the core area of SAE. The most serious counterexample to the generalization according to which languages choose either HYPERs or HYPOs as the tightest construction type is Armenian, which has been argued to be a case in point, given its transitional nature between Europe and Asia.

As to the explanation for this skewed distribution, we argued that, apart from the areal connection, it may be (partly) explained also with some inherent semantic and functional features of nouns: since the prototypical function for nouns is "reference to an object", a coordinating noun-noun compound is typically understood as indicating a set minimally consisting of the two constituents, i.e. a superordinate concept. On the other hand, an adjectival compound would be understood as simultaneously attributing two properties to the same referent, as "modification by a property" is the prototypical function for adjectives; in this case, the interpretation of the whole compound is hyponymic. However, nouns can be also used to indicate properties, as we assume is the case for SAE HYPOs; this is a marked function for nouns, which may explain the limited distribution of (tight) HYPOs. On the other hand, hyponymic adjectival compounds do not show the same skewed distribution of nominal compounds, since this is their basic function. In point of fact, when two nouns are unsuitable for a property reading, a hyperonymic interpretation may be available also for SAE languages with HYPOs as the default type, as shown above (Section 5.1). Also, from the structural point of view, nominal HYPERs and adjectival HYPOs are expected to be less marked (i.e. coordination relation encoded by a lower or

equal number of morphemes) than nominal HYPOs, which is actually the case in nearly all the languages outside the core SAE area which have both HYPERs and HYPOs, whereas the reverse is true for SAE.

As to the relationship between natural coordination and HYPERs, on the one hand, and accidental coordination and HYPOs, on the other hand, we argued that this may be explained by the very function of hyperonymic coordination, which typically is to form a set, as said above; items which often co-occur and can be easily subsumed under a semantic integrator should be preferred for this type of naming units, whereas items with a weak relation should be dispreferred. The same does not appear to hold for HYPOs, since unexpected properties may be attributed to the same referent; however, if the combination is, for lack of a better term, more "plausible", we expect it to be more common, as exemplified by the *actor-director* example.

Lastly, there appears to be (limited) historical evidence in favor of a relatively recent origin for HYPOs in Western European languages, which then arguably spread eastwards; whereas HYPERs typically (but not always) derive from the lexicalization of asyndetic coordinative constructions in syntax, SAE HYPOs might be the evolution of an asymmetric compounding pattern.

To conclude, if the hypothesis sketched here is correct, the areally skewed distribution of tight HYPOs is not surprising: a marked feature (here, a word formation pattern) is expected to be less common than an unmarked one (i.e. HYPERs), but may be strongly entrenched in a family/area. Further research on the diachrony of coordinating structures in the languages of Europe is needed to assess a possible "Europeanness" of this pattern, and, anyway, to shed light on the origin and diffusion of the hyponymic compounding pattern.

Appendix. Sources used for the languages in the non-SAE sample

Language	Source(s)
Modern Hebrew	Gensler (2011), Schwarzwald (2011)
Lele	Frajzingier (2001)
Tamazight	Louali and Philippson (2004)
Degema	Kari (2004)
Kanuri	Hutchinson (1981)
Somali	Puglielli (1998), Saeed (1999)
Turkana	Dimmendaal (1983)
Akan	Appah (2009)
Koyra Chiini	Heath (1999)
Bambara	Creissels (2004)
Hunzib	van Den Berg (1995)
Armenian (M. East.)	Donabédian (2004), Dum-Tragut (2009)
Marathi	Pandharipande (1997)
Khalkha Mongolian	Molomjamts and Csillaghy (1998), Janhunen (2012)
Malayalam	Asher and Kumari (1997)
Japanese	Kageyama (2009), Shimada (2013)
Korean	Sohn (1999)
Georgian	Wälchli (2005)
Persian	Lazard (1992)
Ingush	Nichols (2011)
Mandarin	Ceccagno and Basciano (2009)
Burmese	Soe (1999)
Vietnamese	Thompson (1965), Nguyen (1997)
Cambodian	Antelme (2004), Haiman (2011)
Tagalog	Schachter and Otanes (1972)
Indonesian	Lini (1982), Sneddon (1996)
Tetun	van Klinken (1999)
Meithei	Chelliah (1997)
Lao	Enfield (2007)
Highland Yao	Court (1985)
Warlpiri	Simpson (2009)
Kobon	Davies (1981)
Mangarayi	Merlan (1982)
Korowai	van Enk and de Vries (1997)
West Kewa	Franklin (1971), Franklin et al. (1978)
Dom	Tida (2006)

