AND-languages and WITH-languages

LEON STASSEN

Abstract

Based upon a semantic/cognitive definition of noun phrase conjunction, encoding of this domain is examined in a sample of 260 languages. Languages can be shown to contrast typologically in that they may have two formally different strategies to encode noun phrase conjunction (viz., a Coordinate Strategy and a Comitative Strategy), or just one strategy. This leads to a differentiation between AND-languages and WITH-languages. It is demonstrated that WITH-languages exhibit a tendency to "drift" towards AND-status. Furthermore, the geographical distribution of AND-languages and WITH-languages is described. It turns out that this distribution shows remarkable correspondences to the areal distribution of at least two other binary parameters, viz. the Casedness Parameter and the Tensedness Parameter.

Keywords: agreement, areal linguistics, case, comitative, conjunction, coordination, grammaticalization, head/dependent-marking, noun phrase, tense, word order

1. Introduction

This paper reports on a typological research project which deals with conjunctional strategies in natural language. It focuses on a specific subdomain of this problem area, namely, those conjunctional constructions in which noun phrases are involved. In what follows I will attempt to formulate a number of crosslinguistically valid parameters which govern the encoding of these noun phrase conjunctions in the languages of the world. In doing so I will use a sample of 260 languages, which have been selected from all major linguistic groupings and areas. An alphabetical listing of this sample is presented in the Appendix.

From a methodological (though not necessarily chronological) point of view, the conduct of any typological research project can be split up in a number of successive stages. At the first, DESCRIPTIVE, stage, the researcher aims to
establish a crosslinguistic DATABASE, which constitutes the empirical foundation of the inquiry. Obviously, the construction of a crosslinguistic database presupposes an explicit definition of the DOMAIN of the investigation, that is, a definition of the sort of data which are thought to be relevant to the project. The definition of the domain is not without its problems; I will say more about this point in Section 2. For the moment, let us assume that this question has been settled in a satisfactory manner, so that, for each language in the sample, the relevant data have been gathered and included in the database. Then the researcher will analyze these data, trying to detect crosslinguistic similarities and differences in the structural encoding of the domain. As a result of this analysis s/he will classify constructions and languages into categories which are characterized by a specific, distinct, mode of structural encoding. The final result of this description should be the establishment of a TYPOLOGY, in which the limits of this crosslinguistic variation in encoding are defined.

The typology functions as the input for the second stage of the investigation. On this EXPLANATORY level, we ask for a reason behind the crosslinguistic encoding facts for the domain. A major strategy on this level of research consists in the identification of CORRELATIONAL PARAMETERS. That is, one seeks to match the distinctions in the original typology (and hence, the various types in this typology) with types stemming from the crosslinguistic analysis of other domains. If this matching is wholly or partially successful, one is in a position to formulate correlational statements of the following general type:

(1) If a language belongs to Type X in Typology A, it belongs to Type Y in Typology B.

Typological correlations of this kind are interesting for a variety of linguistic and extra-linguistic reasons. From the perspective of linguistic theory, the main interest of these correlations lies in the fact that they formulate a connection between two bodies of typological facts which were hitherto taken to be unrelated. Thus, the identification of a correlational parameter for some typological distinction enables us to reorganize the typological facts of natural languages in a non-trivial fashion, in that hitherto unrelated typological facts can now be dealt with in terms of a CLUSTER. It will be evident that the establishment of such clusters is one of the most urgent, as well as rewarding, tasks of modern linguistic typology; it is by discovering correlational parameters that linguistic typology is able to lay a factual foundation for theory formation in modern grammar. The scientific work which results in the detection of correlations between typologies has often been underrated, and written off as being "merely descriptive". Clearly, this is unjustified, if for no other reason than that such correlations may often serve as an empirical measure by which competing analyses of a given construction type can be evaluated.
Now, although the identification of a correlational parameter is certainly valuable in itself, it is also arguable that the formulation of a typological correlation is not enough: one will commonly expect more of a correlation than the mere statement of its existence. The reason for this is that, in general, correlations do not present themselves as theoretically significant at first sight. There are various possible reasons why two typological distinctions are correlated with one another, and not all of these reasons are equally interesting. For one thing, one must always be aware of the possibility that correlations, for all their empirical validity, can nonetheless be spurious, that is, a case of purely statistical coincidence. Likewise, it cannot be excluded beforehand that the correlation is indirect, in that it is brought about by virtue of the fact that both typologies at issue are correlated with a third, as yet unidentified, typology. Given these uncertainties, a proposed typological correlation will usually be required to have some degree of explanatory value. The question of what exactly constitutes an explanation is far from clear in typological linguistics, as it is in linguistics in general. However, in my estimation, most linguists would agree that "in order for a typology A to count as an explanation of a typology B, we will generally require that the parameter of A represent some 'deeper-lying', 'more elementary', or 'more fundamental' linguistic property than the parameter upon which typology B is based" (Stassen 1985: 8). If this, admittedly rather vague, requirement is met to some extent, there is some intuitively satisfying way in which we can say that typology B is the way it is because typology A is the way it is. That is, the typological variation in A can now be viewed as the cause, or the motivation, of the typological distinctions manifested in B.

The present paper addresses only the descriptive and correlational stages in the establishment of the typology of NP-conjunction. I think I have succeeded in identifying several correlational parameters for this typology. After that, however, various explanatory questions remain. First, it is doubtful whether the correlational parameters identified here are the only ones to cluster with NP-conjunction; that is, it is by no means certain that the cluster at issue has been defined exhaustively. Furthermore, the mutual relations between the parameters in the cluster are not at all evident, so that the internal structure of the cluster remains opaque for the time being. Finally, it is a mystery to me why the cluster should contain just these members. Thus, I have to concede that, at the present state of my knowledge, I have nothing of any consequence to offer in the way of a "causal explanation" for the typological facts of NP-conjunction.

2. Definition of the domain of investigation

At the very beginning of any typological investigation the researcher will be confronted with the problem of crosslinguistic identification. This problem can
be stated in the following question: how can we be sure that the data which we select from the languages in the sample form a coherent body of facts? In other words, we need a principled way to identify in each language the structural data which are relevant to the project at hand, so that we will not end up working with a database which contains incomparable items. The solution to this problem presupposes a language-independent definition of the domain of the inquiry, that is, a demarcation of the relevant body of facts, which can be applied to any language, regardless of its structural characteristics. It can be argued that such a language-independent definition of the domain cannot be stated in purely formal terms (Stassen 1985: 14–15; Croft 1990: 11–18). By doing so one would only identify those constructions in languages which have a specified formal manifestation, and one would therefore forego the empirical question of determining the formal encoding variation across languages. Given this, a widely accepted research strategy is to define the domain of a typological research project by way of semantic (or functional, or cognitive) notions.

On the basis of these considerations, I have delineated the domain of my typology of NP-conjunction in the following fashion:

(2) Definition of the domain:
A sentence contains a case of NP-conjunction if
(a) it describes a single occurrence of an event (action, state, process, etc.), and if
(b) this event is predicated simultaneously of two (and no more) participant referents, which are conceived of as separate individuals.

To illustrate how this definition delimits the range of constructions to be included in the database, we can take English as an example. First, the requirement that the two participant referents be conceived of as separate individuals rules out sentences with dual, plural, or collective subjects, such as They left or The two men left. Secondly, the definition stipulates that a clausal or sentential conjunction such as (3) will be left outside of the database.

(3) English (Indo-European, West Germanic)
John left and Mary left

It is true that this sentence ascribes the same event type to two different participants, and moreover the sentence has a possible simultaneous reading. Thus, sentence (3) meets the second condition in the definition. However, it fails to meet the first condition, in that it does not describe a single event; instead, it describes two occurrences of the same event-type. In contrast, a sentence like (4) has this “single event” interpretation as one of its readings.
Among other things, the sentence can be interpreted as implying that John and Mary left together, and it is this reading which is meant to be covered by definition (2). As a result, a sentence like (4) will be admitted in our database as one of the typologically relevant facts of English.

It should be noted that, whereas sentence (4) constitutes a relevant fact, a sentence like (5) does not, even though for this sentence a “together”-reading is certainly possible:

\[(4)\] \hspace{0.5cm} \text{English (Indo-European, West Germanic)} \hspace{0.5cm} \text{John and Mary left}\]

The reason for this limitation to cases of just two participants is completely pragmatic. During my investigation I have encountered quite a few languages in which the encoding of “enumerations” exhibits special formal properties which are absent from the encoding of “participant pairs”. Several languages have a specific marker which indicates that the enumeration given need not be exhaustive and hence creates a sort of “et cetera”-reading. In all the relevant languages this marker cannot be employed in constructions with two participants. An example is the Amazonian language Andoke, in which “multiple participant” constructions have the form of a mere juxtaposition of NPs. In enumerations, however, all NPs in the series obligatorily receive the marker -añe:

\[(5)\] \hspace{0.5cm} \text{English (Indo-European, West Germanic)} \hspace{0.5cm} \text{John, Bill, Harry, Mary, and Susan left}\]

Furthermore, in many languages the use of the coordinating particle is different if more than two participants are involved. As we saw in sentence (5), English has the possibility of restricting the use of the conjunctor and in enumerations to the last member of the series, while the other NPs are juxtaposed. This juxtapositional strategy is not an option for a “two participant” sentence like (4). In short, there is good reason to assume that NP-conjunctions with three or more
members form a special case. I have decided to regard these special formal properties of enumerations as a concomitant phenomenon which complicates the typology beyond necessity, and hence I have excluded such constructions from my database. Of course, there is no telling whether this is a correct decision. It may turn out later that the formal properties exhibited by enumerations are of key importance to our understanding of the typology of NP-conjunction as a whole. However, in running such risks linguistic typology does not differ from any other form of scientific endeavour.

A final, and perhaps somewhat unexpected, implication of the domain definition given above is that it admits constructions into the database which generally are not considered to fall under the heading of "NP-conjunction". For example, in addition to a construction as in (4) English has a construction of the type illustrated in (7):

(7) English (Indo-European, West Germanic)
   a. *John left with Mary*
   b. *Mary left with John*

There are, of course, several semantic-functional differences between the constructions in (4) and (7). For one thing, the sentences in (7) have a necessary "together"-interpretation, whereas this interpretation is only one of the options for (4). Furthermore, in the sentences in (7) one of the participants is backgrounded, a feature which is absent from (4). Nevertheless, constructions like those in (7) meet the conditions given in the definition in the same way as constructions such as (4) do. They describe a single event; in fact, a single-event reading is their only option. Also, they ascribe this event simultaneously to two participants. Hence, our database includes two formally different encodings of the domain of NP-conjunction for English. In the terminology adopted in this paper, we will say that English has TWO DIFFERENT CONJUNCTIONAL STRATEGIES. As we will see, there are many languages which resemble English in this respect. On the other hand, however, we will also come across quite a few languages in which the domain defined in (2) is covered by just one strategy.

3. The Coordinate Strategy

Once we have constructed our crosslinguistic database in accordance with the definition given in (2), we can proceed to distinguish various types of encoding for the domain. It is at this stage that we rubricize the constructions sampled into groups, on the basis of formal differences and similarities. Coupled with this constructional typology is a typological grouping of the languages in the sample, on the basis of whether they allow similar or different construction types.
It is clear that the formation of the constructional typology presupposes decisions as to what the relevant formal criteria for type membership are. Now, as anyone who has ever tried to set up a typology is bound to affirm, the selection of the set of criteria is not the result of a mechanical research procedure. In any crosslinguistic database it will be possible to rubricize the constructions along various different parameters, and the applications of these different parameters usually lead to widely differing language groupings. Again, no hard-and-fast rule of methodology can be called upon here. In practice, the researcher will select his/her criteria under the guidance of "hunches" or "gut feelings" which are hard, if not impossible to explicate, and which are based on expectations of what may lead to a revealing research result. As people may differ in their estimations on this point, it is perfectly possible that one and the same database may give rise to several different typologies, none of which has to be intrinsically "better" than the others.

