Relevance of typology to minimalist inquiry.

Jan-Wouter Zwart
University of Groningen

Contribution to special issue of *Lingua*, John Nerbonne, ed.

september 2006.

ABSTRACT

This article discusses the relevance of sample based typological research to central questions of minimalist syntactic theory. The article focuses on the properties of the structure building operation Merge, in particular on the question of whether Merge combines elements in an unordered set or as an ordered pair. It is argued that Merge automatically creates a dependency which may be marked by various dependency indicators, distributed over various components of the grammar. Linear order is the least reliable of these indicators, but may be studied in the restricted domain of juxtapositions and coordinations. A case study illustrates the relevance of sample based typological research, uncovering a universal convergence of linear order and dependency marking in monosyndetic noun phrase coordinations.

Relevance of typology to minimalist inquiry.

1. Introduction

This article considers to what extent sample based typological research may shed light on questions of central concern within the minimalist program of generative grammar (Chomsky1995). To that end I first present what I take to be a core issue of minimalist inquiry, the nature of the structure building operation Merge and the properties of the structures it yields. I formulate a limited number of questions which may be answered through sample based typological research. I then present a case study illustrating the potential of sample based research in this domain. An appendix presents the sample used in these studies and discusses relevant issues of sampling technique.

2. The minimalist program

2.1 Merge

The 'minimalist program' is the latest stage of generative grammar and shares with earlier stages the idea that the object of grammatical inquiry is not (just) linguistic phenomena per se, but the human cognitive capacity generating these linguistic phenomena. It is distinguished from earlier stages of generative grammar in its attempt to reduce all properties of the syntactic component of grammar (referred to as 'narrow syntax') to requirements set by the interfaces between the computational system of human language (CHL) on the one hand and the components of the mind/brain dealing with sound (the articulatory-perceptual interface or A-P interface) and meaning (the conceptual-intentional interface or C-I interface) on the other.

A key assumption of the minimalist program is that CHL involves a system generating structure, called Merge. Merge takes two elements and combines them in a group (a constituent). It is recursive, since the output of Merge may be merged with other elements yielding a further constituent. As Chomsky states, Merge is an operation that comes 'free', in the sense that it is required in some form for any recursive system (Chomsky 2001:4).

The significance of Merge as a property of the faculty of language is underscored by the contention of Hauser, Chomsky and Fitch (2002), who hypothesize that recursion

is the only uniquely human component of the faculty of language. A strong hypothesis on the nature of syntactic relations further emphasizes the centrality of Merge in the minimalist program: as Epstein (1999) conjectures, syntactic relations are established only between elements that are combined through Merge (essentially reducing syntactic relations to the sisterhood configuration). A central question of minimalist inquiry, then, is: what are the properties of Merge, or, more pointedly: what is the simplest form that Merge could take?

There is some indication that phrase structure of natural language is invariably binary branching (cf. Kayne 1984). This can be established using standard constituency tests in most cases. For example, the bracketed ditransitive verb phrase in (1) can be shown to involve a constituent excluding the indirect object illustrated in (2) on the basis of the displacement test illustrated in (3).

(1) Jan heeft [Marie een boek gegeven] (Dutch)
John has Mary a book given
'John gave Mary a book.'

(2) [een boek gegeven]

(3) [Een boek gegeven] heeft Jan Marie __ niet
a book given has John Mary not

If binary branching is a property of all syntactic structures, Merge invariably combines no more than two elements. This much seems uncontroversial in the minimalist literature.

Chomsky (2005a), however, makes the further proposal that Merge combines the two elements it merges into an unordered set. As Chomsky (2005b) notes, there is an alternative, which he rejects as 'more complex', in which Merge creates not a set but an ordered pair. In other words, the two alternatives describe Merge as either symmetric or asymmetric.

In this article I would like to consider the relevance of sample based typological research to this central question of the nature of the structure building process Merge.

2.2 Linear order

One of the issues involved here is how the linear order characterizing linguistic expressions (in the spoken modality, a sequence of sounds ordered in time) relates to the syntactic objects created by Merge. In Chomsky's view, linear order is a property specific to the A-P interface component dealing with sound (Chomsky 2005b). This suggests that the relation between structure and order is either arbitrary or parametrized (e.g. via the head-parameter stipulating per language the order of heads and their complements). This contrasts with the proposal by Kayne (1994), in which linear order is the automatic spell-out of an asymmetric phrase structure.

In Kayne's proposal, phrase structure is asymmetric because of a global condition on possible structures, according to which all terms of a structure must be linearly ordered in the mathematical sense (i.e. the relations among terms must meet the requirements of transitivity, totality and antisymmetry; see Kayne 1994:4f for fuller discussion). A problem with this approach, pointed out by Chomsky (1995, section 4.8) is that the phrase structures needed to make the structure-order correspondence work do not meet the minimalist requirement of 'bare phrase structure' (which prohibits the vacuous projection sometimes needed in Kayne's analysis). However, the more central intuition underlying Kayne's approach, that linear order is a function of asymmetric structure, is not necessarily affected by Chomsky's critique. If Merge yields asymmetric structures (ordered pairs rather than unordered sets), we may take linear order to be an interface-particular interpretation of a more fundamental asymmetry between the members of the pair created by Merge.

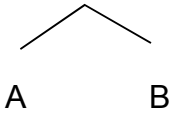
Kayne's intuition that linear order is a function of asymmetric structure was based on a range of empirical observations, mainly involving patterns that are significant by their absence (Kayne 1992). For example, whereas a number of languages show verb second effects (with the verb appearing in the position after any first clausal constituent), no languages seem to exist which display 'verb second to last' effects (with the verb showing up in the position preceding any final clausal constituent). If verb second involves movement of the verb up to a functional position (the structural position for Tense features or for the Complementizer), the observation leads to the conclusion that the functional projections which may play host to a moving verb must all be head-initial. Observations of this kind led Kayne (1994:132) to the striking conclusion "that the specifier-head-complement order is the only order made available by U[niversal]

G[rammar] and consequently that there can be no directionality parameter for word order."

Though some progress has been made in studying the empirical consequences of this conclusion, a fundamental difficulty remains that a particular linear order may be the result of direct Merge (also called 'external merge'), or of a further operation, Move (also called 'internal merge'). To illustrate, the order in (4) may be a straightforward realization of the structure in (5a), resulting from the operation in (5b), or from the structure in (6a), resulting from the sequence of steps in (6b).