(*continued*)

Language	Source(s)
Baruga (Tafota)	Farr et al. (1996)
Tauya	MacDonald (1990)
Kamasau	Sanders and Sanders (1994)
Wardaman	Merlan (1994)
Mohawk	Mithun (2009)
Koasati	Kimball (1991)
Sochiapan Chinantec	Foris (2000)
Chimalapa Zoque	Johnson (2000)
Purapecha	Chamoreau (2003)
Tzeltal	Polian (2006)
West Greenlandic	Fortescue (1984)
Jamul Tiipay	Miller (2001)
Slave	Rice (2009)
Pipil	Campbell (1985)
Maká	Tacconi (2014)
Imbabura Quechua	Cole (1985), Kirtchuk-Halevi (2004)
Hup	Epps (2008)
Wari/Pacaas Novos	Everett and Kern (1997)
Baure	Admiraal and Danielsen (2014)
Jaqaru	Hardman (2000)
Maipure	Zamponi (2009)
Tsimané	Sakel (2004)
Jarawara	Dixon (2004)
Mapudungun	Zúñiga (2001, 2014), Smeets (2008), Baker and Fasola (2009)

References

Admiraal, Femmy, & Swintha Danielsen. 2014. Productive compounding in Baure (Arawakan). In Swintha Danielsen, Katja Hanss & Fernando Zúñiga (eds.), *Word formation in South American languages*, 79–112. Amsterdam & Philadelphia: John Benjamins.

Amoroso, Giovanni, Vincenzo Di Cerbo, & Arturo Maresca. 2009. *Diritto del lavoro: Lo Statuto dei Lavoratori e la disciplina dei licenziamenti*, Vol. II, 3rd ed. Milan: Giuffrè.

Antelme, Michel. 2004. Khmer. In Pierre L. J. Arnaud (ed.), *Le nom composé: Données sur seize langues*, 149–184. Lyon: Presses universitaires de Lyon.

Appah, Clement K. I. 2009. Compounding in Akan. Paper presented at "Universals and Typology in Word-Formation", Košice (Slovakia), 18 August.

Arcodia, Giorgio Francesco. 2010. Coordinating compounds. *Language and Linguistics Compass* 4(9). 863–873.

Arcodia, Giorgio Francesco, Nicola Grandi, & Wälchli Bernhard. 2010. Coordination in compounding. In Sergio Scalise & Irene Vogel (eds.), *Cross-disciplinary issues in compounding*, 177–198. Amsterdam & Philadelphia: John Benjamins.

Asher, Ronald E., & T. C. Kumari. 1997. *Malayalam*. London & New York: Routledge.

Awbery, Gwenllian Mair. 2004. *Welsh* (gallois). In Pierre L. J. Arnaud (ed.), *Le nom compose: Données sur seize langues*, 303–328. Lyon: Presses universitaires de Lyon.

Baker, Mark C., & Carlos A. Fasola. 2009. Araucanian: Mapundungun. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 594–608. Oxford: Oxford University Press.

Bauer, Laurie. 2001a. Compounding. In Martin Haspelmath, Ekkehardt König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language Typology and Language Universals*, 695–707. Berlin & New York: Mouton de Gruyter.

Bauer, Laurie. 2001b. *Morphological productivity*. Cambridge: Cambridge University Press.

Bauer, Laurie, & Elisaveta Tarasova. 2013. The meaning link in nominal compounds. *SKASE Journal of Theoretical Linguistics* 10(3). 2–18.

Becker, Thomas. 1992. Compounding in German. *Rivista di Linguistica* 4(1). 5–36.

Bernini, Giuliano, & Paolo Ramat. 1996. *Negative sentences in the languages of Europe: A typological approach*. Berlin & New York: Mouton de Gruyter.

Bhat, Shankara D. N. 1994. *The adjectival category*. Amsterdam & Philadelphia: John Benjamins.

Bhat, Shankara D. N. 2000. Word classes and sentential functions. In Petra M. Vogel & Bernard Comrie (eds.), *Approaches to the typology of word classes*, 47–64. Berlin & New York: Mouton de Gruyter.

Bisetto, Antonietta, & Sergio Scalise. 2005. Classification of compounds. *Lingue e Linguaggio* 2. 319–332.

Booij, Geert. 2009. Lexical integrity as a morphological universal: A constructionist view. In Sergio Scalise, Elisabetta Magni & Antonietta Bisetto (eds.), *Universals of language today*, 83–100. Berlin: Springer.