In my case, I have chosen to found my typology of the domain of NP-conjunction upon the contrast between the two encoding forms as illustrated for English in sentences (4) and (7). The first of these sentences provides an example of what I will call THE COORDINATE ENCODING STRATEGY of the domain. A fundamental formal characteristic of this strategy is that it encodes the two participants in the construction by way of NPs with equal structural rank. Thus, the two NPs involved are not differentiated as to syntactic function; they have the same thematic role, and in languages in which such NPs receive case marking they will both have the same case. Typically, though not necessarily, the two NPs in such constructions can be seen to form a constituent, viz., a coordinate (plural or dual) NP. As a result of this, they typically govern dual or plural number agreement on predicates, if they have a grammatical function for which this agreement is defined. Furthermore, the two NPs are commonly subject to the Coordinate Structure Constraint as formulated in Ross (1967), which forbids NP-extraction from such constructions:

(8) English (Indo-European, West Germanic)
   a. *Who did you see and Mary?
   b. *The woman that I saw and Mary...

It is possible to subcategorize coordinate NP-structures on the basis of the linking device which they employ. As we have seen in the example from Andoke (see (6a)), there are languages with juxtaposition or zero-marking of such structures. Traditionally, the term "asyndeton" is employed to refer to such constructions. Quite a few languages in my sample can be shown to have this encoding possibility as at least one of their options, but closer inspection reveals that asyndetic NP-coordination is nonetheless a minor strategy. For one thing, obligatory use of this variant is very rare. Apart from Andoke, I have
documented this option in only a few other languages, among which no significant areal cohesion can be defined. Examples are:

(9) Awtuw (Papuan, Sepik)
_Yowmn_ Yawur _du-k-puy-ey_
_Yowmen Yawur DUR-IMPF-hit-IMPF_
‘Yowmen and Yawur are hitting (someone).’ (Feldman 1986: 110)

(10) Gumbainggir (Australian, Pama-Nyungan)
_ŋari ni:gaŋ njammi jaraŋ_
this man woman go.PST
‘This man and woman went together.’ (Smythe 1948: 96)

(11) Maranungku (Australian, Daly)
_mereni kalani ŋeni kili-nya awa_
brother uncle my eat-3PL meat
‘My brother and uncle ate the meat.’ (Tryon 1970: 83)

(12) Mojave (Yuman)
_intay nakut-n²-č ?ahu:t-k idu:-m_
mother father-DEM-NOM good-SS be-NONFUT
‘His mother and father are good.’ (Munro 1976: 162)

(13) Classical Nahuatl (Uto-Aztecan, Aztecan)
__öcenqúiz in ixquich in télpochtli in ichpochtli_
together ART all ART boys ART girls
‘All the youths and the maidens gathered together.’ (Andrews 1975: 367)

(14) Ona-Šelknám (Chon)
_Kačérán, Kormščen viek-án_
Kačérán Kormščen RFL-chase
‘Kačérán and Kormščen chased one another.’ (Tonelli 1926: 72)

(15) Manchu (Altaic, Tungusic)
_muduri tasha daham-bi_
dragon tiger follow-PRS
‘The dragon and the tiger follow.’ (Adam 1873: 66)

In general, zero-marked NP-coordinations vary with an encoding which features one or more overt linking particles. In such cases, the zero option often performs a specific, restricted, function: it is used either in “list-like” enumerations, or in the encoding of NP-pairs which habitually go together and can be said to form some conventionalized whole or “conceptual unit” (Mithun 1988: 332). Thus, pairs like “husband and wife”, “boys and girls”, “horses and cattle”, “bow and arrow”, or “gold and silver” are more likely to be encoded by zero-marking than other, less predictable NP-coordinations. Examples in which asyndetic NP-coordination is limited in this fashion are:
(16) Khalkha (Altaic, Mongolian)
   a. ger baishin
ten building
   'tents and buildings' (Street 1963: 180)
   b. Ceren bolon Bat
Ceren and Batu
   'Ceren and Batu' (Street 1963: 180)

(17) Mandarin (Sino-Tibetan, Sinitic)
   a. zjé yué
sun moon
   'sun and moon' (Mullie 1947: 232)
   b. Lù Wényì gèn wò
Lu Wényì and/with I
   'Lu Wényì and I' (Li & Thompson 1981: 657)

(18) Otomi (Oto-Manguean)
   a. rá dǎda rā nqāq
his father his mother
   'his father and his mother' (Hess 1968: 72)
   b. nú ra ?ánxe ně rā ku
that ART Angela and her brother
   'Angela and her brother' (Hess 1968: 72)

(19) Khasi (Mon-Khmer)
   a. ka yap ka ?im
ART life ART death
   'life and death' (Rabel 1961: 128)
   b. u wadár bad u khun
ART councillor and ART son
   'the councillor and bis son' (Roberts 1891: 142)

(20) Modern Persian (Indo-European, Iranian)
   a. mardhā, zanhā, bacehā
men women children
   'men, women, and children' (Lazard 1957: 200)
   b. dānešjuyān va kārgārān
students and workers
   'students and workers' (Lazard 1957: 200)

As for the areal distribution of zero-marked NP-coordinations, we can note that it appears to be absent from Africa and from at least the western part of Europe. In all other parts of the world it must have been an old encoding option. In fact, it can be demonstrated that its-wane has begun only very recently in
areas such as Central America and Siberia, where the overt coordination markers are clearly of foreign (Spanish or Russian) origin. Even in Indo-European, which in its modern forms allows no or only very marginal zero-coordination, the ancient languages (such as Sanskrit, Old Persian, Ancient Greek, or Latin) and the more "conservative" languages (such as the Baltic languages Latvian and Lithuanian) permit a certain amount of asyndeton. However, the general trend all over the world is that zero-coordination tends to be marginalized into specific functions or is replaced altogether by overt marking strategies. Mithun (1988: 353–357) suggests that this development, which can be attested for NP-coordination and clausal coordination alike, has its source in the global increase in literacy. While the zero-strategy, which is basically intonational, is functionally quite well adapted to spoken language, written language requires a more overt formal marking of syntactic relations.

Overtly marked NP-coordinations can be subcategorized further on the basis of two interacting formal parameters. First, the number of coordinate particles involved gives rise to a distinction between monosyndeton (in which only one marker is present in the construction) and polysyndeton (in which both NPs in the structure have their linking particle). Furthermore, the structural position of the marker or markers may differ from type to type. Among the logical possibilities yielded by combining these two parameters, by far the most prominent option turns out to be the use of a medial connective. In this case, we have one single linking morpheme, which is placed between the two NPs in the construction. For some languages which employ this option, it can be argued that the medial connective has greater structural cohesion with the second NP than with the first, or vice versa. Thus, for example, Ross (1967) and Dik (1968) demonstrate that English and, as illustrated in (4), is in construction with the second NP, rather than with the first NP or with both NPs. However, it must be stressed that this difference in structural cohesion does not lead to a difference in structural rank for the two NPs in the English construction. In other words, and is not a subordinating item in English.

Monosyndetic NP-coordination by way of a medial connective can be encountered all over the globe. It is, of course, the dominating option in the languages of Europe. In some areas, notably Sub-Saharan Africa, the strategy has a minor distribution, but there is hardly any major linguistic area in which it is lacking totally. A number of randomly chosen examples can illustrate this point:

(21) Basque (isolate)

Ander eta Mikel
Ander and Mikel
‘Ander and Mikel’ (Saltarelli 1988: 100)
(22) Finnish (Uralic, Balto-Finnic)
   Pentti ja Pirkko
   Pentti and Pirkko
   'Pentti and Pirkko' (Karlsson 1984: 235)

(23) Cairene Arabic (Afroasiatic, Semitic)
   fagala wi farabijja
   bicycle and car
   'a bicycle and a car' (Gary & Gamal-Eldin 1982: 37)

(24) Dinka (Nilo-Saharan, Nilotic)
   akoon ku miir
   elephant and giraffe
   'the elephant and the giraffe' (Nebel 1948: 110)

(25) Mangbetu (Niger-Kordofanian, Adamawa-Eastern)
   n-ókondó bu n-ɔsùmba
   ART-leopard and ART-rat
   'the leopard and the rat' (Larochette 1958: 33)

(26) Written Mongolian (Altaic, Mongolian)
   qayan kiged qatun
   khan and queen
   'the Khan and the queen' (Poppe 1954: 122)

(27) Turkish (Altaic, Turkic)
   Hasan (ve) Ali ve Zeynep dün sinema-ya
   Hasan and Ali and Zeynep yesterday cinema-to
   gittiler
   go.PAST.3PL
   'Hasan (and) Ali and Zeynep went to the movies yesterday.' (Kornfilt 1997: 115)

(28) Vietnamese (Mon-Khmer)
   ngu-ò-ri giâu cùng ngu-ò-ri nghèo
   people rich and people poor
   'rich people and poor people' (Văn Chinh 1970: 146)

(29) Cebuano (Austronesian, Philippine)
   bir ug tubig
   beer and water
   'beer and water' (Wolff 1967: 164)

(30) Rottinese (Austronesian, East Indonesian)
   boutua-la ma nekahade-la
   syrup.pot-PL and rice.basket-PL
   'syrup pots and rice baskets' (Jonker 1915: 662)
In comparison to medial monosyndeton, other overt coordinate strategies are fairly rare. All strategies at issue feature postposition or suffixation of the connective item or items. Among these options, polysyndeton appears to be the most popular. The strategy can be found in a number of unconnected linguistic areas such as the Caucasus (Abkhaz, Archi, Avar, Chechen, Ubykh), northeastern Africa (Acholi, Amharic, Barea, Beja, Kunama, Nubian, Oromo), Australia and New Guinea (Aghu, Alamblak, Aranda, Asmat, Dyirbal, Kobon, Marind, Yidiny), southern India (Kannada, Tamil, Sinhalese, Vedic) and northeastern Asia ( Ainu, Korean, Japanese). In addition, we find isolated examples in the Americas (Yavapai, Guananano, Imbabura Quechua), in West Africa (Ewe, Temne) and in Burma/eastern India (Burmese, Manipuri). Examples include:

(37) Abkhaz (North-West Caucasian)
\[
s\-\dot{\text{a}}\text{no-}y \quad s\-\dot{\text{a}}\text{ba-}y
\]
my-mother-and my-father-and
‘my mother and my father’ (Hewitt 1979: 65)
(38) Archi (Dagestanian)
_ōs l’onol-u ōs lo-vu_
one wife-and one daughter-and
‘a wife and a daughter’ (Dirr 1928: 266)

(39) Amharic (Afroasiatic, South Semitic)
_qąčč-ēm gošš-ēm_
giraffe-and buffalo-and
‘giraffes and buffaloes’ (Hartmann 1980: 355)

(40) Alamblak (Papuan, Sepik)
_yėn-r-i yėn-t-e_
child-M-and child-F-and
‘a boy and a girl’ (Bruce 1984: 128)

(41) Dyirbal (Australian, Pama-Nyungan)
_bayi yuţi-gara bayi bargan-gara_
CLl-ABS kangaroo-and CLl-ABS wallaby-and
‘a kangaroo and a wallaby’ (Dixon 1972: 230)

(42) Tamil (Dravidian)
_akkaa-vum taŋkacci-yum_
elder.sister-and younger.sister-and
‘elder sister and younger sister’ (Asher 1982: 69)

(43) Sinhalese (Indo-European, Indic)
_noona-y daru-o-y_
wife-and children-and
‘the wife and children’ (Gair 1970: 42)

(44) Ainu (isolate)
_okkay ka menoko ka_
men and women and
‘men and women’ (Refsing 1986: 163)

(45) Korean (Altaic, isolate)
_na hago ne hago kagesso_
1SG and 2SG and go.FUT
‘You and I will go.’ (Ramstedt 1939: 156)

(46) Burmese (Sino-Tibetan, Tibeto-Burman)
_hsi-yò hsaŋ-yò_
oil-and rice-and
‘oil and rice’ (Okell 1966: 191)
Monosyndetic postposing of coordination markers may, of course, in principle take two different forms, depending on whether the item is postposed to the first or to the second NP in the coordination. Both types of construction occur in the sample, but their frequency is low. Moreover, the languages which present one (or both) of these options typically also allow a construction of the polysyndetic type, so that these monosyndetic constructions are best regarded as variants in which one of the markers in the polysyndetic construction can be optionally deleted. Given this, it will be clear that the two monosyndetic variants occur in roughly the same areas as have been listed for the polysyndetic construction. Examples of the "first-NP" subtype are Beja (northeastern Africa), Modern Persian, and Ainu (northern Japan). The "second-NP" subtype is illustrated by NP-coordinations from Latin, Pitjantjatjara (Australian), and Tubu (Saharan).