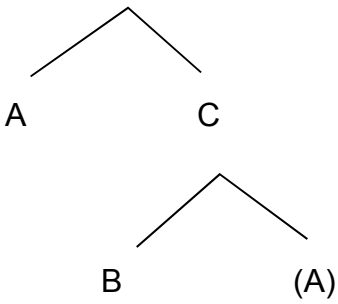
(4) / A B /

(5) a.



b. Merge A to B

(6) a.



b. Merge B to A (yielding C) (external merge)
Merge A to C (= Move) (internal merge)

This uncertaintly makes it imperative that empirical data be analysed in considerable detail before their relevance to any theoretical issue can be fully ascertained. Consequently, it becomes difficult, if not impossible, to bring to bear on these theoretical issues the large scale data collections underlying current typological research.

This illustrates the main difficulty of using large scale data collections for fundamental syntactic research: the questions we are interested in (the nature of the structure building operation Merge and the properties of the structures it yields) can only be studied if we are certain about the structural analysis of the data, which requires an in depth study of the phenomena, practically impossible in sample based typological research.

Nevertheless, I believe it is possible to address the fundamental question of the nature of Merge using large language samples if the research domain is carefully circumscribed. In particular, the research should focus on types of constructions where we are relatively certain that movement played no role in their derivation. In the next section, we consider how such 'restricted' construction types lend themselves to a potentially large scale investigation of syntactic dependency. In section 4, then, I present more detailed research on coordination as a first illustration of how sample based typological research may shed light on questions central to minimalist inquiry.

3. Dependency as a function of Merge

The noun phrase coordination phenomena to be reviewed in section 4 suggest that two elements combined by Merge invariably realize a dependency, expressed by linear order and morphology (the conjunction functioning as a linking particle). This finding may be interpreted in one of two ways. First, we may presume that Merge still operates in a symmetric fashion, merging two elements in an unordered set. The dependency then would have to be imposed at the A-P interface, as part of a spell-out procedure. Second, we may assume that Merge operates in an asymmetric fashion, creating an ordered pair rather than an unordered set. Linear order would then be one of a number of modality specific expressions of the asymmetry between the members of the ordered pair.

I argue elsewhere that there are theoretical and conceptual considerations favoring the second view (Zwart 2004). This has to do with the detailed workings of Merge, viewed as an assignment operation assigning an element from the Numeration to the current derivation. If this is the way Merge works, the output of the operation invariably consists of an 'old' element (the previous state of the current derivation) and a new element, the newly Merged object. As a result, the members of the newly created pair

are asymmetric in terms of the history of the derivation (cf. also Jaspers 1998, Fortuny Andreu 2007).

Here, however, I would like to consider a line of empirical argumentation addressing the question of symmetric vs. asymmetric Merge. The argument is based on the observation that indicators of dependency are distributed over all components of the grammar, not restricted to the A-P interface, and that these indicators tend to identify the dependent and the nondependent category in a consistent way. The relevance to the topic of this article is that these dependency indicators can be readily inspected from standard reference grammars.

3.1 Dependency

In the structuralist tradition, dependency is generally defined in terms of phrase structure, capitalizing on the head/nonhead distinction (e.g. Tesnière 1959, Nichols 1986). Since a defining property of heads is that they may select nonheads within their project, this approach views dependency as the function of a semantic selection relation.

I believe this definition is both too narrow and too much dependent on particular views of phrase structure. For example, the question of whether the subject is a dependent of the verb is answered differently depending on the hypothesized base position of the subject (inside the VP or not) and on the hypothesized structure of the functional domain (involving separate heads for functional categories like Tense or not).

A more basic approach to dependency would start from the core case in (7):

(7) δ is a dependent of α if δ expresses in its morphology a feature of α

On this view, a target element δ whose morphology is controlled by some other element α is by definition the dependent of α (cf. Corbett 2003). More generally, we can say that in this situation δ is set off from α by some marked feature, in this case morphology. But other features may also be used to set off one element from another, most notably pitch, linear order, and semantic interpretation. For instance, in the colloquial American English pejorative expression of the type illustrated in (8)(cf. Grohmann and Nevins 2004), *schmoney* is set off from *money* in at least four ways:

(8) Money schmoney

Morphologically, *schmoney* is dependent on *money* in that it reduplicates its root. Intonationally, *schmoney* is set off by a marked (higher) pitch on its first syllable. In linear order, the reduplicating (target) element invariably follows the reduplication source (the controller). And semantically, *schmoney* predicates a particular property of *money*, mainly the property of irrelevance.

Generalizing, we state the following definition of dependency:

(9) δ is a dependent of α if δ is set off from α by some marked property

Since on this view of dependency, dependency indicators are distributed over the components of the grammar, some belonging to the A-P side of the grammar (morphology, intonation, linear order) and others belonging to the C-I side (semantic interpretation), we would like to know whether the various indicators express the dependency in a consistent way. If so, the source of the dependency must not be associated with the interface components A-P or C-I, but with the component feeding both interface components, CHL.

3.2 Indicators of dependency

It is clear that both within a single language and across languages, linear order is the least reliable indicator of dependency. Thus, whereas in the Dutch example (10a) all indicators of dependency converge, as detailed in (11b), (11) from the same language and (11b), has inverted the linear order, but arguably not the dependency.

(10) a. Vogel-s VLIEG-en (Dutch)
bird-PL fly-PL
'Birds fly.'

b. morphology: [PLURAL] is an inherent feature of the noun phrase *vogels*, expressed on the verb *vliegen* (i.e *vogels* = α, *vliegen* = δ)
linear order: δ follows α
intonation: δ has high pitch, α has neutral pitch
semantics: δ is interpreted as a predicate of α (and not as another noun phrase; NB *vliegen* could also be interpreted as the plural noun *flies*, but not here)

(11) Toch VLIEG-en vogel-s (Dutch)
yet fly-PL bird-PL
'Yet birds fly.'

In (11), morphological, intonational and semantic dependency indicators still identify *vliegen* 'fly' as the dependent of *vogels* 'birds', in spite of the inverted linear order.

Examples illustrating the unreliable character of linear order as a dependency indicator can easily be multiplied. (12), for instance, shows examples of N-V compounds in French and Dutch:

(12) a. ouvre-bouTEILle (French)
open-bottle

b. FLES-open-er (Dutch)
bottle-open-NOM

both: 'bottle opener'

In both cases, the element translated by 'bottle' is the dependent of the element translated by 'open', as 'bottle' is interpreted as the element affected by 'open'. This dependency is expressed by pitch accent in both French and Dutch, but by linear order only in French. We find similar facts within a single language, in the expression of geographical terms:

(13) a. de berg HOreb (Dutch)
the mountain Horeb
'mount Horeb'

b. het ATlas-gebergte
the Atlas-range
'the Atlas range'

In (13a) we find the geographical term realized in a construct state syntax, with the name specifying the geographical object following it and set off via pitch accent. In (13b), a similar geographical object is described in a compound, with the specifying element preceding the element specified, but still carrying the pitch accent indicative of its dependent character.