Borg, Albert, & Marie Azzopardi-Alexander. 1997. *Maltese*. London & New York: Routledge.

Brucale, Luisa. 2012. Latin compounds. *Probus* 24(1). 93–117.

Bybee, Joan. 2003. Mechanisms of change in grammaticization: The role of frequency. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 602–623. Oxford: Blackwell.

Campbell, Lyle. 1985. *The Pipil language of El Salvador*. Berlin & New York: Mouton de Gruyter.

Ceccagno, Antonietta, & Bianca Basciano. 2009. Sino-Tibetan: Mandarin Chinese. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 478–490. Oxford: Oxford University Press.

Chamoreau, Nadine. 2003. *Grammaire du Purepecha*. Munich: Lincom Europa.

Chelliah, Shobhana L. 1997. *A grammar of Meithei*. Berlin & New York: Mouton de Gruyter.

Cole, Peter. 1985. *Imbabura Quechua*. London: Croom Helm.

Court, Christopher A. F. 1985. *Fundamentals of Iu Mien (Yao) Grammar*. Berkeley, CA: University of California dissertation.

Creissels, Denis. 2004. Bambara. In Pierre L. J. Arnaud (ed.), *Le nom compose: Données sur seize langues*, 21–46. Lyon: Presses universitaires de Lyon.

Croft, William. 2000. Parts of speech as language universals and as language-particular categories. In Petra M. Vogel & Bernard Comrie (eds.), *Approaches to the typology of word classes*, 65–102. Berlin & New York: Mouton de Gruyter.

Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Dabašinskienė, Ineta. 2010. Acquisition of compounds in Lithuanian. *Baltic Linguistics* 1. 51–82.

Dahl, Östen. 1990. Standard Average European as an exotic language. In Johannes Bechert, Giuliano Bernini & Claude Buridant (eds.), *Toward a typology of European languages*, 3–8. Berlin & New York: Mouton de Gruyter.

Davies, John. 1981. *Kobon*. Amsterdam: North Holland.

Dietz, Klaus. 2015. Historical word-formation in English. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe*, Vol. III, 1914–1930. Berlin & Boston: De Gruyter Mouton.

Dimmendaal, Gerrit J. 1983. *The Turkana language*. Dordrecht: Foris.

Dixon, Robert M. W. 2004. *The Jarawara language of Southern Amazonia*. Oxford: Oxford University Press.

Donabédian, Anaïd. 2004. Arménien. In Pierre L. J. Arnaud (ed.), *Le nom compose: Données sur seize langues*, 3–20. Lyon: Presses universitaires de Lyon.

Dryer, Matthew S. 1989. Large linguistic areas and language sampling. *Studies in Language* 13(2). 257–292.

Dryer, Matthews S. 1992. The Greenbergian word order correlations. *Language* 68(1). 81–138.

Dum-Tragut, Jasmine. 2009. *Armenian*. Amsterdam & Philadelphia: John Benjamins.

Enfield, Nicholas J. 2007. *A grammar of Lao*. Berlin & New York: Mouton de Gruyter.

Epps, Patience. 2008. *A grammar of Hup*. Berlin & New York: Mouton de Gruyter.

Everett, Daniel L., & Barbara Kern. 1997. *Wari: The Pacaas Novos language of Western Brazil*. London & New York: Routledge.

Fabri, Ray. 2009. Compounding and adjective-noun compounding in Maltese. In Bernard Comrie, Ray Fabri, Elizabeth Hume, Manwel Mifsud, Thomas Stolz & Martine Vanhove (eds.), *Introducing Maltese linguistics*, 207–232. Amsterdam & Philadelphia: John Benjamins.

Farr, Cynthia J. M., Bomi Terrence Furoke, & James Buyers Farr. 1996. Tafota Baruga grammar notes. https://www.eva.mpg.de/lingua/tools-at-lingboard/pdf/Farr_TBGRammarNotes.pdf.

Forgács, Tamás. 2015. From Old Hungarian to Modern Hungarian. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe*, Vol. III, 2079–2096. Berlin & Boston: De Gruyter Mouton.

Foris, David Paul. 2000. *A grammar of Sochiapan Chinantec*. Dallas: SIL International.

Fortescue, Michael. 1984. *West Greenlandic*. London: Croom Helm.

Fradin, Bernard. 2009. IE, Romance: French. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 417–435. Oxford: Oxford University Press.

Frajzyngier, Zygmunt. 2001. *A grammar of Lele*. Stanford, CA: CSLI.