(47) Beja (Afroasiatic, Cushitic)
   a. ani-wā barūk-wā
      1SG-and 2SG-and
      'you and I' (Reinisich 1893: 195)
   b. mēk-wā lága
donkey-and calf
      'a donkey and a calf' (Reinisich 1893: 195)

(48) Old Persian (Indo-European, Iranian)
Pārsam-čā Mādam-čā
Persia-and Media-and
'Persia and Media' (Meillet & Benveniste 1931: 224)

(49) Modern Persian (Indo-European, Iranian)
mardhā-o zanhā-o bacehā
men-and women-and children
'men, women, and children' (Lazard 1957: 200)

(50) Ainu (isolate)
   a. hapo ka mici ka
      mother and father and
      'mother and father' (Refsing 1986: 163)
   b. urki ka taiki
      flea and louse
      'fleas and lice' (Refsing 1986: 162)

(51) Latin (Indo-European, Italic)
    senatus populus-que
    senate people-and
    'the senate and the people' (Jan Brouwers, personal communication)
(52) Pitjantjatjara (Australian, Pama-Nyungan)

*Henry-ku mama ngunytju puru*

Henry-GEN mother father and

'Henry’s father and mother' (Glass & Hackett 1970: 66)

(53) Tubu (Nilo-Saharan, Saharan)

a. *túrku ye mɔ́lɔ́fúr ye*

jackal and hyena and

'the jackal and the hyena' (Lukas 1953: 166)

b. *wúdén arko ye*

gazelle goat and

'the antilope and the goat' (Lukas 1953: 166)

To round off the discussion of the various manifestations of the Coordinate
Strategy, I can note that monosyndetic preposing on the first NP is not attested
at all in the sample. That is, there do not seem to be languages which conform
to the AND-NP NP scheme. Monosyndetic preposing on the second NP can of
course be claimed to exist in those languages in which the medial connective
seems to be in construction with the second NP; thus, English would be a case
in point. For a number of such languages, there is a polysyndetic preposed
variant of the type AND-NP AND-NP. This variant, which can, among others, be
attested in various languages of Europe, usually has an emphatic or contrastive
function: the English *both-and* construction is a fairly representative instance
of this strategy. Languages in which this preposed polysyndeton is the only
option do not occur in my sample.

In my estimation, the variation in the formal encoding of coordinate NPs is a
self-contained subject which merits a separate typological inquiry, with its own
correlational parameters and its own specific explanatory principles. Since it
is not my aim to present such an investigation here, I will confine myself to
a few general observations. First, it is to be expected that the types in this
typology will turn out to exhibit at least some partial correlations with word
order options. In my sample, the two correlations formulated in (54) receive
substantial empirical support:

(54) Word order correlations for Coordinate NP-types:

a. If a language has a (monosyntetically or polysyntetically) post-
posed coordination marker, then that language is verb-final.

b. If a language is verb-initial, and if it has an overtly marked Coor-
bine Strategy for NPs, that strategy will involve a medial con-
nective.

Furthermore, it is almost certain that this typology has a diachronic dimen-
sion. Mithun (1988) shows that NP-coordinators derive from various sources
by way of a process of grammaticalization. In particular, she mentions NP-coordinators which have their origin in grammaticalized comitative markers ("with"), or in grammaticalized sentence adverbials with the original meaning 'also, too, as well', or 'furthermore, then, moreover'. To this list I can add several other diachronic possibilities. There are languages in which a coordinating particle for NPs clearly arises from numerals or quantifiers like 'two', 'both', or 'all', which are grammaticalized to a greater or lesser degree. A concentration of these cases is found in the Australian-Papuan area, but incidental instances are attested for other areas as well.

(55) Waskia (Papuan, Adelbert Range)
\textit{ane kadi mu ili}
1SG man the all
'the man and I' (Ross & Natu Paol 1978: 40)

(56) Tok Pisin (English-based creole, Melanesia)
\textit{Yoannes tupela Yosep}
Yoannes two Yosep
'John and Joseph' (Mihalic 1957: 172)

(57) Arrernte (Australian, Pama-Nyungan)
\textit{Augustine therre Duncan therre}
Augustine two Duncan two
'Augustine and Duncan' (Wilkins 1989: 371)

(58) Written Mongolian (Altaiic, Mongolian)
\textit{bi ci qoyar}
1SG 2SG two
'you and I' (Poppe 1954: 111)

(59) Luiseño (Uto-Aztecan)
\textit{hunwut 'awaal weh}
bear dog both
'the bear and the dog' (Hyde 1971: 46)

Related to this "numeral" strategy is a "pronominal" strategy, in which dual or plural personal pronouns are gradually grammaticalized into coordinative markers for NPs:

(60) Waropen (Austronesian, New Guinea)
\textit{mangha kisi bingha}
man.ART 3DU woman.ART
'the man and the woman' (Held 1942: 90)

(61) Tiwi (Australian, Tiwi)
\textit{wuta Tapara Waijai}
3PL Tapara Waijai
'Tapara and Waijai' (Osborne 1974: 72)
AND-languages and WITH-languages

(62) Sedang (Mon-Khmer)
prei klá prei koa
3DU tiger 3DU turtle
the tiger and the turtle' (Smith 1979: 92)

(63) Siuslaw (isolate, Oregon)
sqūma' lqal'ōā'm-a"x
pelican sea.gull-3DU
'the pelican and the sea gull' (Frachtenberg 1922: 610)

(64) Chukchi (Chukotko-Kamchatkan)
Gi'philin e'rrn Tiₐ'irgin
Sunset 3PL Dawn
'Sunset and Dawn' (Bogoras 1922: 856)

A further source for NP-coordinators lies in non-finite forms of verbs meaning 'to be' or 'to exist'. The item mi-cha 'and' in Choctaw is a participial or switch-referential form of the verb mi 'to be' (Nicklas 1974: 257). In Korean, one of the items used as connectives between NPs has its origin in a (simultaneous) converb of the verb ha 'to be, to do'. Similarly, Classical Mongolian employs converbal forms of the verb stems'ki 'to do' or bollbu 'to be, to exist'.

(65) Choctaw (Muskogean)
ano micha sashki
1SG and my.mother
'my mother and I' (Nicklas 1974: 257)

(66) Korean (Altaic, isolate)
na hago ne hago
1SG and 2SG and
'you and I' (Ramstedt 1939: 156)

(67) Written Mongolian (Altaic, Mongolian)
bayşı büged šabi
teacher and pupil
'the teacher and the pupil' (Poppe 1954: 122)

Finally, there are a few cases where the NP-coordinator seems to be a specialization of the function of a general focus-marking particle:

(68) Manam (Austronesian, Melanesian)

a. moáne-be áine di-púra
man-and woman 3PL-arrive
'The men and the women arrived.' (Lichtenberk 1983: 365)

b. wabubu-ló-be i-púra
night-at-FOC 3SG-come
'It was at night that he came.' (Lichtenberk 1983: 483)
(69) Kabyl (Afroasiatic, Berber)
a. *agerfiou d' oubarer’*
   raven and fox
   ‘the raven and the fox’ (Hanoteau 1906: 91)
b. *netsa d’ agellid’ en temourth agi*
   3SG FOC king of country this
   ‘It is him who is the king of this country.’ (Hanoteau 1906: 88)

As a general conclusion, we can state that the sources of coordinative elements for NPs seem, at least at first sight, to be heterogeneous. It is reasonable to assume that these different origins may still have their bearing on the synchronic status of coordinate markers in individual languages, and that the syntactic behaviour of such markers is still to a certain degree influenced by formal properties of their source-items. In my project I have decided to ignore the formal variation in coordinate NP-marking, so that a construction will be classified as coordinate regardless of its individual morphosyntactic make-up.

4. The Comitative Strategy

In opposition, and often in addition, to the Coordinate Strategy the domain can be encoded by what I will call THE COMITATIVE STRATEGY. In its essential features, this strategy is diametrically opposed to the Coordinate Strategy. Thus, under the Comitative Strategy the two participants in the event are morphosyntactically encoded as NPs of unequal structural rank. While one of the NPs can take any case role, the other NP is invariably encoded as the head of an oblique NP. A prototypical characteristic of comitative structures is that the two NPs involved are not part of the same constituent. As a result, they typically do not force dual or plural agreement on predicates, and neither of the two NPs is subject to restrictions on extraction rules as defined by the Coordinate Structure Constraint.

In contrast to the Coordinate Strategy, the Comitative Strategy is much more uniform formally across languages. In the overwhelming majority of sampled languages, the Comitative Strategy manifests itself by way of an oblique marker ‘with’ on one of the participant NPs. If the language has a coordinate NP-strategy as well, the comitative marker is, in the typical case, not identical to the item used as an NP-coordinator. Depending on general morphosyntactic features of the language the comitative marker may take the form of an adposition or an affix. Since this manifestation of the Comitative Strategy is basically the same as in English, I trust a few random examples may suffice as illustration:
AND-languages and WITH-languages

(70) Gallilarese (Papuan, Halmahera)
mo liho-ka dé ma baba
3SG.F return-PF with her father
'She returned with her father.' (van Baarda 1908: 156)

(71) Shuswap (Salish)
ckícx-Ø ml-kʰ*séltkn-s
come-3SG with-friend-his
'He came with his friend.' (Kuipers 1974: 147)

(72) Kobon (Papuan, East Highlands)
ne aip ar-nab-in
you with go-FUT-1SG
'I will go with you.' (Davies 1981: 73)

(73) Bilin (Afroasiatic, Cushitic)
kū-di fardiyā
you-with go.1SG.FUT
'I will go with you.' (Reinisch 1882: 100)

Although "dependent marking" (Nichols 1986) is the predominant encoding option in comitative constructions, there are some languages which deviate from this pattern. A "head-marking" strategy, in which the comitative marker is incorporated into the predicate, is encountered in languages of the North-West Caucasus:

(74) Abkhaz (North-West Caucasian)
à-č’kʰ’on so-yò-c-ce-yt’
the-boy 1SG-3SG-with-go-AOR
'I went with the boy.' (Hewitt 1979: 115)

(75) Ubykh (North-West Caucasian)
ỳò a-u-ji-k’a-qa
2SG 3SG-2SG-with-come-PF
'He came with you.' (Dumézil 1931: 17)

In some cases, this incorporation leads to transitivity of the predicate, so that the "comitative" NP takes on the syntactic status of a direct object. Assorted languages of the Americas present this option:

(76) West Greenlandic (Eskimo-Aleut)
miiqqa-t tiki-uti-nngil-akka
child-PL come-with-NEG-1SG.ACT/3PL.PAT
'I didn’t come with the children.' (Fortescue 1984: 215)
(77) Peñoles Mixtec (Mixtecan)
kwáʔ?-ndiʔ?-dé ʔadiʔ?i
    'He goes with his wife.' (Daly 1973: 63)

(78) Otomi (Oto-Manguean)
bi-mé-wi ra ku xá ra tāí
    'He went to the market with his brother.' (Hess 1968: 147)

(79) Ona-Šelknám (Chon)
Kkokós telken okel-enen
    'Kokós goes with the boys.' (Tonelli 1926: 58)

That comitative constructions are liable to turn into (or, alternatively, may have their source in) transitive patterns is manifested even more explicitly in languages where the comitative NP is constructed as the direct object of a verb 'to accompany/to have/to hold/to take/to follow' in a serialization construction. Obviously, such cases will be found in areas where verb serialization is rampant anyway. Thus, West Africa, Eastern Austronesia, and Sino-Tibetan are places to search for this phenomenon:

(80) Igbo (Niger-Kordofanian, Kwa)
    há sò anyí gá-á 'Ábá
    'They went to Aba with us.' (Welmers 1973: 369)

(81) Keiese (Austronesian, Moluccan)
    jaw oe-hoev o
    1SG.go 1SG-accomp 2SG
    'I will go with you.' (Geurtjens 1921: 59)

(82) Mandarin (Sino-Tibetan, Sinitic)
    wò gēn tā kāi-wānxìào
    'I am joking with him/her.' (Li & Thompson 1981: 78)

As with the Coordinate Strategy, I take it that the internal formal variation exhibited by the Comitative Strategy has an independent motivation, and that the correlational parameters and explanatory principles involved constitute a research subject of its own. For our purposes, then, we can conclude that the formal encoding of the domain defined in (2) is covered by two conjunctional strategies, which contrast in the following lexical, morphological, and syntactic characteristics:
The typological observations presented in the following sections will be based upon the dichotomy between these two encoding options. It should be noted in advance that, in quite a few languages, a differentiation between these two strategies cannot be stated with razor-sharp precision. The formal features listed in (83) formulate prototypical properties of the two strategies, but one has to make allowance for in-between cases. In other words, just like virtually any worthwhile distinction in typological linguistics, the contrast between the Coordinate Strategy and the Comitative Strategy must be regarded as a formulation of the extreme, and focal, positions on a continuum. In the next section, some of the factors which underlie the intermediate cases on this scale will be elucidated.