The circumstance that linear order is an unreliable indicator of dependency is reflected in the more or less even distribution of head-initial and head-final languages in the sample (Table 1). Taking a complement to invariably be a dependent of the head it is merged with, the very existence of the head-initial/head-final typology indicates that linear order does not reflect dependency directly. In terms of the structure building operation Merge, linear order alterations may be realized by remerging one of the elements merged to the derivation in a previous step (movement; cf. (5)-(6)).

To abstract away from the effects of movement, we may study indicators of dependency in a range of constructions where we are certain that no remerge takes place. Some Dutch examples of these 'restricted' expressions are listed in (14).

(14)	a.	sports result	1-0	één-NUL	[one-nil]	(Dutch)
	b.	digit sequence	1, 2, 3	één-twee-DRIE	[one-two-three]	
	c.	numbers	21	een-en-TWINtig	[one and twenty]	
	d.	the time	1:30	half TWEE	[half two]	
	e.	the amount	2,50	twee-VIJFtig	[two-fifty]	
	f.	reduplication	zozo	zo-ZO	'so-so'	
	g.	titles		luitenant-koloNEL	'wing commander'	
	h.	acronyms	PvdA	payvaydayAH	[socialist party]	

These examples involve mere juxtapositions, and the only dependency indicators which

appear to be consistently present are linear order and intonation. In each case, the second element in the linear order (rightmost on paper, following in time) is set off by higher pitch.

Morphological dependency indicators appear to be generally absent in this type (except, possibly, in reduplications of the kind in (8)). Semantic dependency may be detected, however. Consider the example of sports results (14a). The leftmost term, *één* 'one' seems to have a simple cardinality reading (n_1=1), but the rightmost terms appears to be ambiguous between the cardinality reading (n_2=0) and a reading that profiles the number against that of the leftmost term (n_2<n_1). This reading is only possible if the rightmost term is intonationally set off from the leftmost term. When both terms are pronounced equal, the relative reading of the rightmost term disappears. This indicates that the interpretation of the second of the two terms is dependent on that of the first. Other types in which some semantic dependency may be identified are (14d,e), where the second term is interpreted as a fraction of the first, and (14g), where the second term specifies a subrank of the first.

If restricted expressions of the type in (14) are created by Merge, it is interesting to hypothesize that the convergence of the dependency indicators that are available in this type reflect the basic asymmetric organization of pairs created by Merge. That this conclusion might be justified is suggested by the observation that slightly less restricted juxtaposition types appear to maintain the convergence of dependency indicators, most notably the correlation of intonation and linear order. This can be seen in the cases listed in (15)(Zwart 2003).

(15)	a.	coordinations	john and MAry		
	b.	asyndetics	me TARzan		
	c.	construct state	huis USHer	[house (of) Usher]	(Dutch)

In these types, the semantic contribution of the second member is often that of predication of specification. In addition to the basic dependency indicators of linear order and intonation, we see the beginnings of morphological dependency indicators, via linking particles (the conjunction in (15a), the functional preposition *of* in the English version of the construct state construction (15c)).

From here it appears to be a small step to full fledged predication and complementation structures, where the predicate/complement are intonationally set off

from their sisters by the Nuclear Stress Rule, which puts pitch accent on the complement (Chomsky & Halle 1968, Cinque 1993), or, in the spirit of Zubizarreta's (1998:43) reformulation, on the dependent element in an asymmetric sister pair. In addition, we see the emergence of a range of morphological dependency indicators with predication and complementation, the exact distribution of which may very well be studied in large scale language samples. The only complication with predication and complementation is that here (especially with complementation), the syntactic derivation may involve movement, eliminating linear order as a reliable dependency indicator.

The relevance of this discussion of dependency indicators to the general question at hand is the following. If we take the dependency indicators linear order, morphology, and intonation to properly belong to the A-P interface domain, and the semantic dependency indicator to the C-I interface domain, we observe that:

(16) a. dependency indicators tend to converge across interface domains,
b. divergence of dependency indicators takes place within the A-P interface domain, and
c. the cause of divergence is a syntactic operation belonging to neither of the interface domains

These observations strongly suggest that the source of the asymmetries underlying the dependency indicators is not contained within one of the interface domains, but is instead centrally located in the component feeding both interface domains, CHL (narrow syntax).

4. Coordination

We now turn to a description of sample based research on coordination addressing the question of whether coordination constructions are generated in a symmetric or asymmetric fashion.

4.1 Preliminaries

Coordination constructions may be described as symmetric or asymmetric based on the

type of coherence relation between the members of the coordination (e.g. Kraak and Klooster 1968:252f., Kehler 2000: 539f.). In (17), the coherence is of a balanced type (symmetric), whereas in (18), the coordinated members are organized in a contiguous (18a) or even resultative (18b) association (asymmetric):

(17) John likes Mary and Bill hates Susan

(18) a. John came home and grabbed a sandwich

b. Federer won the US Open and became the year-end number one for the third year in succession

In the study reported on here, I looked only at coordinations that could be characterized as 'symmetric' in this sense. More particularly, while the typological survey covered coordinations of all categories, conclusions were drawn on the basis of noun phrase coordinations only.

Based on the number of conjunctions *N* we can distinguish asyndetic (*N* = 0), monosyndetic (*N* = 1) and polysyndetic (*N* = >1) coordination types. We concentrate on the monosyndetic type here, which may in principle be expected to appear in one of three positions (where *&* is the conjunction and *A,B* are the coordinands):

(19) a. & A B

b. A & B

c. A B &

Since type (19a) does not occur at all (cf. Haspelmath 2000), I use the terms 'initial conjunction' and 'final conjunction' for the types in (19b) and (19c), respectively.

It should be noted that certain textbook examples of final conjunctions on closer inspection turn out to be initial conjunctions, suffixed to the second coordinand's first word. This is the case in Latin, where the bound morpheme *-que* only appears to be an example of a final conjunction:

(20) arma virum-que canō (Latin)
weapon:PL.ACC man:SG.ACC-and sing:1SG.PRES
'I sing of war and of a man.'

However, where the second coordinand contains more than one word, we see that *-que* must appear as a second position element within the second coordinand:

(21) ingenia fecunda totius-que naturae
mind:PL-ACC fertile:PL.ACC all:SG.GEN-and nature:SG.GEN

capacia (Latin)
grasping:PL.ACC

'minds that are fertile and able to grasp the entire universe.'