Franklin, Karl J. 1971. *A grammar of Kewa, New Guinea*. Canberra: Pacific Linguistics.

Franklin, Karl J., Joice Franklin, & Yapua Kirapeasi. 1978. *A Kewa dictionary, with supplementary grammatical and anthropological materials*. Canberra: Australian National University.

Gavrilidou, Zoe. 2013. NN combinations in Greek. *Journal of Greek Linguistics* 13(1). 5–29.

Gensler, Orin D. 2011. Morphological typology of Semitic. In Stefan Weninger (ed.), *The Semitic languages: An international handbook*, 303–329. Berlin & New York: De Gruyter Moutonr.

Göksel, Asli. 2009. Compounds in Turkish. *Lingue e Linguaggio* 8(2). 213–235.

Goldman, Robert P., & Sally J. Sutherland-Goldman. 1999. *Devavāṇīpraveśikā, An introduction to the Sanskrit language*. 3rd ed. Berkeley, CA: Center for South Asia Studies, University of California.

Gouesse, Marie-Josèphe. 2004. Le nom compose: Données sur seize langues. In Pierre L. J. Arnaud (ed.), *Hongrois*, 131–148. Lyon: Presses universitaires de Lyon.

Grossmann, Maria. 2012. Romanian compounds. *Probus* 24(1). 147–173.

Haiman, John. 2011. *Cambodian: Khmer*. Amsterdam & Philadelphia: John Benjamins.

Hardman, Martha J. 2000. *Jaqaru*. Munich: Lincom Europa.

Haspelmath, Martin. 2004. Coordinating constructions: An overview. In Martin Haspelmath (ed.), *Coordinating constructions*, 3–40. Amsterdam & Philadelphia: John Benjamins.

Heath, Jeffrey. 1999. *A grammar of Koyra Chiini*. Berlin & New York: Mouton de Gruyter.

Hualde, José Ignacio. 2003. Compounds. In José Ignacio Hualde & Jon Ortiz de Urbina (eds.), *A grammar of Basque*, 351–362. Berlin & New York: Mouton de Gruyter.

Hutchinson, John P. 1981. *The Kanuri language: A reference grammar*. Madison, WI: University of Wisconsin.

Janhunen, Juha A. 2012. *Mongolian*. Amsterdam & Philadelphia: John Benjamins.

Johnson, Heidi Anna. 2000. *A grammar of San Miguel Chimalapa Zoque*. Austin, TX: University of Texas dissertation.

Kageyama, Taro. 2009. Isolate: Japanese. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 512–526. Oxford: Oxford University Press.

Kari, Ethelbert Emmanuel. 2004. *A reference grammar of Degema*. Cologne: Köppe Verlag.

Kasar, Sündüz Ö. 2004. Le nom composé: Données sur seize langues. In Pierre L. J. Arnaud (ed.), *Turc*, 269–286. Lyon: Presses universitaires de Lyon.

Kimball, Geoffrey. 1991. *Koasati grammar*. Lincoln, NE: University of Nebraska Press.

Kiparsky, Paul. 2010. Dvandvas, blocking, and the associative: The bumpy ride from phrase to word. *Language* 86(2). 302–331.

Kirtchuk-Halevi, Pablo-I. 2004. Quechua. In Pierre L. J. Arnaud (ed.), *Le nom composé. Données sur seize langues*, 221–232. Lyon: Presses universitaires de Lyon.

Lang, Ewald. 1991. Koordinierende Konjunktionen. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, 597–623. Berlin & New York: Mouton de Gruyter.

Langacker, Ronald W. 1987. Nouns and verbs. *Language* 63(1). 53–94.

Larsson, Jenny H. 2002. Nominal compounds in the Baltic languages. *Transactions of the Philological Society* 100(2). 203–231.

Lass, Roger. 1994. *Old English: A historical linguistic companion*. Cambridge: Cambridge University Press.

Lazard, Gilbert. 1992. *A grammar of contemporary Persian*. Costa Mesa, CACosta Mesa, CA: Mazda Publishers.

Lieber, Rochelle, & Štekauer Pavol. 2009. Introduction: Status and definition of compounding. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 3–18. Oxford: Oxford University Press.

Lini, Lorenzo. 1982. *Manuale di conversazione italiano-indonesiano*. Bologna: EMI.

Louali, Naima, & Gérard Philippson. 2004. Tamazight. In Pierre L. J. Arnaud (ed.), *Le nom composé. Données sur seize langues*, 233–248. Lyon: Presses universitaires de Lyon.