5. AND-languages and WITH-languages

The distinction between the Coordinate Strategy and the Comitative Strategy can be put to use in the assessment of the typological status of the sampled languages in regard to the encoding of the defined domain. A first observation to make is that, with only a few exceptions, all languages in the sample appear to have the possibility of employing the Comitative Strategy. Thus, having a Comitative Strategy does not constitute an interesting crosslinguistic parameter. What is interesting is the fact that a number of languages appear to use this Comitative Strategy as the only way to encode the domain. That is, in these languages the only way to encode the situation in which a single event is ascribed simultaneously to two different participants is to use a non-balanced, non-constituent, construal of the two NPs involved: coordination of NPs, in the structural ("balanced") sense, is not an option. Languages of this type will be called WITH-languages. In the next section I will attempt to delineate this language type in areal terms. For the moment, I will limit myself to the presentation of a number of clear examples. The crucial fact of each of the languages exemplified in (84)-(95) is that the distinction between 'John and Mary left' and 'John left with Mary' is not structurally recognized. Instead, all these languages take the comitative encoding as the only available option.③
Leon Stassen

(84) Mandarin (Sino-Tibetan, Sinitic)
Lu Wényí gēn wǒ qù huá-chuán
Lu Wenyi with 1SG go row-boat
'Lu Wenyi and I went rowing/Lu Wenyi went rowing with me.' (Li & Thompson 1981: 657)

(85) Lushai (Sino-Tibetan, Tibeto-Burman)
mi pahnih ka hnēhna lo-va an kal
man two 1SG with field-to 3PL go
'The two men and I went to the field/The two men went to the field with me.' (Lorrain & Savidge 1898: 33)

(86) Samoan (Austronesian, Polynesian)
o lo'o 'a'ai Malia ma loane i fa'i
PROG eat Mary and/with John at banana
'Mary and John are eating bananas/Mary is eating bananas with John.' (Marsack 1975: 119)

(87) Chamorro (Austronesian, Philippine)
maapos si Juan yan si Maria
leave.PST ART Juan and/with ART Maria
'Juan and Maria left/Juan left with Maria.' (Topping 1973: 146)

(88) Nkore-Kiga (Niger-Kordofanian, Bantoid)
n-ka-za-yo na Mugasho
1SG-PST-go-there and/with Mugasho
'I and Mugasho went there/I went there with Mugasho.' (Taylor 1985: 58)

(89) Akan (Niger-Kordofanian, Kwa)
Kwesi nye Amba a-ba
Kwesi with Amba PST-come
'Kwesi and Amba have come/Kwesi has come with Amba.' (Balmer & Grant 1929: 150)

(90) Haida (isolate)
studjá'gadān-ai da ñat la qá'xuls
woodpecker-DEF with he go.out
'He and the woodpecker went out/He went out with the woodpecker.' (Swanton 1911: 245)

(91) Chatino (Oto-Manguean)
nší?yu32 ne?3 yka3 lo?o1 ta?a23
cut.IMPF he wood with relative
'He and his brother were cutting wood/He was cutting wood with his brother.' (Pride 1965: 82)
AND-languages and WITH-languages

(92) Jacaltecc (Mayan)
\[\text{chin to boj hach}\]
1SG go and/with 2SG
'You and I go/ I go with you.' (Craig 1977: 30, 32)

(93) Goajiro (Arawakan)
\[\text{ayatájirrásh Jurránsiku nü-má Juan}\]
work.RECIP Francisco him-with Juan
'Juan works with Francisco/Juan and Francisco are working together.'
(Jusayú 1975: 47)

(94) Kaingang (Ge)
\[\text{tä panh vý ?inh mrè tî mû}\]
your father 3SG.SBJ 1SG with go DYN
'Your father and I will go/Your father will go with me.'
(Wiesemann 1972: 145)

(95) Urubu-Kaapor (Tupi)
\[\text{Nasui riki ihê namô i-hon}\]
Nasui EMPH 1SG with 3-go
'Nasui and I went/Nasui went with me.' (Kakumasu 1986: 349)

Opposed to WITH-languages, there are, of course, many languages like English, in which there is a clear differentiation between the comitative and the coordinative encoding of the domain. Such languages will be called AND-languages. Practically all the languages of Europe are members of this class, but the phenomenon can be encountered in many other linguistic families and areas. A small selection of examples, given in (96)–(106), will demonstrate this fact.

(96) Xanty (Uralic, Ugric)
\[\text{a. mä pä äsem matsamôn}\]
1SG and father.my go.1DU.PST
'My father and I went out.' (Rédei 1965: 81)
\[\text{b. mä üt-n pilxü-em piln jax-lam}\]
1SG wood-LOC friend-my with walk-1SG.PRS
'I walk in the wood with my friend.' (Rédei 1965: 76)

(97) Manchu (Altaic, Tungusic)
\[\text{a. muduri tasha daham-bi}\]
dragon tiger follow-PRS
'The dragon and the tiger follow.' (Adam 1873: 66)
\[\text{b. tere irgen i niyalma i emgi te-he}\]
that people GEN men GEN with stay-PF
'He stayed with the men of,that people.' (Adam 1873: 60)
(98) Maltese (Afroasiatic, Semitic)
   a. *is-sikkina u l-furketta ssaddu*
      ART-knife and ART-fork grow.rusty.3PL.PF
      ‘The knife and the fork have grown rusty.’ (Aquilina 1965: 189)
   b. *gie ghandi b-il-kelb mieghu*
      come.3SG.PF at.me with-ART-dog with.him
      ‘He came to my house with his dog.’ (Aquilina 1965: 109)

(99) Somali (Afroasiatic, Cushitic)
   a. *waraabe iyo dawaco ayaa wada ugaadshsi*
      hyena and jackal FOC together hunt
tegey
go.3M.PST
      ‘A hyena and a jackal went hunting together.’ (Serzisko 1984: 165–166)
   b. *na la ra‘*
      1PL with come.IMP
      ‘Come with us!’ (Kirk 1905: 75)

(100) Tamil (Dravidian)
   a. *Raaman-um Murukan-um vantaagka*
      Raman-and Murugan-and come.PST.3PL
      ‘Raman and Murugan came.’ (Asher 1982: 67)
   b. *naan appaa-kuu{e Cennekki pooreen*
      1SG father-with Madras.DAT go.PRS.1SG
      ‘I am going to Madras with my father.’ (Asher 1982: 112)

(101) Thai (Austro-Asiatic, Kam-Tai)
   a. *lug læ? pûa duu Toorá?Tát*
      uncle and aunt watch television
      ‘Uncle and aunt watch television.’ (Warotamasikkhadit 1972: 43)
   b. *p%m kàp mæe*
      father go with mother
      ‘Father goes with mother.’ (Warotamasikkhadit 1972: 44)

(102) Malagasy (Austronesian, West Indonesian)
   a. *misotro Rabe sy Rakoto*
      drink Rabe and Rakoto
      ‘Rabe and Rakoto drink.’ (Ed Keenan, personal communication)
   b. *niaraka tami-ny aho*
      go.PST with-him I
      ‘I went with him.’ (Malzac 1960: 83)
(103) Arrernte (Australian, Pama-Nyungan)
   a. Kwementyaye Gabriella petye-ke
      Kwementyaye Gabriella come-PST
      ‘Kwemetyaye and Gabriella came.’ (Wilkins 1989: 405)
   b. Les Kathy-ngə petye-rnə
      Les Kathy-with come-PST
      ‘Les arrived with Kathy.’ (Wilkins 1989: 186)

(104) Yaqui (Uto-Aztecan)
   a. Peo into Maria aman saha-k
      Peo and Maria there go-PF
      ‘Peo and Maria went there.’ (Lindenfeld 1973: 122)
   b. ine-po in abachi-mak weye
      1SG my brother-with go
      ‘I am going with my brother.’ (Lindenfeld 1973: 19)

(105) Imbabura Quechua (Andean)
   a. Marya-pash Juzi-pash Utavalu-man ri-rka
      Maria-and José-and Otavalo-to go-3PST
      ‘Maria and José went to Otavalo.’ (Cole 1982: 80)
   b. ūnaka-ka wambra-wan puri-ni
      1SG-TOP child-with walk-1PRS
      ‘I walk with the child.’ (Cole 1982: 80)

(106) Guarani (Tupi)
   a. huā ha Maria o-jo-hayhu
      Juan and Maria 3-RECIP-love
      ‘Juan and Maria love one another.’ (Krivoshein de Canese 1983: 95)
   b. a-há ta ne-nivé
      1SG-go FUT 2SG-with
      ‘I will go with you.’ (Gregores & Suárez 1967: 143)

The above examples show that both for WITH-languages and for AND-languages clear, “pure”, instances can be found in the sample. At the same time, however, we must observe that the typological status of these two types is probably not equally well-established. First, there is a marked difference of frequency between the two types, as there are roughly twice as many AND-languages in the sample as there are WITH-languages. Secondly, and more importantly, there is a notable discrepancy in the stability of the types. In general, AND-languages can be said to be stable diachronically and “pure” in their synchronic state: there is a sharp delineation of the two available strategies, along the lines of the oppositions listed in (83). On the other hand, “pure” instances
of WITH-languages are relatively rare. For a considerable number of such languages, some process of grammaticalization of the comitative encoding can be attested. The general outcome of this process in all relevant languages is that it effectuates a shift from a monolithic encoding of the domain towards a dualism of encoding. To put it in rather informal, anthropomorphic, and therefore unquotable terms: WITH-languages do not have a Coordinate Strategy, but they would like to have one. To this end, they tend to differentiate the structural features of the Comitative Strategy, by changing one or more features of that strategy in some contexts in the direction of the features of the Coordinate Strategy. In this way, the language acquires a two-strategy encoding of the domain, in which one of the strategies is still purely comitative, while the other is some hybrid between the comitative and the Coordinate Strategy.

The grammaticalization of the Comitative Strategy in WITH-languages is of course a gradual process. Moreover, the structural features which are changed in the course of this process are different for various groups of WITH-languages. In structural terms, the grammaticalization of a comitative encoding pattern into a “coordination-like” construction prototypically involves the creation of a single constituent, in which both the ‘with’-phrase and the non-comitative NP are included, and in which the two NPs gradually come to be regarded as being of equal structural rank. However, there are several routes by which this end result can come about. In some cases, the creation of a coordinate NP involves “movement” of the ‘with’-NP from its canonical position in the sentence. In others, differentiation in verb agreement (mainly between singular and dual/plural) may be the main formal manifestation of the process. Still other WITH-languages signal the creation of a coordinate NP-structure by “doubling” the comitative marker in coordinations, thereby overtly indicating the equality in rank of the two NPs; and there are, of course, various conceivable combinations of these structural processes. In the end, however, all these languages come to exhibit a split between the two strategies in all relevant features except the lexical linking item. Thus, these mixed WITH-languages evolve towards AND-status. Indeed, one might call them AND-languages, if it were not for the fact that the particle employed in their coordinate NP-structure is still lexically identical to their comitative marker.

There are reasons to assume that the grammaticalization process of the Comitative Strategy is sensitive to general typological characteristics. To be precise, there are two typological oppositions which seem to direct the exact form of the process. First, the grammaticalization in WITH-languages which allow Person-Number-Gender (PNG) agreement on verbs is typically different from the process in languages which lack such agreement. Furthermore, basic word order seems to play a part, in that the process in SVO languages typically has a formal outcome which differs from that in SOV or VSO languages. In the following paragraphs I will examine the interaction of these two parameters.
and their influence on the creation of mixed WITH-languages. In doing so, I will employ a number of "pattern schemes", which represent the starting point for the grammaticalization of the comitative construction in various groups of WITH-languages. I assume that, in these pattern schemes, the 'with'-NP has the form of an adverbial phrase and that, therefore, it will occupy the structural position which is canonical for adverbial phrases. Furthermore, to keep matters simple I limit myself here to sentences with intransitive predicates. The non-comitative NP in the structure is constructed as the grammatical subject.