This type of 'second position initial conjunction' is found in at least 13 languages in the sample (Bella Coola, Evenki, Fon, Hausa, Jacaltec, Kalasha-ala, Lezgian, Shipibo, Turkish, Wardaman, West Greenlandic, Zaghawa, Zay), mostly with clausal coordination, but sometimes with noun phrase coordination (where the type is potentially confusing) as well.

Coordinations may be formally marked by various types of elements (ignoring intonation, which Mithun 1988:331f. identifies as a major coordination marking device). Next to *true conjunctions* of the type of English *and*, which are not currently used in any other function than coordination, we find *comitative* coordination markers (cf. Stassen 2000), illustrated in (22), and *summary* coordination markers (cf. Haspelmath 2000, section 6.3), illustrated in (23).

(22) *comitative strategy*
Péédoró-mútsi-kye Jóáá-**ma** ájtyúmɨ́ɨ́be (Bora)
Pedro-DU-ACC Juan-with see:1SG
'I see Pedro and Juan.' (Thiesen 1996:75)

(23) *summary strategy*

Péédoro-o, Jóáa-á, Perípe-é, **éhdume** péé téhullévu (Bora)
Pedro-RED Juan-RED Felipe-RED this quantity go away
'Pedro, Juan, and Felipe went away.' (Thiesen 1996:75)

The comitative coordination marking strategy uses a marker otherwise used to express comitativity or instrumentality ('with'), mostly an adposition or a particular case marker. As we will see, the comitative marker may develop into a true conjunction (cf. Mithun 1988: 339).

In the summary coordination marking strategy, coordinands are listed and then resumed by some element identifying the listed elements as a single participant in the event. This summarizing function may be performed by a range of elements, including pronouns, quantifiers, number markers (plural or dual), copulas, focus markers (such as *also, too*), adverbs of the *together* type, etc. As discussed in Mithun (1988: 337), this coordination strategy may be regarded as a minimal elaboration of a more basic, asyndetic type, where the coordinands are merely juxtaposed.

4.2 Word order generalizations

What follows replicates the findings of Zwart (2005), now based on an enlarged sample of 214 languages (see the Appendix). As the bottom line we may state that true conjunctions are invariably of the initial type (19b).

First of all, if we define a language as head-initial (or final) based on the position of a verb and/or adposition with respect to its complement (where a language is head initial if a head precedes its complement), the languages in the sample are about evenly partitioned, as illustrated in Table 1.

HEAD-INITIAL	HEAD-FINAL	SPLIT	UNCLEAR
96	**91**	**10**	**17**

Table 1: head position in a 214 language sample

In contrast, initial conjunctions (of any type) are found much more frequently than final

conjunctions, as Table 2 illustrates.

INITIAL	FINAL	MIXED	OTHER
135	**12**	**26**	**39**

Table 2: conjunction position in a 214 language sample

In the mixed group, 11 languages have both initial and final conjunctions and 15 languages show a preference for polysyndetic coordination; of these, 5 languages have final conjunctions as a monosyndetic alternative. In the 'other' group, 16 use polysyndetic coordination, and data from the remaining 23 are either missing or unclear.

It follows that a sizeable number of head-final languages uses initial conjunctions. Conversely, no strict head-initial languages use final conjunctions (although some languages with split headedness do use final conjunctions, and some head-initial languages use final summary elements as a secondary strategy). Table 3 shows the distribution of initial and final conjunctions over head-initial and head-final languages.

INITIAL CONJUNCTION			FINAL CONJUNCTION		
135			**12**		
head-initial	split	head-final	head-initial	split	head-final
85	**3**	**47**	**0**	**2**	**10**

Table 3: conjunction position in head-initial vs. head-final languages

The following table shows that the 12 languages using final conjunctions exclusively invariably employ either the comitative or the summary strategy:

LANGUAGE	SUMMARY	COMITATIVE	TRUE
35:3 Slave	X	X	
56:2 Yaqui		X	
58:1 Ika		X	
64:1 N Junin Quechua		X	
65:1 Jaqaru		X	
71:1 Yagua	X		
72:1 Bora	X	X	
75:1 Sanumá	X		
76:1 Barasano	X		
76:2 Retuarã	X		
82:1 Paumarí	X		
86:1 Trío	X		

Table 4: languages using final conjunctions exclusively

To illustrate, the Ika (Aruak, Chibchan) coordination marker *-sin* in (24a) has a primary use as an instrumental or comitative element in (24b) (Frank 1990:37-38).

(24) a. ribru rapi-sin pa ú (Ika)
book pencil-with put down AUX
'Put down the book and pencil!' (Frank 1990:38)

b. kʌnsia-sin si aʔsir-i ...
vine-with string tie-while
'He tied it with a vine ...' (Frank 1990:37)

Likewise, the Barasano (East Tucanoan) noun phrase conjunction *kẽde* (25a) has an independent use as a focus marker 'also' (25b) (Jones & Jones 1991:30).

(25) a. ũbʊ-a rõbi-a dake-rã kẽde yã-ka-bã ĩdã
male-PL female-PL young-AN.PL also be-far:PAST-3PL 3PL
(Barasano)
'There were men, women and children there.' (Jones & Jones 1991:133)

b. to bahi-ro yi-ya bũ kẽde
that be-NOM do-PRES 2SG also
'You will do that too.' (Jones & Jones 1991:174)

Also to consider are the 16 languages in the mixed group of table 2 (11 using both initial and final conjunctions, 5 using final conjunctions next to a polysyndetic type; two of these, Baram Kayan and Wari', appear to be head-initial). Here, too, the true conjunction is exceedingly rare.

LANGUAGE	SUMMARY	COMITATIVE	TRUE
3:1 Logbara	X	X	(X?)
7:6 Kalasha-ala	X		X*
10:1 Kolyma Yukaghir		X	
12:1 Ket		X	
15:2 Kham	X		
19:21 Baram Kayan	X		
33:4 W Desert Lg	X	X	
33:5 Kayardild	X		
35:2 Navaho		X	
46:1 Hualapai	X	(X)	
56:1 Shoshone		X	
64:2 Imbabura Quichua	X		
73:1 Pirahã	X		
83:1 Tariana		X	
85:1 Wari'	X		
94:1 Kwaza			X*
* also used as initial conjunction			

Table 6: languages using final conjunctions not exclusively

If we look closer at the relevant cases of true conjunctions in final position, it turns out that none of them is above suspicion.

In Logbara (Central Sudanic, Nilo-Saharan), the final conjunction is the element *pie*, illustrated in (26).