MacDonald, Lorna. 1990. *A grammar of Tauya*. Berlin & New York: Mouton de Gruyter.

Manolessou, Io, & Symeon Tsolakidis. 2009. Greek coordinated compounds: Synchrony and diachrony. *Patras Working Papers in Linguistics* 1. 23–39.

Masini, Francesca. 2006. Binomial constructions: Inheritance, specification and subregularities. *Lingue e Linguaggio* 5. 207–232.

Mathiassen, Terje. 1997. *A short grammar of Latvian*. Columbus, OH: Slavica Publishers.

Mauri, Caterina. 2008. *Coordination relations in the languages of Europe and beyond*. Berlin & New York: Mouton de Gruyter.

Merlan, Francesca C. 1982. *Mangarayi*. Amsterdam: North Holland.

Merlan, Francesca C. 1994. *A grammar of Wardaman, a language of the Northern Territory of Australia*. Berlin & New York: Mouton de Gruyter.

Miller, Amy W. 2001. *A grammar of Jamul Tiipay*. Berlin & New York: Mouton de Gruyter.

Mithun, Marianne. 2009. Iroquoian: Mohawk. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 564–583. Oxford: Oxford University Press.

Molomjamts, Jargal, & Andrea Csillaghy. 1998. *Introduzione Alla lingua Mongola*. Udine: Forum.

Müller, Peter O. 2015. Historical word-formation in German. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe*, Vol. III, 1867–1914. Berlin & Boston: De Gruyter Mouton.

Navickaitė-Klišauskienė, Agnė. 2016. Word-formation: An international handbook of the languages of Europe. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Latvian*, Vol. V, 3107–3124. Berlin & Boston: De Gruyter Mouton.

Neef, Martin. 2009. IE, Germanic: German. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 386–399. Oxford: Oxford University Press.

Newmark, Leonard, Philip Hubbard, & Peter Prifti. 1982. *Standard Albanian: A reference grammar for students*. Stanford, CA: Stanford University Press.

Nguyen, Dinh Hoa. 1997. *Vietnamese*. Amsterdam & Philadelphia: John Benjamins.

Nichols, Johanna. 2011. *Ingush grammar*. Berkeley, CA: University of California Press.

Niemi, Jussi. 2009. Compounds in Finnish. *Lingue e Linguaggio* 8(2). 237–256.

Olsen, Susan. 2014. Zwischen Kern und Peripherie: Untersuchungen zu Randbereichen in Sprache und Grammatik. In Antonio Machicao, Andreas Nolda Priemer & Athina Sioupi (eds.), *Coordinative structures in morphology*, 269–286. Berlin & Boston: De Gruyter.

Olsen, Susan. 2015. Composition. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe*, Vol. I, 364–385. Berlin & Boston: De Gruyter Mouton.

Pandharipande, Rajeshwari. 1997. *Marathi*. London & New York: Routledge.

Polian, Gilles. 2006. *Elements de grammaire du Tseltal: Une langue maya du Mexique*. Paris: L'Harmattan.

Puglielli, Annarita. 1998. *Dizionario italiano-somalo*. Rome: Carocci.

Radimský, Jan. 2015. *Noun+Noun Compounds in Italian: A corpus-based study*. České Budějovice: Jihočeská univerzita.

Rainer, Franz, & Soledad Valera. 1992. Compounding in Spanish. *Italian Journal of Linguistics* 4 (1). 117–142.

Rainer, Franz, & Claude Buridant. 2015. From Old French to Modern French. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe*, Vol. III, 1975–2000. Berlin & Boston: De Gruyter Mouton.

Ralli, Angela. 1992. Compounding in modern Greek. *Rivista di Linguistica* 4(1). 143–174.

Ralli, Angela. 2013. *Compounding in modern Greek*. Dordrecht: Springer.

Ramat, Paolo. 2005. Per una definizione di 'parola'. In Paolo Ramat (ed.), *Pagine linguistiche*, 106–121. Rome: Laterza.

Rice, Keren. 2009. Athapaskan: Slave. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 542–563. Oxford: Oxford University Press.

Richteróva, Klára, Alena Macuróva, & Nováková Radka. 2016. Kinship terminology in Czech Sign Language. In Ulrike Zeshan & Keiko Sagara (eds.), *Semantic fields in sign languages: Colour, kinship and quantification*, 163–208. Berlin & Boston: De Gruyter Mouton.

Saeed, John. 1999. *Somali*. Amsterdam & Philadelphia: John Benjamins.