First, let us consider the case of WITH-languages with basic SVO word order. Depending on whether such languages lack or allow subject-agreement marking on predicates, the pattern scheme for such languages can be formulated as either (107a) or (107b):

(107) a. NP1 V with-NP2 
b. NP1 V-SG with-NP2

Common to both types is the feature that the two "participant" NPs are not contiguous; they are separated from one another by the predicate. Since grammaticalization of comitative structures requires the creation of a single constituent, and since constituency presupposes the contiguity of lexical items, grammaticalization in SVO WITH-languages will typically involve a "shift" of the comitative NP2. Thus, the comitative phrase is allowed to occupy a position in front of the verb, which makes it contiguous to NP1. For many languages of this type, this shift of the 'with'-NP is sufficient in itself to create a new, coordinate, sentence pattern, as it will typically be the case that comitative phrases are the only adverbial phrases which are permitted to undergo this fronting into pre-verbal position. Once comitative phrases are allowed to appear contiguous to NP1, speakers may start to reanalyze the string NP1 with-NP2 as a single constituent. It is reasonable to assume that this reanalysis is a gradual process, and languages may differ as to the extent in which it has proceeded. If the language has subject agreement, the completion of the process may be marked by allowing dual/plural marking on the predicate instead of singular marking. In short, by way of grammaticalization SVO WITH-languages may come to acquire the structural patterns (108a–b) in addition to the patterns (107a–b):

(108) a. [ NP1 with-NP2 ] V 
b. [ NP1 with-NP2 ] V-DU/PL

Major areas in which a grammaticalization process of this type seems to be under way, or may even have been completed, are Sub-Saharan Africa (including Sranan, a creole language from Surinam which has a clear West-African substratum) and the languages of the Indonesian archipelago. Furthermore, the phenomenon is attested for at least some Mon-Khmer languages, such as Khasi, Sedang, and Vietnamese. A selection of examples includes:
(109) Hausa (Afroasiatic, Chadic)

a. *Audù yā zō dà ṭuba-n-sà*
   Audu 3SG.M.PST come with father-his
   ‘Audu came with his father.’ (Gerrit Jan Dimmendaal, personal communication)

b. *Audù dà ṭuba-n-sà su-n zō*
   Audu and/with father-his 3PL.PST come
   ‘Audu and his father came.’ (Gerrit Jan Dimmendaal, personal communication)

(110) Sango (Niger-Kordofanian, Ubangian)

a. *lo goe na kóli só*
   3SG go with man this
   ‘She goes with this man.’ (Samarin 1967: 73)

b. *lo na zo vokó a-gá*
   3SG and man black SBJ-come
   ‘He and the Black Man came.’ (Samarin 1967: 88)

(111) Babungo (Niger-Kordofanian, Bantoid)

a. *Làmbi gò tāa yiwiŋ ghọ Ndùlá*
   Lambi go.PF to market with Ndula
   ‘Lambi went to the market with Ndula.’ (Schaub 1985: 87)

b. *Làmbi ghọ Ndùlá gò tāa yiwiŋ*
   Lambi and/with Ndula go.PF to market
   ‘Lambi and Ndula went to the market.’ (Schaub 1985: 87)

(112) Shona (Niger-Kordofanian, South-East Bantu)

a. *wa-ka-dzoka na-ye*
   3SG-PST-return with-3SG
   ‘He returned with him.’ (Fortune 1955: 399)

b. *Sadza na-Gutu va-uya*
   Sadza and/with-Gutu 3PL-come.PF
   ‘Sadza and Gutu have come.’ (Fortune 1955: 419)

(113) Sranan (English-based creole)

a. *mi kojri nanga mi brada*
   1SG stroll with my brother
   ‘I took a walk with my brother.’ (Donicie 1954: 119)

b. *mi nanga mi brada kojri*
   1SG and/with my brother stroll
   ‘My brother and I took a walk.’ (Donicie 1954: 107)

(114) Bahasa Indonesia (Austronesian, West Indonesian)

a. *Ali datang dengan ibu-nja*
   Ali come with mother-his
   ‘Ali came with his mother.’ (Kähler 1965: 155)
b. *Siti dengen Amat*

'Siti and Amat' (Macdonald 1976: 109)

(115) Toradja (Austronesian, East Indonesian)

a. *siko da ndeku pai tau se’e*

2SG must go.up with people DEM

'You must go to the highlands with these people.' (Adriani 1931: 397)

b. *asu pai wawu mombe-luku*

do.g and/with pig RECIP-fight

'Dogs and pigs fight with each other.' (Adriani 1931: 418)

(116) Gallilarese (Papuan, Halmahera)

a. *mo liho-ka dé ma bāba*

3SG.F return-PF with her father

'She returned with her father.' (van Baarda 1908: 156)

b. *ai baba dé ngoi awa i tupu*

my father and/with my mother they.me burn

'My father and my mother have burnt me.' (van Baarda 1908: 148)

(117) Vietnamese (Mon-Khmer)

a. *Giáp di chọi vó-i Āt*

Giap go stroll with At

'Giap went out for a walk with At.' (Văn Chinh 1970: 241)

b. *Giáp vó-i Āt là anh em ruợt*

Giap and/with At COP brothers

'Giap and At are brothers.' (Văn Chinh 1970: 146)

A few special cases of the process can be documented. In Tera, Acholi, and Tolai, shift of the comitative NP does not seem to be possible. However, these languages nonetheless differentiate between a singular and a dual-plural agreement strategy:

(118) Tera (Afroasiatic, Chadic)

a. *Ali wà da varan ndә woy-a*

Ali PF move himself with boy-the

'Ali ran away with the boy.' (Newman 1970: 55)

b. *Ali wà da vanda ndә woy-a ku*

Ali PF move themselves and/with boy-the PL

'Ali and the boy ran away.' (Newman 1970: 55)
(119) Acholi (Nilo-Saharan, Nilotic)

a. Œkeelọ ó-cítọ paàco gün ki láminne
   Okeelo 3SG-go home 3PL with sister.his
   ‘Okeelo went home with his sister.’ (Crazzolara 1955: 66)

b. Œkeelọ gi-cítọ paàco gün ki láminne
   Okeelo 3PL-go home 3PL with sister.his
   ‘Okeelo went home with his sister.’ (Crazzolara 1955: 66)

(120) Tolai (Austronesian, Melanesian)

a. nam ra tutana i ga rovoi ma ra pap
   DEM ART man 3SG PF hunt with ART dog
   ‘That man hunted with his dog.’ (Mosel 1984: 176)

b. Telengai dir rovoi ma ra pap
   Telengai 3DU hunt and/with ART dog
   ‘Telengai and the dog hunted.’ (Mosel 1984: 176)

Conversely, in spite of their basic SVO order Khasi and Acehnese appear to allow pre-verbal position of comitative NPs. Again, it is by way of a difference in number agreement that the coordinate and comitative patterns are distinguished:

(121) Khasi (Mon-Khmer)

a. u nonghikai bad ki khynnah u la wan
   ART teacher with ART children 3SG PF come
   ‘The teacher has come with the children.’ (Roberts 1891: 143)

b. u nonghikai bad ki khynnah ki m put
   ART teacher and/with ART children 3PL NEG yet kloi
   ready
   ‘The teacher and the children are not ready yet.’ (Roberts 1891: 142)

(122) Acehnese (Austronesian, West Indonesian)

a. lôn ngön-adék lôn-jak bak-sikula
   1SG with-younger.brother 1SG-go to-school
   ‘I go to school with my younger brother.’ (Durie 1985: 124)

b. lôn ngön-adék meu-jak bak-sikula
   1SG and/with-younger.brother 1PE-go to-school
   ‘My younger brother and I go to school.’ (Durie 1985: 124)

In contrast to languages with SVO word order, WITH-languages with basic SOV or VSO word order do not face the contiguity problem. Comitative phrases, like all adverbial phrases, are typically positioned at the same side of the predicate as subjects are, so that the pattern schemes for such languages can be formulated as in (123)–(124):
The WITH-languages which take one of these patterns as a starting point for the development of a “coordinate-like” syntagm are situated in a number of different areas. Among verb-final languages, we can point to a concentration of the phenomenon in Tibeto-Burman (Tibetan, Newari, Lushai, Burmese), in northeastern Asia and northern North America (Even, Nenets, Yakut, Japanese, Inuktitut), in northeastern Africa (Kanuri, Kunama, Nubian, Bilin), in northern New Guinea (Amele, Kapauku-Ekagi, Kobon), and in some other languages of North America (Navaho, Chemehuevi) and South America (Cuzco Quechua, Piro, Canela-Krahô). The phenomenon is considerably less frequent in verb-initial languages. We encounter it mainly in a number of languages of the Americas, some of which are situated on the west coast of North America (Kwakwala, Shuswap), while others belong to groupings in Central America (Mixtec, Zapotec). Finally, the area of northeastern Africa is, apart from the SOV-languages mentioned above, represented by the verb-initial Nilotic language Turkana.

It will be clear that WITH-languages which would want to grammaticalize their comitative structures on the basis of the patterns (123a) or (124a) can resort neither to differentiation in the position of the ‘with’-phrase nor to differentiation in terms of PNG-agreement. In other words, the morphosyntactic means for marking the creation of a new “coordinate-like” syntagm for NPs are severely limited for such languages. It is my impression that, especially in verb-initial languages, these circumstances lead to the preservation of the original “pure” WITH-status. Alternatively, of course, further analysis of these cases may show that there are differences in the degree of structural cohesion between the subject-NP and the ‘with’-phrase which just happen to be morphologically unmarked.

A restricted number of WITH-languages with a pattern of the type (123a) have developed a special morphosyntactic device to circumvent the limitations imposed on their type. In Burmese, Tibetan, Newari, Japanese, and Canela-Krahô the creation of a coordinate NP-structure involves the “doubling” of the comitative marker, so that this marker becomes associated with both NPs instead of just with NP2. Thus, these languages have developed a pattern like (125) in addition to their comitative pattern (123a):

(125) NP1-with NP2-with V
The double occurrence of the comitative marker in (125) signals the equality in rank, and hence the essentially coordinate status, of the syntactic relation between the two NPs. Examples are:

(126) Japanese (Altaic, isolate)
   a. John ga Mary to benkyoosita
      John SBJ Mary with studied
      ‘John studied with Mary.’ (Kuno 1973: 103)
   b. John to Mary to Tom to ga kita
      John with Mary with Tom with SBJ came
      ‘John, Mary, and Tom came.’ (Kuno 1973: 112)

(127) Burmese (Sino-Tibetan, Tibeto-Burman)
   a. ñme-ne’ tba-te
      mother-with go-NONFUT
      ‘(I) went there with (my) mother.’ (Stewart 1955: 33)
   b. Kou Miñ Lwiñ-ne Kou Myou Nyuñ-ne
      Ko Min Lwin-with Ko Myo Nyunt-with
      ‘Ko Min Lwin and Ko Myo Nyunt’ (Okell 1966: 190)

For those verb-final or verb-initial WITH-languages which allow or require marking for number on their predicates, an obvious possibility for creating a distinction between comitative and coordination-like constructions lies in a contrast between singular and non-singular verb agreement. This grammaticalization strategy can be encountered in both SOV and VSO languages of the type. Examples are:

(128) Even (Altaic, Tungusic)
   a. Paca ńari-ńun gerka-n
      Paca boy-with walk-3SG.PRS
      ‘Paca walks with the boy.’ (Benzing 1955: 65)
   b. Anna Miko-ńun tul-lā gerka-r
      Anna Miko-and/with street-LOC walk-3PL.PRS
      ‘Anna and Miko walk in the street.’ (Benzing 1955: 65)

(129) Yakut (Altaic, Turkic)
   a. en Ivan-ńun barayn
      2SG Ivan-with go.2SG.PRS
      ‘You are going with Ivan.’ (Krueger 1962: 87)
   b. Liza Sonya-lun oyytutar
      Liza Sonya-and/with fall.down.3PL.PST
      ‘Liza and Sonya fell down.’ (Krueger 1962: 87)
(130) Chemehuevi (Uto-Aztecan)

a. *puŋkuc ohovi-wa tirawiʔi-kwai-vi-∅*
   dog bone-with dash-away-PST-SG
   'The dog ran away with the bone.' (Press 1979: 105)

b. *puŋkuc aŋ pusi-wa nukwi-ji-pim*
   dog that cat-and/with run-PRS-PL
   'The dog and that cat are running.' (Press 1979: 105)