(26) à mu ɛ̀ri pie àkú-a (Logbara)
we go he and home-to
'I and he go home.' (Crazzolara 1960:100)

In this example, the first coordinand is a so-called inclusory pronominal, semantically singular ('I') but formally marked plural (Haspelmath 2000: section 6.2, Lichtenberk 2000). The pronoun is called 'inclusory' because, as Ladusaw (1989) argues, the pronoun encodes a set which includes the referent of the second coordinand. The first coordinand, then, functions as the head of the construction, and the second as an adjunct. Not surprisingly, the inclusory strategy commonly employs a comitative marker to link up the two coordinands. A true conjunction, on the other hand, appears to be rare (cf. Haspelmath 2000).

The element *pie* is not a comitative marker, however, but it may be related to the element *pi*, a verb meaning 'to complete', and the element *pi* (with mid-low tone) which, when following a noun phrase, means '*cum suis*' (Crazzolara 1960: 101). Example (27) illustrates the use of *pi*, as well as the comitative marker used in Logbara coordination:

(27) ɛ̀tɔ́ɔ p`ì mù dɱ̀ a'ú-ā̃ bɛ (Logbara)
hare c.s. go then fowl-DIM with
'the hare and a small fowl went together.' (Crazzolara 1960:101)

If *pi* is a '*cum suis*' pluralizer, (27) represents another example of the inclusory strategy, not with pronouns but with full noun phrases (similar to the case of Margi reported by Haspelmath 2000). In all, there is reason to believe that *pie/pi* is not a true conjunction but a pluralizer of some kind, which would imply that (26) is not true coordination but coordination of the summary type.

In Kalasha-ala (Nuristani, Indo-Aryan), the final conjunction *ye* is apparently found with fixed expressions only. Whereas (28a) simply means 'a man and a woman', (28b),

which may be further reduced, refers to a man-woman pair, such as ‘husband and wife’:

(28) a. e meši ye e muša (Kalasha-ala)
a woman and a man
‘a woman and a man’ (Degener 1998:166)

b. meši moša ye (> meši-moša-y)
women men and
‘man and woman’ (ibid.)

As far as I have been able to ascertain, the final conjunction only occurs in fixed expressions of this type, suggesting that we are dealing with a special case.

Finally, the final conjunction in Kwaza (unclassified, Brazil) is what Van der Voort (2004) terms a ‘cosubordinator’ (CSO):

(29) si xyi-a-’ta oja-’nã-a-ki Ba’hozo-nã (Kwaza)
I you-1PL-CSO go-FUT-1PL-DEC Barroso-LOC
‘I and you we are going to Barroso.’
[lit: me, you, we being, we go to Barroso] (Van der Voort 2004:707)

The cosubordinating suffix *’ta/tja* is analyzed as a medial clause same subject marker, i.e. as belonging to the switch reference marking system. Switch reference markers are used in coordination constructions in a number of languages in the sample, including Kiowa, Nateni, Shipibo, and Western Desert Language. Like in Kiowa (Watkins 1980:288), the switch reference marker seems to have developed into a true (initial) conjunction, in examples like (20) (the declarative marker is a verbalizer, cf. Van der Voort 2004: 208).

(30) ’meza-na kore’ja-tja di’hu-tse (Kwaza)
table-LOC knife-CSO spoon-DEC
‘On the table are a knife and a spoon.’ (Van der Voort 2004:706)

But in other constructions, such as (29), the switch reference marker seems to have retained the original function of linking a (clausal) constituent up with a matrix clause.

In this respect the appearance of the 1PL suffix *-a* seems relevant, leading to the literal translation indicated, where the coordination acts like an adjunct clause. This case, too, then does not appear to be a case of a true conjunction in final position.

If the evaluation of the final conjunctions in Logbara, Kalasha-ala and Kwaza is correct, none of the 214 languages in the sample features true conjunctions in final position. Moreover, a substantial number of languages have developed initial true conjunctions out of final comitative markers (17 languages in the sample: Burmese, Dogon, Haida, Japanese, Ket, Kinnauri, Kokborok,Korean, Ladakhi, Lavukaleve, Mikir, Navaho, Northern Qiang, Shipibo, Slave, Suppire, Turkish), or out of other final elements such as the switch reference markers of Kiowa and Kwaza (cf. (29)-(30)). This is illustrated in (31) for Kinnauri (Western Himalayish, Sino-Tibetan).

(31) a. əṅ rəṅ doː chaṅ due (Kinnauri)
1SG:GEN with 3SG:GEN son be:3PAST
'His son was with me.' (Sharma 1988:91)

b. gə rəṅ ki bi-tič
1SG:DIR and you:HON go-FUT:1DU.INCL.HON
'I and you will go.' (Sharma 1988:182)

Finally, whereas conjunction borrowing is a common phenomenon (for reasons discussed in Mithun 1988), no language seems to have come up with a final conjunction as the result of borrowing (based on 13 languages in the sample with borrowed conjunctions).

Mithun (1988), discussing the origin of formal coordination markers, notes that formally marked noun phrase conjunction is disfavored in spontaneous discourse, and suggests that the emergence of conjunctions may be a function of the development of literacy. She describes the summary and comitative strategies as minimal elaborations of the more basic type of asyndetic coordination, in fact as types "originally used to circumvent coordinate noun phrases" (Mithun 1988:339). Indeed we regularly find the comitative or summary strategy described as a (marked) alternative to asyndetic juxtaposition (e.g. Nedjalkov 1990:90 on Evenki, Koroma 1994:76 on Gola, Enrico 2003:1079 on Haida, Evans 1995:394 on Kayardild, Davies 1981:72 on Kobon, Smeets 1989:180 on Mapadungu, De Smedt & Mostaert 1964:163 on Monguor, Dayley

1989:336 on Shoshone, Aikhenvald 2004:484 on Tariana, Lewis 1986:206 on Turkish, Glass & Hackett 1970:64-65 on Western Desert Language, Payne 1985:97 on Yagua, Dedrick & Casad 1999:360 on Yaqui, Maslova 2003:313f. on Kolyma Yukaghir, among others).

This suggests that a distinction be made between juxtaposition, comitative coordination and summary coordination on the one hand, and true coordination on the other. The generalization that arises from the languages in the sample, then, is that true coordination is invariably marked on the second conjunct, more precisely, at the second conjunct's left edge.

4.3 Discussion

The coordination phenomena reviewed in section 3.2 show up a clear asymmetry between the two members of a coordinated structure: a coordination marked by a true conjunction (i.e. not a comitative or summary element) invariably singles out the second conjunct for expressing the coordination.

