

Elements of complex structures, where recursion isn't

Masayoshi Shibatani
Rice University

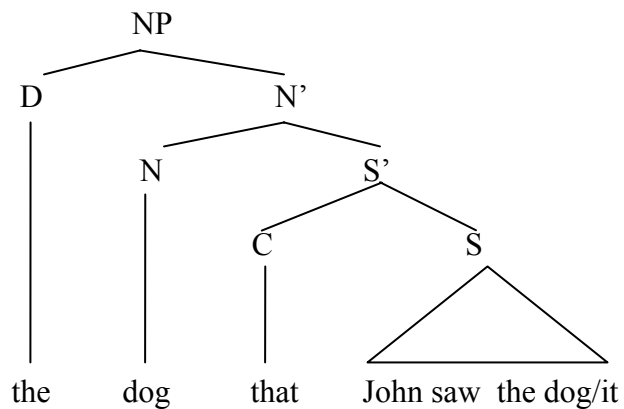
1. Chomskyan notation of recursion and syntactic structures

As far as I can see, there are two negative consequences of the traditional Chomskyan notation of recursion of the following form or its variants accounting for complex syntactic structures.

- (1) a. $S \rightarrow NP VP$
- b. $NP \rightarrow (D)N' (S')$
- c. $VP \rightarrow V (S')$
- d. $S' \rightarrow (C) S$

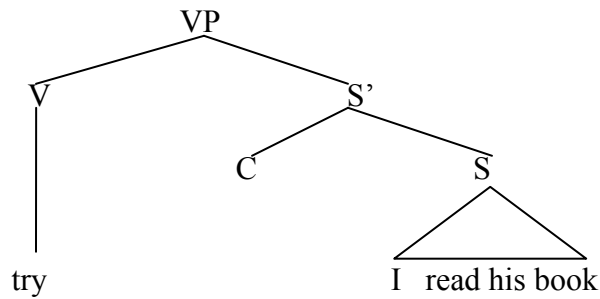
One is the implication that what is embedded within a NP or a VP is the same object as the main clause, namely a **sentence**. The other, related assumption is that a full sentence **with all its arguments** underlies a clausal complement and a modification structure such as a relative clause. These have had a profound effect in our thinking about the nature of complex syntactic structures and their analysis in both synchronic and diachronic dimensions. Synchronically, we have been taught that relative clauses and verb complement constructions, for example, have the following underlying structures.

(2) a. Relative clause



(the dog that [John saw])

b. Verb complement



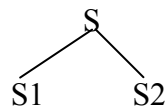
(I try [to read his books])

Many interesting research results have been achieved based on the assumptions embodied in the Chomskyan recursive rules and the syntactic structures they countenance. For example, Keenan and Comrie's (1977) seminal work on the universal constraints on relativization is predicated on the assumption that underlying relative clause structures have the form of (1b) above. Such an assumption allows them to talk about "the NP in the restricting **sentence** that is coreferential with the head noun as the NP relativized on" (64; emphasis added), and to establish the concept of accessibility to relativization in terms of which the proposed universals on relativization are formulated.

The other, diachronic arena is also fraught with ideas that a sentence might be integrated with another one giving rise to such complex structures as relative clauses, clausal complements, subordinate adjuncts, and serial verb constructions. For example, Hopper and Traugott (1993:169) note that "[f]rom the point of view of language change, the initial formation of a complex structure involves the combining into one integrated structure of two separate and autonomous nuclei [sentences] that are mutually relevant" as depicted in (3) and (4) below:

(3) $S1 \Leftrightarrow S2$

(4)



Whereas Hopper and Traugott (1993) are a bit more careful in their description of the transition from the paratactic to other more integrating patterns of clause combination along the cline of parataxis > hypotaxis > subordination, others have been less so. For example, Heine and Kuteva (2007: 214) have recently suggested two channels through which clause subordination arises, namely "[e]ither Expansion, that is, the reinterpretation of a nominal as a clausal (propositional) participant, or via the integration of two independent **sentences** within one sentence" (emphasis added), and have sketched the two patterns of development in the following manner:

- (5) a. S [NP] > S₁ [S₂] Expansion (complements and adverbial clauses)
 b. [S₁ + S₂] > S₁ [S₂] Integration (relative clauses)