(131) Turkana (Nilo-Saharan, Nilotic)

a. *ā-iryam-ŋ a-yāŋ ka` ñēsi`*
   1SG-meet-ASP 1SG with him
   'I have met him.' (Dimmendaal 1983: 366)

b. *ki-ruk-ir` a-yāŋ ka` e-kapolonŋ*
   1PLL-go-ASP 1SG and/with chief
   'The chief and I went.' (Dimmendaal 1983: 366)

It can be observed, incidentally, that the agreement contrast in languages with patterns (123b) or (124b) is sometimes accompanied by a doubling of the comitative marker in the coordination-like construction. Although, strictly speaking, this double marking is redundant, it stresses the fact that in the plural/dual-agreement construction the structural cohesion between the two NPs is stronger than in the singular-agreement construction. Examples include:

(132) Nubian (Nilo-Saharan, East Sudanic)

a. *ai Ali-dan safār wē-ka fa-safir-ir*
   1SG Ali-with journey one-ACC FUT-travel-1SG
   'I will make a journey with Ali.' (Reinisch 1879: 121)

b. *Esmann-dān Ali-dān tagoran*
   Esman-with Ali-with come.3PL.PF
   'Esman and Ali have come.' (Armbruster 1965: 224)

(133) Cuzco Quechua (Andean)

a. *maisi-kuna-wan mitikarkan*
   comrade-PL-with flee.3SG.PF
   'He fled with his comrades.' (von Tschudi 1884: 374)

b. *ńoka-wan kam-wań wasi-yki-mań risu-ntšix*
   1SG-with 2SG-with house-your-to go-1PL.FUT
   'You and I will go to your house.' (von Tschudi 1884: 467)

(134) Kobon (Papuan, East Highlands)

a. *ne aip ar-nab-in*
   2SG with go-FUT-1SG
   'I will go with you.' (Davies 1981: 73)
Juab aip Minöp aip kale be ar-bil
Juab with Minöp with 3DU forest go-3DU.PF
‘Juab and Minöp have gone to the forest.’ (Davies 1981: 73)

Gradual grammaticalization from a comitative construction is not the only way by which the pure WITH-status of a language can be undermined. In addition, we find cases in which the rise of a coordinate-like structure must have been more “direct”. As we have seen in Section 3, coordinating particles are recruited from a variety of sources, among which a comitative construction is only one of the options. Thus, coordinate NP-structures can arise as the result of the grammaticalization of sentential adverbs, numerals, focus markers, and probably several other construction markers. Now, it is hardly likely that the creation of NP-coordinations from one of these non-comitative sources is limited to AND-languages; there is no reason why such processes should not occur in WITH-languages as well. Thus, a language can lose its “pure” WITH-status not only by way of a gradual grammaticalization of its comitative structure, but also when a coordinate structure is forced upon it “from outside”. An example of a language in which this latter process must have happened very recently is Bahasa Indonesia. As we have seen above, this is a mixed WITH-language, in which the comitative preposition dengan ‘with’ is gradually grammaticalized into an NP-coordinator:

(135) Bahasa Indonesia (Austronesian, West Indonesian)

a. Ali datang dengan ibu-nja
   Ali come with mother-his
   ‘Ali came with his mother.’ (Kähler 1965: 155)

b. Siti dengan Amat
   Siti and Amat
   ‘Siti and Amat’ (MacDonald 1976: 109)

However, the literature indicates that the use of dengan as an coordinating particle between NPs is felt to be old-fashioned and literary (Fokker 1951: 229). Instead of (135b), modern colloquial Bahasa Indonesia has (136). These sentences feature the connective dan, a particle which is also in use as the coordinator of verb phrases and sentences, but which cannot be used as a marker of comitative phrases:

(136) Bahasa Indonesia (Austronesian, West Indonesian)

a. ayah saya dan ayah Dodi mengunjungi pengarang
   father my and father Dodi visit author
terkenal itu
   well-known the
   ‘My father and Dodi’s father visited the well-known author.’
   (Kwee 1965: 12)
b. *Mini menyanyi dan Irma bermain piano*
   Mini sing and Irma play piano
   'Mini sang and Irma played the piano.' (Kwee 1965: 125)

c. *andjing menjalak dan meraung*
   dog bark and cry
   'The dog barked and whined.' (Kähler 1965: 167)

In other words, Bahasa Indonesia started out as a WITH-language in which the comitative construction was grammaticalized into a coordination, but this "coordination"-like comitative was subsequently superseded by a "true" coordinate structure. Thus, in its present-day colloquial form, Bahasa Indonesia is an AND-language.

The development sketched for Bahasa Indonesia can be encountered in some form in other WITH-languages as well. In Vietnamese we see that a grammaticalized comitative competes with an "outside" coordinative structure, the marker of which has its source in an erstwhile sentence adverbial. It is my impression that, in this competition, the comitative-based construction is slowly losing ground, in that it tends to get restricted to combinations which are tightly related by convention and form some sort of conceptual unit.

(137) Vietnamese (Mon-Khmer)

a. *Giáp đi chòi vôi Ăt*
   Giap go stroll with At
   'Giap went out for a walk with At.' (Văn Chinh 1970: 241)

b. *Giáp vôi Ăt là anh em ruột*
   Giap and/or At COP brothers
   'Giap and At are brothers.' (Văn Chinh 1970: 146)

c. *Giáp và Ăt là anh em ruột*
   Giap and At COP brothers
   'Giap and At are brothers.' (Văn Chinh 1970: 146)

d. *Cây này là cây mo và cây kia là cây dào*
   tree this COP tree apricot and tree that COP tree peach
   'This tree is an apricot tree and that tree is a peach tree.' (Văn Chinh 1970: 72)

Essentially the same situation as in Vietnamese is documented for some other areas, such as North-East and Central Asia (Uzbek, Khalkha, Korean) and several American Indian languages (Miskito, Dakota, Mojave). Examples are:
36  **Leon Stassen**

(138) **Uzbek (Altaic, Turkic)**

a. *siz Bilon*
   2PL with 'with you' (von Gabain 1945: 57)

b. *on-am bilan o't-am bərdilar*
   mother-my and/with father-my go.3PL.PST
   'My mother and my father left.' (Sjoberg 1963: 63)

c. *partada perslar wa rüčka bóra*
   desk.LOC pen-points and pen-holder exist
   'There are pen-points and a pen-holder in the desk.' (Sjoberg 1963: 137)

d. *oquvči-lar oqiydilar wa yozalılar*
   student-PL read.3PL.PST and write.3PL.PST
   'The students were reading and writing.' (Sjoberg 1963: 156)

(139) **Korean (Altaic, isolate)**

a. *Yong-i Mia-wa nol-ko isstä*
   Yong-SUBJ Mia-with play-GER be.PRES
   'Yong is playing with Mia.' (Chang 1996: 93)

b. *sasneγ-gwa hagšeq*
   teacher-and pupil
   'a teacher and a pupil' (Lee 1989: 70)

c. *na do ne do ka-gesso*
   1SG and 2SG and go-FUT
   'You and I will go.' (Ramstedt 1939: 114)

d. *na do ka-gesso*
   1SG also go-FUT
   'I will go too.' (Ramstedt 1939: 123)

(140) **Assiniboine (Siouan)**

a. *ne kícι wa-kτί*
   3SG with 1SG-return
   'I returned with him.' (Levin 1964: 116)

b. *wikόška košká kícι Ø-wóetaka-pi*
   girl boy and/with 3-sit.together.to.eat-PL
   'A girl and a boy sat together to eat.' (Levin 1964: 136)

c. *wicáxco ka təwícu wóyate Ø-yutá-pi-šι*
   old.man and wife dinner 3-eat-PL-NEG
   'The old man and his wife did not stay for dinner.' (Levin 1964: 96)
(141) Miskito (Chibchan)

a. *Juan ai lakra wal tawan-ra wa-bia*
   Juan his sister with/and village-to go-3FUT
   ‘Juan will go to the village with his sister/Juan and his sister will go to the village.’ (CIDCA 1985: 214)

b. *Carlos wihki Ignacio wal*
   Carlos and Ignacio with
   ‘with Carlos and Ignacio’ (CIDCA 1985: 196)

In fact, the competition between ‘with’-coordination and ‘and’-coordination manifests itself to a limited extent even in Indo-European. This family typically consists of “pure” AND-languages, but in the earlier forms some degree of ‘with’-coordination for NPs must have been possible, witness the data on Old Irish in (142). In the present, the phenomenon can be attested for some areally related languages from eastern Europe, including Lithuanian, Polish, Russian, and – to a limited degree – Hungarian.

(142) Old Irish (Indo-European, Celtic)

a. *luid co n-a muintir*
   go.3SG.PST with the-his family
   ‘He went with his family.’ (Dottin 1913: 90)

b. *lá co n-óidche*
   day and/with ART-night
   ‘Day and night’ (Dottin 1913: 90)

c. *in fer ocus in ben*
   the man and the woman
   ‘The man and the woman’ (Dottin 1913: 240)

(143) Lithuanian (Indo-European, Baltic)

a. *tēvas iř sūnūs*
   father.NOM.SG and son.NOM.SG
   ‘Father and son’ (Senn 1966: 476)

b. *jūs/jūdu su žmōna*
   2NOM.PL/2NOM.DU with wife.INST.SG
   ‘You and your wife’ (Senn 1966: 477)

(144) Polish (Indo-European, West Slavonic)

a. *Jan i Wanda podróżują po Polscie*
   Jan.NOM and Wanda.NOM travel.3PL.PRES through Poland.OBL
   ‘Jan and Wanda are traveling through Poland.’ (Stone 1980: 15)

b. *syn z siostrą przybyli*
   son.NOM.SG with sister.INST.SG arrive.3PL.PST
   ‘The son and the sister arrived.’ (Meckelein 1926: 53)
(145) Russian (Indo-European, East Slavonic)

a. *moj brat i sestrá studenti*
   
   my brother.NOM.SG and sister.NOM.SG student.NOM.PL
   
   'My brother and sister are students.' (Fennell 1961: 47)

b. *my s nim byl v teátre*
   
   1PL.NOM with 3SG.M.INST be.PST at theatre.LOC
   
   'He and I were at the theatre.' (Raptschinsky 1946: 203)

(146) Hungarian (Uralic, Ugric)

a. *én és a vezetőm elindultunk*
   
   1SG and ART guide.my leave.1PL.PST
   
   'My guide and I left.' (Hetzron 1973: 493)

b. *elindultam a vezetőm-mel*
   
   leave.1SG.PST ART guide.my-with
   
   'I left with my guide.' (Hetzron 1973: 493)

c. *elindultunk a vezetőm-mel*
   
   leave.1PL.PST ART guide.my-with
   
   'I left with my guide/My guide and I left.' (Hetzron 1973: 493)

As a general conclusion, we can state that the relatively low frequency of "pure" WITH-languages is the result of an apparent "drive" in languages towards the creation of a coordinate structure, in which the two NPs are balanced in rank. It is, of course, quite justified to ask why languages from all over the world should exhibit this drive. As a speculation on this point I offer the following considerations. Pure WITH-languages, in which no balanced NP-coordination is available, have the disadvantage of presenting a mismatch between form and meaning. In terms of meaning, the comitative construction must be interpreted as ascribing the same semantic role to the two NPs: the referents of these two NPs "do the same thing". However, this similarity in semantic role is not mirrored by a formal encoding in which the two NPs are of equal structural rank and therefore have the same formal case marking. In other words, pure WITH-languages are not iconic in the sense of Haiman (1980). It can be hypothesized that speakers will, in general, not be happy with such a situation. Hence, they will start to develop an additional coordinative encoding option for NPs, either by differentiating the uses of the comitative construction, or by recruiting an "outside" coordinate structure, or both. The terminal point of this development can be AND-status for the language. From the point of view of iconicity, AND-languages can be rated as optimal. First, they possess a construction which mirrors the equal semantic status of the two NPs. Furthermore, by retaining the comitative construction they formally reflect a functional contrast in the encoding of the domain. Although the two NPs always have the same semantic role, their functional status may be different. It is possible to view the two participants in the event as equally salient or
functionally important, but one may also regard one of the participants as the main protagonist, with the other participant functioning as some kind of "side show". By the fact that AND-languages have both a balanced and an unbalanced encoding strategy for the domain they are able to mirror these functional possibilities through a formal strategy contrast.