As discussed by Munn (1993), the English conjunction *and* forms a constituent with the second member, witness its position in discontinuous coordinations of the type in (32b,c):

(32) a. I saw [John and Bill] the other day
b. I saw [John] the other day [and Bill]
c. * I saw [John and] the other day [Bill]

I am not aware of any languages differing from English in this respect. This allows us to proceed on the assumption that the conjunction properly belongs to the second coordinand.

Munn's conclusion, adopted by Kayne (1994), is that the conjunction actually functions as a head, taking the second conjunct as its complement. If this is correct, the facts surveyed indicate that true coordination is invariably head-initial, suggesting that even clear head-final languages are not entirely without head-initial syntactic organization (Zwart 2005).

Alternatively, we may describe the conjunction in more neutral terms as a linking device. Since the linker marks the second conjunct, this leads to the conclusion that in

noun phrase coordination the second conjunct is invariably the dependent of the first conjunct. The findings then suggest an order-dependency correlation, such that the nondependent invariably preceeds the dependent.

If this is correct, and if coordination is established via the structure building operation Merge, we are allowed to conclude that Merge invariably creates a dependency between the elements merged, hence that Merge operates in an asymmetric fashion.

5. Conclusion

I have argued in this article that sample based typological research may be instrumental in addressing central issues in minimalist syntactic theory. Concretely, I have shown that there is reason to believe that the structure building operation Merge assumed within the Minimalist Program (Chomsky 1995) works in an asymmetric fashion, generating ordered pairs rather than unordered sets. I have suggested that a fruitful area of investigation relevant to this claim is provided by restricted syntactic constructions, including juxtapositions of various types, and slightly more complicated structures, such as coordinations marked by a (true) conjunction, asyndetic predications and construct state constructions. The research might focus on the realization of various indicators of dependency in these construction types, both belonging to the A-P (phonological) and C-I (semantic) side of the grammar. In particular, the question whether these indicators of dependency are realized in a consistent way across languages and constructions may be fruitfully studied using the empirical data provided by standard reference grammars. As an example, I reviewed a sample based study of noun phrase coordination, showing that true coordination invariably involves morphological marking of the left edge of the second coordinand.

Appendix.

The sample used in the research discussed here is a variety sample of 214 languages, created for the study of morphosyntactic variation across the languages of the world. Its aim is to arrive at maximal coverage of linguistic variation, not to estimate the probability of certain features (co)occurring. In the terminology of Bell (1978), we

determine as the universe of the sample: the total of all possible human languages, or universal grammar; and as the frame of the sample: the total of actual extant human languages. The frame was partitioned in 71 language families, following the classification in 96 families of Gordon 2005 (the Ethnologue) (25 very small families could not be included; the number of families in the Ethnologue has since increased to 108, and the sample may be adjusted accordingly). The number of languages selected from each family was determined by size and estimated diversity, and the selection of individual languages was determined first by subgrouping and then by convenience (i.e. the availability of excellent reference grammars).

The following families and languages were included in the sample:

(i) 1. Khoisan, 2 languages (!Kung, tbd); 2. Niger-Congo, 27 lgs (Pulaar, Gola, Kalabari, Kikuyu, Nkore-Kiga, Ewondo, Tikar, Tiv, Limbum, Ibibio, Yoruba, Degema, Igbo, Duka, Jukun, Birom, Dogon, Grebo, Fon, Baule, Mundang, Ngbaka, Nateni, Suppire, tbd, Bobo Dioula, Soninke); 3. Nilo-Saharan, 5 lgs (Logbara, Dilling, Lango, Zaghawa, Songhai); 4. Afro-Asiatic, 8 lgs (Tamasheq, Margi, Lele, Hausa, Iraqw, Kafa, Gulf Arabic, Zay); 5. South Caucasian, 1 lg (Georgian); 6. North Caucasian, 2 lgs (Lezgian, Kabardian); 7. Indo-European, 9 lgs (Albanian, Lithuanian, Breton, Dutch, Gojri, Kalasha-ala, Dimli, Portuguese, Russian); 8. Uralic, 2 lgs. (Nenets, Hungarian); 9. Dravidian, 4 lgs. (Kolami, Brahui, Abujhmaria, Tamil); 10. Yukaghir, 1 lg. (Kolyma Yukaghir); 11. Altaic, 3 lgs. (Monguor, Evenki, Turkish); 12. Yenisei Ostyak, 1 lg. (Ket); 13. Japanese, 1 lg. (Japanese); 14. Chukotko-Kamchatkan, 2 lgs. (Chukchi, Itelmen); 15. Sino-Tibetan, 12 lgs. (Cantonese, Ladakhi, Kinnauri, Kham, Kokborok, Eastern Kayah Li, Mao Naga, Burmese, Meithei, Karbi, Digaru, Northern Qiang); 16. Miao-Yao, 1 lg. (tbd); 17. Austroasiatic, 6 lgs. (Temiar, Chrau, Khmer, Khasi, Vietnamese); 18. Daic, 2 lgs. (Kam, Thai); 19. Austronesian, 31 lgs. (Sedeq, Paiwan, Tsou, Kambera, Paulohi, Roti, Loniu, Fijian, Samoan, Iaai, Ponapean, Tiri, North Efate, Sie, Kwaio, Hoava, Manam, Kilivila, East Makian, Malagasy, Baram Kayan, Ida'an, Chamorro, Tagalog, Ivatan, Sama, West Bukidnon Manobo, Mongondow, Muna, Acehnese, Toba Batak); 21. Andamanese, 1 lg. (A-Pucikwar); 24. East Papuan, 1 lg. (Lavukaleve); 28. Sepik-Ramu, 3 lgs. (Yimas, Abelam, tbd); 30. Torricelli, 1 lg. (Arapesh); 31. Trans New Guinea, 9 lgs. (Tauya, Amele, Central Asmat, Kobon, Nabak, Marind, Daga, Eipo, Nimboran); 32. West Papuan, 1 lg. (Abun); 33. Australian, 6 lgs. (Gooniyandi, Wardaman, Dyirbal,