Sakel, Jeanette. 2004. *A grammar of Mosetén*. Berlin & New York: Mouton de Gruyter.

Sanders, Arden, & Joy Sanders. 1994. *Kamasau (Wand Tuan) grammar: Morpheme to sentence*. https://pnglanguages.sil.org/resources/archives/31072

Scalise, Sergio, & Antonietta Bisetto. 2009. Classification of compounds. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 49–82. Oxford: Oxford University Press.

Scalise, Sergio, Antonietta Bisetto, & Emiliano Guevara. 2005. Selection in compounding and derivation. In Wolfgang U. Dressler, Dieter Kastovsky, Oskar E. Pfeiffer & Franz Rainer (eds.), *Morphology and its demarcations*, 97–108. Amsterdam & Philadelphia: John Benjamins.

Schachter, Paul, & Fe T. Otanes. 1972. *Tagalog reference grammar*. Berkeley, CA: University of California Press.

Schonberg, Harold C. 1987. *The great pianists: From Mozart to the present*. New York: Simon & Schuster.

Schwarzwald, Ora. 2011. Modern Hebrew. In Stefan Weninger (ed.), *The Semitic languages: An international handbook*, 523–536. Berlin & New York: Mouton de Gruyter.

Shimada, Masaharu. 2013. Coordinated compounds: Comparison between English and Japanese. *SKASE Journal of Theoretical Linguistics* 10(1). 77–96.

Simpson, Jane. 2009. Pama-Nyungan: Warlpiri. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 609–622. Oxford: Oxford University Press.

Smeets, Ineke. 2008. *A grammar of Mapuche*. Berlin & New York: Mouton de Gruyter.

Sneddon, James Neil. 1996. *Indonesian: A comprehensive grammar*. London & New York: Routledge.

Soe, Myint. 1999. *A grammar of Burmese*. Eugene: University of Oregon dissertation.

Sohn, Ho-Min. 1999. *The Korean language*. Cambridge: Cambridge University Press.

Štichauer, Pavel. 2009. Compounds in Czech. *Lingue e Linguaggio* 8(2). 293–314.

Stifter, David. 2010. Early Irish. In Martin J. Ball & Nicole Müller (eds.), *The Celtic languages*, 55–115. London & New York: Routledge.

Stundžia, Bonifacas. 2016. Lithuanian. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe*, Vol. V, 3089–3107. Berlin & Boston: De Gruyter Mouton.

Tacconi, Temis L. 2014. Towards a characterization of compounding in Maká. In Swintha Danielsen, Katja Hanss & Fernando Zúñiga (eds.), *Word formation in South American languages*, 33–50. Amsterdam & Philadelphia: John Benjamins.

Ten Hacken, Pius. 2000. Derivation and compounding. In Geert Booij, Christian Lehmann & Joachim Mugdan (eds.), *Morphologie-Morphology*, 349–360. Berlin-New York: Mouton de Gruyter.

Thompson, Laurence C. 1965. *A Vietnamese grammar*. Seattle, WA: University of Washington Press.

Tida, Syuntarô. 2006. *A grammar of the Dom language: A Papuan language of Papua New Guinea*. Kyoto: University of Kyoto dissertation.

Uotila, Eeva. 1980. Asyndeton in the Baltic and Finnic languages: An archaic construction in its typological periphery. *Journal of Baltic Studies* 11. 86–92.

van Den Berg, Helma. 1995. *A grammar of Hunzib: With texts and lexicon*. Munich: Lincom.

van Enk, Gerrit J., & Lourens de Vries. 1997. *The Korowai of Irian Jaya: Their language in its cultural context*. Oxford: Oxford University Press.

van Klinken, Catharina Lumien. 1999. *A grammar of the Fehan dialect of Tetun: An Austronesian language of West Timor*. Canberra: Pacific Linguistics.

Villoing, Florence. 2012. French compounds. *Probus* 24(1). 29–60.

Wälchli, Bernhard. 2005. *Co-compounds and natural coordination*. Oxford: Oxford University Press.

Zamponi, Raoul. 2009. Arawakan: Maipure-Yavitero. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 584–593. Oxford: Oxford University Press.

Zúñiga, Fernando. 2001. *Mapudungun*. Munich: Lincom.

Zúñiga, Fernando. 2014. Nominal compounds in Mapudungun. In Swintha Danielsen, Katja Hanss & Fernando Zúñiga (eds.), *Word formation in South American languages*, 11–32. Amsterdam & Philadelphia: John Benjamins.