At least some part of this iconicity hypothesis receives further support from the facts in various Papuan languages. These are steadfast AND-languages, in that there is a sharp delineation between coordinate and comitative encoding, along the lines of the feature distinctions given in (82). However, even in these languages we see that the comitative construction requires or allows plural/dual agreement on verbs, instead of the singular agreement which is prototypical of comitative encoding. Thus, it seems that languages may choose to honour the semantic fact that the two NPs in the comitative construction have the same semantic role by using dual/plural agreement, regardless of the fact that they already have a coordinate construction, in which this role equality is reflected optimally.

(147) Aghu (Papuan, Central & South)
   a. *nu ko eke ko*
      1SG and 3SG and
      'he and I' (Drabbe 1957: 6)
   b. *efe n'angang gumu fomo-ghenä*
      3SG wife with search.food-3PL.PST
      'He searched food with his wife.' (Drabbe 1957: 42)

(148) Alamblak (Papuan, Sepik)
   a. *yimar-i yënrr yi-më-f*
      man-and boy go-PST-3DU
      'The man and the boy went.' (Bruce 1984: 251)
   b. *yënrr yimar-pnë yi-më-f*
      boy man-with go-PST-3DU
      'The boy went with the man.' (Bruce 1984: 202)

(149) Awtuw (Papuan, Sepik)
   a. *Yowmnôn Yawur du-k-puy-ey*
      Yowmnôn Yawur DUR-IMPF-hit-IMPF
      'Yowmen and Yawur are hitting someone.' (Feldman 1986: 110)
   b. *tey wan-ek t-øy-re*
      3SG.FEM 1SG-with DU-go-FUT
      'She will go with me.' (Feldman 1986: 73)
(150) Usan (Papuan, Madang)
   a. *Munon eng wonou umour nob mor*
      man the his brother with house.in
      *is-orei*
      descend-3SG.PST
      ‘The man entered the house with his brother.’ (Reesink 1987: 85)
   b. *Munon eng wonou umour nob mor*
      man the his brother with house.in
      *is-umirei*
      descend-3PL.PST
      ‘The man entered the house with his brother.’ (Reesink 1987: 85)

(151) Kobon (Papuan, East Highlands)
   a. *ne aip ar-nab-in*
      2SG with go-FUT-1SG
      ‘I will go with you.’ (Davies 1981: 73)
   b. *ne aip ar-nab-ul*
      2SG with go-FUT-1DU
      ‘I will go with you.’ (Davies 1981: 73)

6. The areal distribution of AND-languages and WITH-languages

In this section I will make some observations about the distribution of AND-languages and WITH-languages around the globe. Due to the widespread occurrence of mixed WITH-languages, it is not always possible to state this distribution in a completely clear and uncontroversial fashion. I have committed myself to the following guideline: I will rate a language as a WITH-language if there is minimally a lexical identity of the comitative marker and the coordinate marker. Thus, for example, Literary Bahasa Indonesia will be categorized as a WITH-language, while Modern Colloquial Bahasa Indonesia is an AND-language. Obviously, WITH-status is a gradable property, with “pure” WITH-languages at the end of the spectrum.

Looking first at AND-languages, we note that there are two large areas where this type is concentrated. The first of these “mega-areas” comprises what is called the Old World in Nichols (1992). It includes all the languages of Europe, Central Asia, and Siberia; in the North-East, it also includes some, though not all, of the Austronesian languages of the Philippines. To the South-East, the area stretches as far as India. Finally, it includes the Middle East and Northern Africa. The major language phyla found in the area are Indo-European, Uralic, Altaic, Dravidian, Semitic, and Kartvelian. Moreover, AND-status can be argued for a number of isolate languages in the area, such as Basque, Ket, and
Burushaski. The only cases of doubt are the North-West Caucasian languages and the Dagestanian languages; both Caucasian language groups deviate from the straightforward AND-status of Old World languages to a certain extent.

A second mega-area of AND-languages is formed by (most of the) languages of Australia and New Guinea. Especially along the coast-lines of both islands, where Austronesian influence is notable, there are a number of counter-examples, but the central highlands of New Guinea and the bulk of the Australian mainland contain almost uniform AND-encoding.

Outside of these two main AND-areas, pure AND-encoding can be found in several other places, in particular in the Americas, but the distribution of these AND-languages seems to be rather whimsical. Perhaps the best case for a third AND-area can be made for the languages of the southern part of Central America and the eastern part of South America. This area, which one might call Meso-Andean America, comprises part of the Uto-Aztecan languages of Mexico, the Oto-Mangean languages of Southern Mexico and Honduras, the Chibchan languages of El Salvador, Nicaragua, Costa Rica, and Panama, and the languages of the Andes.

Large, unbroken, concentrations of WITH-languages are encountered in Africa, Asia, and the Americas. With the possible exception of Khoisan, all the languages of Africa in and below the Sahara exhibit (some degree of) WITH-encoding. A second WITH-area is made up of East and South-East Asia (including Tibet, China, Korea, and Japan) and the islands of Indonesia, Melanesia, and Polynesia. On the American continent, WITH-status is normal for at least the Far North (Alaska, Canada), and the Deep South (non-Andean South America).

As we have seen, grammaticalization of comitatives into coordinate-like structures is a process which may have proceeded further in some languages than in others. It can be predicted that the more advanced stages of this process will be encountered in the border areas between the various heartlands of the two types, and this is roughly what we find in the data. Mixed WITH-languages are particularly prominent in Ethiopia and Sudan, and in the African Sahel territory which separates North Africa from Sub-Saharan Africa. Furthermore, the North-East of Asia, where the Old World AND-area meets with the North-American and East-Asian WITH-areas, shows quite a large number of “almost” AND-languages. A third conspicuous case in point is formed by the North Coast of New Guinea, on the interface of the New Guinean AND-area and the Austronesian WITH-area.

Some special attention must be paid to the situation in North America. For many of these languages a categorization on the AND/WITH parameter is problematic. At first sight, one might want to rate the majority of these cases as AND-languages, as there seems to be a clear structural differentiation between coordinate and comitative strategies. Closer inspection, however, reveals that
both of these strategies manifest themselves formally in a way which is definitely deviant from their prototypical encoding as defined in (83). The Comitative Strategy in languages of North America very often involves the use of specific ‘accompany’-verbs, or the formation of applicative verbs, by which the argument structure of the verb is altered. Although neither of these formal variants of the Comitative Strategy is limited to occurring in North America only, the density of these variants in this area is certainly conspicuous. Some examples are:

(152) Biloxi (Siouan)
\[n-y-i \text{ˈno} \text{″pa} \quad n-da \quad dande]\n1SG-2SG-be.two 1SG-go FUT
‘I will go with you.’ (Dorsey & Swanton 1912: 238)

(153) Blackfoot (Algonquian)
\[omistsiks pokaiks ixp-\text{ˈit}-saks-iu\]
this.PL child.PL with-there-go.out-3SG.AN
‘He then went out with those children.’ (Uhlenbeck 1938: 87, 206)

(154) Ojibwa (Algonquian)
\[wi\text{ˈt-o} \text{ˈppu-ma} \text{ˈt}\]
with-eat-3SG.ACT/3SG.PAT
‘He eats with him.’ (Bloomfield 1958: 95)

(155) Quileute (Chimakuan)
\[ki\text{ˈtax} \quad yik \quad á\text{ˈt} \text{ˈcit.t} \quad hé\text{.ol-ic} \quad kaki\text{″}\]
go.PROG ART chief’s wife accompany-INGR ART
dsisters
tsitsi’itskwa’”a
‘The chief’s wife went out with her daughters.’ (Andrade 1938: 279)

For those North American constructions which might be considered as the functional equivalent of coordinations, deviance from the norm formulated in (83) can be noticed as well. In the typical case, the two NPs in such constructions do not form a constituent: they are not (or do not have to be) contiguous, and verbs typically have singular forms, agreeing only with one of the NPs. Often there is no overt marker to signal coordination of NPs. If there is, the item is usually a “loose” particle with the meaning ‘also/then/again’, which is not in construction with any of the two NPs. Furthermore, we find various languages in which the two NPs cannot even be part of the same clause. Biloxi and Yavapai, among others, are languages which show a preference for an encoding in which either NP is constructed as a constituent of its own clause, which contains an occurrence of the same predicate. The equality in rank between the two NPs is signalled by the fact that the two NPs fulfill the same case role in their respective clauses.
AND-languages and WITH-languages

(156) Biloxi (Siouan)

aye 'ki ūṅk-tcu'-di ato poǐtcka ūṅk-tcu'-di
corn 1SG-plant-PST potatoes 1SG-plant-PST
‘I planted corn and potatoes.’ (Dorsey & Swanton 1912: 22)

(157) Yavapai (Yuman)

maː-c m-yuʔa-cʔ -yu-e: hlo ?-maː-c-km
2-SBJ 2-be 1-SBJ 1-be-DS rabbit 1-eat-PL-lMPF
‘You and I were eating rabbit.’ (Kendall 1976: 194)

For Yavapai, Kendall explicitly states: “It looks very much as if Yavapai has no way to conjoin [...] NPs” (1976: 199). In other words, the functional range of NP-coordinations in other languages is often covered by some form of sentential coordination in North American languages. In Biloxi and Yavapai, this sentential coordination is still completely transparent; that is, it is not grammaticalized at all. In other languages, some incipient stage of grammaticalization may be at hand, in that one of the predicates has been left out. However, in such cases the singular verb agreement, the non-contiguity of the two NPs, and the “loose” use of sentential adverbs as coordinative markers point to the basic sentential status of the construction. The conclusion can be defended that such languages do not have the option of coordinative NP-encoding, at least not if we hold on to the definition of the domain as formulated in Section 2. This, then, would have as its consequence that, by definition, these languages are to be classified as WITH-languages. However, since comitative encoding in North American languages is non-standard and puzzling in itself, this consequence should not be drawn too easily. In particular for the languages of North America the old cliché holds that further research will be definitely needed.

7. Correlational parameters of the AND/WITH-distinction

Now that the typology of the encoding of our domain has been established, it is time to look for possible correlational parameters. A good way to start this search is to consider typological studies in which the areal distribution of the parameter(s) described more or less matches the areal distribution of the AND/WITH-parameter. Thus far, I have been able to find two cases which seem promising in this respect.

Nichols (1992) contains a description of the geographical spread of Head-Marking vs. Dependent-Marking of various grammatical categories. The map which she draws matches the AND/WITH-map in several interesting features. In particular, the delineation of the AND-areas in the Old World and Australia-New Guinea, as well as the WITH-area of South-East Asia and Austronesia, find their counterparts in the map of the Head/Dependent parameter. Of course, the match of the two maps is far from perfect, but there is enough of a correspondence to justify closer examination. Up to now, I have investigated the corre-
lation of AND/WITH with just one of the manifestation of the Head/Dependent parameter, namely, the presence or absence of case marking on argument NPs. For this I have employed a distinction between “cased” and “non-cased” languages, which is defined as follows:

(158) A language is cased if the difference between core argument NPs is obligatorily marked by way of bound morphology on these argument NPs.

The Casedness Parameter is, as is usual with parameters, gradual in itself. To name just one hierarchical factor, one often finds that the marking on nouns has been lost, while it still lives on in the pronominal system: English is an obvious case in point. So far I have found it useful to take the least restricted interpretation of the definition as my guideline. Hence, even languages like English, which have a lexical limitation on case marking in that only a select group of words exhibit it, are reckoned to be cased.

When we correlate the Casedness Parameter with the AND/WITH Parameter, we ask ourselves whether or not the four logically possible combinations of the values for these parameters will have roughly similar frequencies, and we hope that they do not. The results of the investigation provide some support for our expectations. It turns out that the combination of AND and Cased, and the WITH–NonCased combination, are highly frequent. As for the other two options, the AND–NonCased combination has some frequency, but it is definitely minor. The WITH–Cased option hardly ever occurs at all. On the basis of these results, it is possible to formulate the following correlational tendencies:

(159) Tendencies in the Casedness–AND/WITH correlation:
  a. If a language is Cased, it will tend to have AND-status.
  b. If a language has WITH-status, it will tend to be NonCased.