Western Desert Language, Kayardild, Djingili); 34. Eskimo-Aleut, 2 lgs. (Aleut, West Greenlandic); 35. Na-Dene, 4 lgs. (Haida, Navaho, Slave, Tlingit); 36. Algic, 1 lg. (Eastern Ojibwe); 38. Wakashan, 1 lg. (Nootka); 39. Salish, 2 lgs (Bella Coola, Thompson); 41. Siouan, 1 lg. (Lakota); 42. Caddoan, 1 lg. (Wichita); 43. Iroquoian, 1 lg. (Tuscarora); 44. Penutian, 2 lgs. (Yokuts, Coos); 46. Hokan, 2 lgs. (Hualapai, Karok); 48. Kiowa-Tanoan, 1 lg. (Kiowa); 50. Muskogean, 1 lg. (Koasati); 52. Totonacan, 1 lg. (Misantla Totonac); 53. Mixe-Zoque, 1 lg. (Chimalapa Zoque); 54. Mayan, 3 lgs. (Tzotzil, Jacaltec, Cakchiquel); 55. Subtiaba-Tlapanec, 1 lg. (Tlapaneco); 56. Uto-Aztecan, 2 lgs. (Shoshone, Yaqui); 57. Oto-Manguean, 5 lgs. (Sochiapan Chinantec, Chalcatongo Mixtec, Mezquital Otomi, Southern Popoloca, Zoogocho Zapotec); 58. Chibchan, 1 lg. (Ika); 61. Choco, 1 lg. (Saija); 63. Zaparoan, 1 lg. (Zaparo); 64. Quechuan, 2 lgs. (North Junín Quechua, Imbabura Quichua); 65. Aymaran, 1 lg. (Jaqaru); 66. Aracaunian, 1 lg. (Mapudungu); 71. Peba-Yaguan, 1 lg. (Yagua); 72. Witotoan, 1 lg. (Bora); 73. Mura, 1 lg. (Pirahã); 75. Yanomam, 1 lg. (Sanumá); 76. Tucanoan, 2 lgs. (Barasano, Retuarã); 78. Nambiquaran, 1 lg. (Sabanês); 79. Maku, 1 lg. (Dâw); 81. Tupi, 3 lgs. (Kanoé, Guaraní, Yuqui); 82. Arauan, 1 lg. (Paumarí); 83. Arawakan, 1 lg. (Tariana); 85. Chapacura-Wanham, 1 lg. (Wari'); 86. Carib, 2 lgs. (Trió, Hixkaryána); 89. Mataco-Guaicuru, 1 lg. (Pilaga); 90. Panoan, 1 lg. (Shipibo); 91. Tacanan, 1 lg. (Araona); 92. Macro-Ge, 2 lgs. (Borôro, Canela); 93. Isolates, 6 lgs. (Ainu, Basque, Burushaski, Korean, Mosetén, Nivkh); 94. Unclassified, lg. (Kwaza); 95. Creoles, 3 lgs. (Ndyuka, Mauricien, Cape Verdean Creole); 96. Pidgins, 1 lg. (Naga Pidgin); 97. Deaf Sign Languages, 1 lg. (American Sign Language).

If we set the total number of languages in the world, excluding isolated/unaffiliated languages, pidgins and creole languages and signed languages, at 6871 (Gordon 2005), the sample as a whole represents 3.1 % of that total number. The 25 groups not represented in the sample total 112 languages, i.e. 1.6 %. The only sizeable group among the nonrepresented families is the Geelvink Bay group at 33 languages, a situation which needs to be remedied. The reason these groups are not represented is the current unavailability of good reference grammars.

Underrepresentation occurs where a language family's representation in the sample, calculated as the percentage of languages of that family included in the sample, is under 3.1 %. This typically occurs with large language families, notably Niger-Congo

(1514 lgs, repr 1.8 %), Indo-European (445 lgs, repr 2.0 %), Afro-Asiatic (375 lgs, repr 2.1 %), Australian (263 lgs, 2.3 %), and Austronesian (1268 lgs, repr 2.4 %). In other cases, underrepresentation is the result of families or subgroups being poorly described, as with Arawkan (1.6 %), Trans-New Guinea (1.6%), and Torricelli (1.9%). Underrepresentation may also occur with very small language groups with estimated low diversity (i.e. few or no subdivisions). Overrepresentation is not necessarily problematic in a variety sample, where no correlations are sought. In the sample, small language groups are typically overrepresented, with a view to covering the family internal diversity.

In what follows I briefly discuss the issue how the quality of a variety sample of this type might be assessed, considering questions of diversity and coverage.

Diversity is to some extent a function of size: a small family of four languages, divided over two branches, shows more diversity than an arbitrary set of four languages belonging to a larger subgroup. In the first case we want representation of two out of four languages, in the second we want the subgroup to be represented, for which a single language may suffice. Quantification of diversity is done by Bell (1978) and Rijkhoff et al. (1993). Bell's diversity assessment is based on the projected number of groups per phylum in separation at 3500 years before current time. We do not adopt this approach, for two reasons. First, for a number of phyla, the historical record does not allow for an informed estimate. Second, historical accident may cause deviations from the presupposed steady rate diversity increase. Rijkhoff et al.'s *diversity value* is a function of the complexity of genetic language trees (with weighting of higher versus lower branches), arrived at by calculating the average number of nonroot/nonterminal nodes at each intermediate level in the tree. This approach is more promising, we feel, but an additional factor that might be taken into account is the size of the terminal groups. We proceed from the idea that diversity arises when an opposition between (groups of) languages necessitates branching. It follows that two factors contribute significantly to diversity: a) the size of the terminal groups (where small = diverse), and b) the number of splits in the genetic language family tree. Note that the approach to estimating diversity assumes that branching reflects diversity in a consistent way, an assumption which is certainly false. This is an inherent weakness of any approach to diversity based on configurational properties of linguistic affiliation trees (see also Croft 1990:22).

More important to a variety sample than genetic (or regional) bias is the effective

coverage. This can be seen as a function of the proportion of the branches within each family represented in the sample. Sampling technique should be directed at gaining maximal coverage out of a limited number of languages. We propose to measure coverage by calculating for each phylum the average percentage of the branches of each layer of the genetic tree represented in the sample. If we number the layers of the tree from the bottom up from 1 to *n*, additional weighting can be done (in the spirit of Rijkhoff et al. 1993) by multiplying the percentage for each layer by its number, and dividing the sum of the percentages by *n*!. In the current sample, coverage values range from .34 (Austronesian) to 1.00 (Yukaghir). This method of estimating coverage provides an easy tool for improving the sample, as a low coverage value is typically caused by nonrepresentation of one or more high level branches. For instance, the relatively low coverage value of Nilo-Saharan (.44) is explained by the circumstance that the Nilo-Saharan family branches directly into a high number of subgroups, most of which are small and difficult to represent for lack of available descriptions. Incorporation of one or more languages from these branches would effect an immediate improvement. As it stands, certain coverage values are unacceptably low, but the expectation is that with the adition of another 30-40 languages, the situation might be significantly improved.

References.

Aikhenvald, A. 2004. *A grammar of Tariana*. Cambridge: Cambridge University Press.

Bell, A. 1978. Language sampling. In *Universals of human language, Vol. 1*, J.H. Greenberg, ed., 125-156. Stanford: Stanford University Press.

Chomsky, N. 1995. *The minimalist program.* Cambridge, Mass.: MIT Press.

Chomsky, N. 2001. Derivation by phase. In *Ken Hale: a life in language*, M. Kenstowicz, ed., 1-52. Cambridge, Mass.: MIT Press.

Chomsky, N. 2005a. Three factors in language design. *Linguistic Inquiry* 36, 1-22.

Chomsky, N. 2005b. On phases. Ms., MIT.

Chomsky, N. and M. Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Cinque, G. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24, 239-297.

Corbett, G. 2003. Agreement: terms and boundaries. In *The role of agreement in natural language: TLS 5 proceedings*, W.E. Griffin, ed., 109-121. Austin: University of Texas.
http://uts.cc.utexas.edu/~tls/

Croft, W. 1990. *Typology and universals*. Cambridge: Cambridge University Press.

Gordon, R.G., Jr. 2005. *Ethnologue: languages of the world (15th edition)*. Dallas: SIL International.
http://www.ethnologue.com/

Crazzolara, J.P. 1960. *A study of the Logbara (Ma'di) language: grammar and vocabulary*. London: Oxford University Press.

Davies, J. 1981. *Kobon*. Amsterdam: North Holland Publishing Company.

Dayley, J.P. 1989. *Tümpisa (Panamint) Shoshone grammar*. Berkeley: University of California Press.

De Smedt, A. and A. Mostaert 1964. *Le dialecte monguor parlé par les Mongols du Kansou occidental, 2ème partie: grammaire*. The Hague: Mouton.

Dedrick, J.M. and E.H. Casad. 1999. *Sonora Yaqui language structures*. Tucson: University of Arizona Press.

Degener, A. 1998. *Die Sprache von Nisheygram im afghanischen Hindukusch*. Wiesbaden: Harrassowitz.

Enrico, J. 2003. *Haida syntax*. Lincoln: University of Nebraska Press.

Epstein, S.D. 1999. Un-principled syntax: the derivation of syntactic relations. In *Working minimalism*, S.D. Epstein and N. Hornstein, eds., 317-345.

Evans, N. 1995. *A grammar of Kayardild with historical-comparative notes on Tangkic*. Berlin: Mouton de Gruyter.

Fortuny Andreu, J. 2007. *The emergence of order in syntax*. University of Barcelona dissertation.

Frank, P. 1990. *Ika syntax*. Dallas: Summer Institute of Linguistics.

Glass, A. and A. Hackett. 1970. *Pitjantjatjara grammar: a tagmemic view of the Ngaayatjara (Warburton ranges) dialect*. Canberra: Australian Institute of Aboriginal Studies.

Grohmann, K.K. and A. Nevins. 2004. Echo-reduplication: when too-local movement requires PF-distinctness. *University of Maryland Working Papers in Linguistics* 13, 84-108.

Haspelmath, M. 2000. Coordination. To appear in *Language typology and linguistic description*

Hauser, M.D., N. Chomsky, and W.T. Fitch. 2002. The Faculty of Language: what is it, who has it, and how did it evolve? *Science* 298, 1569-1579.

Jaspers, D. 1998. Categories and recursion. *Interface* 12, 81-112.

Jones W. and P. 1991. *Barasano syntax*. Dallas: Summer Institute of Linguistics.

Kayne, R.S. 1984. *Connectedness and binary branching*. Dordrecht: Foris.

Kayne, R.S. 1992. Word order. Guest lecture at the GLOW Colloquium, Lisbon, April 14.

Kayne, R.S. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Kehler, A. 2000. Coherence and the resolution of ellipsis. *Linguistics and Philosophy* 23, 533-575.

Koroma, R. 1994. *Die Morphosyntax des Gola*. Cologne: Institut für Afrikanistik.

Kraak, A. and W. Klooster. 1968. *Syntaxis*. Culemborg: Stam-Robijns.

Ladusaw, W. 1989. Group reference and the plural pronoun construction. In *Papers on the plural pronoun construction and comitative coordination*, 1-7. Santa Cruz: Syntax Research Center, University of California at Santa Cruz.

Lewis, G.L. 1986. *Turkish grammar*. Oxford: Oxford University Press.

Lichtenberk, F. 2000. Inclusory pronominals. *Oceanic Linguistics* 39, 1-32.

Maslova, E. 2003. *A grammar of Kolyma Yukaghir*. Berlin: Mouton de Gruyter.

Mithun, M. 1988. The grammaticization of coordination. In *Clause combining in grammar and discourse*, J. Haiman and S.A. Thompson, eds., 331-359. Amsterdam: John Benjamins.

Munn, A. 1993. *Topics in the syntax and semantics of coordinate structures*. University of Maryland dissertation.

Nedjalkov, I. 1990. *Evenki*. London: Routledge.

Nichols, J. 1986. Head-marking and dependent-marking grammar. *Language* 62, 56-119.

Payne, D.L. 1985. *Aspects of the grammar of Yagua: a typological perspective*. UCLA dissertation.

Rijkhoff, J., D. Bakker, K. Hengeveld, and P. Kahrel. 1993. A method of language sampling. *Studies in language* 17, 169-203.

Sharma, D.D. 1988. *A descriptive grammar of Kinnauri*. Delhi: Mittal Publications.

Smeets, I. 1989. *A Mapuche grammar*. Leiden University dissertation.

Stassen, L. 2000. AND-languages and WITH-languages. *Linguistic Typology* 4, 1-55.

Tesnière, L. 1959. *Éléments de syntaxe structurale*. Paris: Klincksieck.

Thiesen, W. 1996. *Gramática del idioma bora*. Yarinacocha: Instituto Linguïstico de Verano.

Van der Voort, H. 2004. *A grammar of Kwaza*. Berlin: Mouton de Gruyter.

Watkins, L.J. 1980. *A grammar of Kiowa*. University of Kansas dissertation.

Zubizarreta, M.L. 1998. *Prosody, focus and word order*. Cambridge, Mass.: MIT Press.

Zwart, J.W. 2003. What's in a name? Syntactic and asyntactic accentuation in Dutch. In *Grammatik i focus/Grammar in focus: Festschrift for Christer Platzack 18 November 2003*, L-O. Delsing, C. Falk, G. Josefsson, and H.Á. Sigurðsson, eds., Vol. II, 395-401. Lund: Department of Linguistics.

Zwart, C.J.W. 2004. The format of dependency relations. Bloomington 2004 Syntax Fest lectures, Indiana University, June 22 - July 1.

Zwart, J.W. 2005. Some notes on coordination in head-final languages. In *Linguistics in the Netherlands 2005*, J. Doetjes and J. van de Weijer, eds., 232-241. Amsterdam: John Benjamins.