Thus, Cased is a good indicator for AND-status, while WITH is a good indicator for NonCased. The fact which makes the statements in (159) tendencies rather than universals is that a language may lack Case but still can have AND-status. As I stated in the introduction, no attempt at an explanation of the correlations in (159) will be presented here. However, as a possible suggestion, I want to venture the thought that the reason why these two parameters appear to cluster lies in a general typological decision which any natural language has to make. Both parameters concern the question of whether or not a language must formally acknowledge the similarity and difference in semantic role for two (nuclear) NPs. Seen from this perspective, it can be explained why AND–Cased and WITH–NonCased are the two major combinations in the cluster. Pure WITH-languages do not formally recognize the semantic role identity
of the two NPs in the conjunctional construction. Likewise, they do not formally mark the difference in semantic role between Agent and Patient NPs in transitive sentences. **AND-languages** are iconic, in that they match the identical semantic role status of NPs in conjunctional constructions by the equal formal encoding of these NPs. In transitive sentences, **AND-languages** are iconic as well, since they reflect the contrast between Agent and Patient in their formal system.

A second parameter which seems to provide a correlation for the **AND/WITH** parameter is Tensedness. This parameter, which has been developed in Wetzler (1996) and Stassen (1997), presents a distinction between tensed and non-tensed languages. It can be defined as follows:

(160) A language is tensed if

a. it has an obligatory marking for the distinction in Past versus Non-Past, and

b. this marking is effectuated by bound morphology on verbs.

Thus, tensed languages are those languages which, in the terms of Comrie (1985), have a grammatical category of tense. It will be noted that definition (160) contains both a semantic and a formal aspect of the notion. The requirement of a Past–NonPast distinction rules out those languages which have no tense distinctions at all (such as Cambodian), or only have a Future–NonFuture distinction (such as Burmese), or only have aspectual distinctions (such as Choctaw). The requirement of bound morphological marking on verbs rules out a language like Tigak. This language has obligatory marking for Past–NonPast, but this marking takes place by means of so-called “subject–tense pronouns”, which are not morphologically fused with the verb. In contrast to all these non-tensed languages, English is a clear example of a tensed language.

(161) Cambodian (Austro-Asiatic, Mon-Khmer)

\[\text{kee maok Phnom-Penh}\]

3PL come Phnom-Penh

‘They come/came/will come to Phnom-Penh.’ (Huffman 1967: 228)

(162) Burmese (Sino-Tibetan, Tibeto-Burman)

a. \[\text{eiñ pyañ\text{-thwà-te}}\]

home return-go-\text{l.csc NONFUT}

‘(He) goes/went home.’ (Okell 1969: 389)

b. \[\text{li'-me}\]

vanish-FUT

‘(I) will vanish.’ (Okell 1969: 467)
(163) Choctaw (Muskogean)
   a. *pisa-li*
      look.at.PF-1SG.ACT
      'I see/saw it.' (Nicklas 1974: 75)
   b. *pinsa-li*
      look.at.IMPF-1SG.ACT
      'I am/was looking at it.' (Nicklas 1974: 75)

(164) Tigak (Austronesian, Melanesian)
   a. *gi ima*
      3SG.PRS come
      'He is coming.' (Beaumont 1980: 74)
   b. *ga ima*
      3SG.PST come
      'He came.' (Beaumont 1980: 74)

(165) English (Indo-European, West Germanic)
   a. *John laugh-s*
   b. *John laugh-ed*

   The geography of the Tensedness Parameter is remarkably similar to the ge-
   ography of the AND/WITH Parameter, at least in as far as the "core areas" are
   concerned. The mega-areas of AND-distribution, viz. the Old World, Australia-
   New Guinea, and the Meso-Andean area, are also the most prominent con-
   centrations of Tensed languages. Conversely, South-East Asia, Austronesia,
   Sub-Saharan Africa, and the majority of American languages are clearly non-
   tensed, and it is in these areas that WITH-encoding is concentrated. As with
   the Casedness Parameter, there thus seems to be a preference for two diametri-
   cally opposed combinations on the Tensedness Parameter and the AND/WITH
   parameter. We can formulate the following two tendencies:

(166) Tendencies in the Tensedness–AND/WITH correlation:
   a. If a language is Tensed, it will tend to have AND-status.
   b. If a language has WITH-status, it will tend to be nonTensed.

   Again, what keeps these tendencies from being bi-directional is the fact that
   there is a minor option for a combination of NonTensedness and AND-status.
   Thus, it is not impossible for a non-tensed language to have AND-encoding. In
   opposition, a combination of Tensedness and WITH-status hardly ever occurs
   in my sample.

   The Casedness Parameter and the Tensedness Parameter thus present the
   same picture in their correlation with the AND/WITH Parameter. For both pa-
   rameters, a clear matching of values with one of the values on the AND/WITH
   Parameter can be established. But in both cases the picture is somewhat blurred
by the fact that AND-encoding turns out to be more widespread than would be predictable from the correlations. However, this fact can be explained on the basis of the observations made in Section 4. We have seen there that languages have a special motivation for the development of balanced conjunctival constructions. Since the “drive” behind this development is probably not present in the realms of Casedness and Tensedness, it is to be expected that AND-status will, in a minority of languages, be acquired by languages which would not be entitled to it on the grounds of Casedness or Tensedness.

In conclusion, we can state that the hunt for correlational parameters of the AND/WITH-distinction has turned out to be an at least potentially fruitful enterprise. The AND/WITH-distinction can be shown to cluster with two other typological distinctions which, at least at first sight, seem to be independent of it. It would, however, be wildly premature to formulate any far-reaching explanatory principles for this clustering. For one thing, it is absolutely conceivable, and even likely, that the three parameters reviewed here are not the only members of the cluster. Furthermore, it is far from clear what the mutual relations between the three parameters in the cluster are. For example, the relation between Casedness and Tensedness, if indeed there is any, has not been clarified at all. Hence, much additional descriptive work will be required before questions of this nature can even begin to be formulated in a sensible fashion.

Received: 26 February 1999
Revised: 11 February 2000

Appendix

The sample used in this study includes the following languages: Abkhaz, Acehnese, Acholi, Aghu, Agul, AINU, Akan, Alabama, Alamblak, Albanian, Amarakaeri, Amele, Amharic, Andoke, Apalai, Cairene Arabic, Classical Arabic, Archi, Classical Armenian, Arrernte, Avar, Awtuw, Babungo, Bahasa Indonesia, Bambara, Banda, Banggai, Barea, Bari, Basque, Karo Batak, Beja, Bilin, Biloxi, Birom, Blackfoot, Bororo, Breton, Bribri, Buli, Burmese, Burushaski, Cambodian, Canela-Krahó, Car, Surinam Carib, Cebuano, Chacobo, Chamorro, Chatino, Chechen, Chemehuevi, Choctaw, Highland Chontal, Chukchi, Coos, Coptic, Cornish, Dafla, Daga, Dakota/Assiniboine, Dinka, Diyari, Duala, Dumaki, Dutch, Dyirbal, English, Even, Ewe, Falor, Fijian, Finnish, Fulani, Gaelic, Gallu, Gallilarese, Modern Georgian, Gidabal, Goaiiro, Gothic, Grebo, Ancient Greek, Modern Greek, Guajajara, Guanano, Guaraní, Gumbainggir, Gurene, Haida, Haitian Creole, Hausa, Biblical Hebrew, Modern Hebrew, Hindi, Hixkaryana, Hungarian, Icelandic, Igbo, Inuktitut, Modern Irish, Old Irish, Jabem, Jacaltec, Japanese, Kabyl, Kaingang, Kalispel, Spoken Kannada, Kanuri, Kapauku-Ekagi, Kète, Keiese, Ket,

Notes

Correspondence address: ATD/KUN, Erasmusplein 1, 6525 GG Nijmegen, The Netherlands; e-mail: l.stassen@let.kun.nl

This paper is dedicated to Ekkehard König, on the occasion of his 60th birthday.

Part of the research reported here was conducted within the context of the EUROTYP Programme, which was sponsored by the European Science Foundation (ESF) in 1990–1994. I am grateful to the members of the Noun Phrase Group (Grev Corbett, David Gil, Jim Hurford, Sasha Kibrik, Masha Koptjevskaja-Tamm, Edith Moravcsik, John Payne, Frans Plank), as well as to Martin Haspelmath, for their stimulation and support. I am also indebted to my graduate students René van Gessel and Michael Cysouw for their help in gathering and discussing the data, and to Wolfgang Schellinger for crosschecking some of them.

Abbreviations: 1, 2, 3 1st, 2nd, 3rd person, I, II, ... noun class marker, ABS absolutive case, ACC accusative case, ACT actor marker, AN animate gender, AOR aorist tense/aspect, ART article, ASP aspect marker, CL class, COM comitative case, COP copula, DAT dative case, DEF definiteness marker, DEM demonstrative, DS different-subject marker, DU dual number, DUR durative aspect, EMPH emphasis marker, ENUM enumerative marker, F feminine gender, FOC focus marker, FUT future tense, GEN genitive case, GER gerundial marker, IMP imperative mood, IMPF imperfective aspect, INCL inclusive marker, INGR ingressive aspect, INST instrumental case, LOC locative case, M masculine gender, NARR narrative form, NEG negation marker, NMNL nominalizer, NOM nominative case, NONFUT non-future
tense, OBL oblique case, PAT patient marker, PE exclusive plural, PF perfective aspect, PI inclusive plural, PL plural, PROG progressive aspect, PRS present tense, PST past tense, RECIP reciprocal marker, RFL reflexive marker, SBJ subject marker, SG singular, SS same-subject marker, STAT stative form, TOP topic marker.

1. All examples which are not my own data are quoted in the original orthography. In some cases, however, I have made some changes or alterations to the material presented in the source. On various occasions I have left out parts of the quoted sentence, as I deemed those to be immaterial to the discussion at hand. Glosses have, in a number of cases, been altered in order to streamline the set of abbreviations used in this paper. Finally, in a few cases I have changed the original translation somewhat, or added some clarification of my own. Given all this, I must ask the reader not to quote linguistic material from this paper; instead, one should always consult the original source.

2. As will become clear from the data presented further on, I have tried to restrict myself to cases in which the participants are formalized as NPs with a full nominal head. Thus, cases in which one or both of the participants are represented by a pronoun will be avoided as much as possible. The reason for this is that several languages have a special conjunctional strategy when pronouns are involved. Some examples are:

(i) Dinka (Nilo-Saharan, Nilotic)
   a. wa ke ma
      my.father and my.mother
      ‘my father and my mother’ (Nebel 1948: 93)
   b. Ok o yin
      1PL 1PL 2SG
      ‘you and I’ (Nebel 1948: 93)

(ii) Kalispei (Salish)
   a. in-lo'eu u i-skúí
      my.father and my.mother
      ‘my father and my mother’ (Vogt 1940: 171)
   b. l-qe'-Alík
      with-1PL-Alík
      ‘Alik and I’ (Vogt 1940: 66)

(iii) Kpelle (Niger-Kordofanian, Mande)
   a. surɔŋ 'tɔ nɛnɛ 'tì pà
      man 3PL.INCL woman 3PL come
      ‘A man and a woman came.’ (Welmers 1973: 306)
   b. kwà Sumo kà pà
      1PI Sumo 1PL come
      ‘Sumo and I came.’ (Welmers 1973: 306)

(iv) Khalkha (Altaic, Mongolian)
    ter bid choër
    3SG 1PL two
    ‘he and I, lit. he we two’ (Street 1963: 181)
This special treatment of pronouns is definitely an interesting typological phenomenon, which, in a full exposition of NP-conjunction, should certainly have its place. For the present, however, I have decided to ignore it. Whenever pronouns are featured in the example sentences, it should be understood that, in the language in question, the behaviour of pronouns in coordination is identical to that of full nominal phrases.

3. The existence of WITH-languages is hardly a new finding; it was already established by an author as early as Trombetti (1923). I am grateful to Thomas Stolz for pointing this out to me.

4. Grammaticalization of comitatives into coordinations has been argued for in Mithun (1988), and in various later publications, such as Heine, Claudi, & Hünnefelder (1991) and Stolz (1998).

5. For a detailed exposition of the syntactic status of various coordination-like constructions in Japanese see Ono (1993).

6. Plural agreement in a comitative encoding is only possible in Hungarian if at least one of the participants in the construction is pronominal. Thus, while sentence (146c) is grammatical, sentence (i) is not:

(i) Hungarian (Uralic, Ugric)

*A Őiři elindultak a vezetőjével
the man leave.3PL.PST the guide.his.COM

'The man left with his guide.' (Hetzron 1973: 493)

References


AND-languages and WITH-languages


Landaburu, Jon (1979). La langue des Andoke. (Langues et Civilisations à Tradition Orale, 36.)