While it is true that a sentence may consist of two or more sentences, as at the paratactic stage shown in (4), where the two subparts are not structurally integrated, the transition from a paratactic structure to a more integrating hypotactic and subordinate one seems to require greater cognitive processing, not shown in (5), than the formal hierarchical integration of two sentences into a complex structure. This is easy to see if we look at what is involved in converting a direct quote into an indirect one in a language, such as (Old) Japanese, which has numerous predication features, e.g. discourse particles, honorifics, tense-marking, distinguishing between a sentence and a(n embedded) clause. Indeed, the essence of clause integration seems to be the cognitive ability to convert a sentence into a non-sentential object, which can then be legitimately integrated into a matrix sentence. In what follows, I would like to show this largely on the basis of relative clause formation in Austronesian languages and others, where a full clause, let alone a sentence, is not involved, contrary to what is suggested by the syntactic structure shown in (2a) countenanced by the recursive phrase structure rules, or the schematic representation such as (5b) (see Deutscher's symposium paper for a very similar view). The discussion below is basically synchronic but diachronic studies such as Heine and Kuteva (2007) and others are only as good as their match with the synchronic patterning of language.

2. Austronesian relativization

Despite the reduction in the morphological contrast and even a total loss of such contrast in some dialects, Sasak of Lombok Island in eastern Indonesia is typical of Western Malayo-Polynesian languages in maintaining the structural contrast between so-called Actor-focus (AF) constructions and Patient-focus (PF) constructions. This is clear from the way relativization works in these dialects. That is, consistent with other relevant Austronesian languages, only the primary argument, referred to variably in the literature as "topic", "subject", "pivot" or "trigger", can be relativized on—the fact that has been construed as the "subjects-only" constraint by Keenan and Comrie (1977) and that underlies one of the proposed universal constraints on relativization formulated as "subjects are universally the most relativizable of NPs" (Keenan 1985:158) or "every language can relativize on subjects" (Comrie 1989: 158).

(6) Pancor ngeno-ngené Sasak

- a. **dengan mame ino** mantok loq Ali (AF)
 person male that N.hit ART Ali
 'That man hit Ali.'
- a'. dengan mame [si Ø mantok loq Ali] batur=meq
 person male REL N.hit ART Ali friend=2SG
 'That man who hit Ali is your friend.' (Topic A relativized)
- a". *Loq Ali [si dengan mame ino mantok Ø] batur=meq
 ART Ali REL person male that N.hit friend=2SG
 'Ali, whom that man hit, is your friend.' (Non-Topic P relativized)
- b. **Loq Ali** pantok=na siq dengan mame ino (PF)

ART Ali Ø.hit=3SG by person male

‘That man hit Ali.’

b’. loq Ali [si Ø pantok=na siq dengan mame ino] batur=meq

ART Ali REL hit=3SG by person male that friend=2SG

‘Ali, whom that man hit, is your friend.’ (Topic P relativized)

b’’. *dengan mame [si Ali pantok=na Ø] batur=meq

person male REL Ali hit=3SG friend=2SG

‘The man who hit Ali is your friend.’ (Non-Topic A relativized)

The gaps in the relative clauses above indicate the position of relativization in the Keenan-Comrie type approach that assumes a full sentence status for relative clauses. Comrie and Horie (1995) and Comrie (1998), on the basis of the parallelism between relative clauses with gaps and ordinary sentences with similar gaps in Japanese seen below, argue that relative clauses (in Japanese) are no different from ordinary sentences with pronominal gaps, and that RC gaps are not the ones created by extraction or movement of the relativized NPs as in the standard generative analysis.

(7) Japanese

a. kore=ga [Ø kinoo Ø katta] hon desu.

this=NOM yesterday bought book COP

‘This is the book that (I) bought Ø yesterday.’

b. Ø kinoo Ø katta.

yesterday bought

‘(I) bought (it) yesterday.’

While (7b) is a perfect independent sentence of Japanese that answers a question (such as “Have you bought the book already?”), the parallelism between relative clauses and independent sentences seen above is deceptive, and a similar situation may not obtain in other languages. For example, to the question in (8a) below, the appropriate answer in Pancor ngeno-ngené Sasak would be either (8b) with full pronouns or (8c) with pronominal clitics, while the Japanese answer would have gaps for “I” and “it”.

(8) Pancor ngeno-ngené Sasak

a. Kumbeq=meq buku=no?

what.do=2 book=that

‘What did you do with that book.’

b. **Aku** nulak-ang **ia** tipak perpustakaan

I return-APPL it to library

‘I returned it to the library.’

c. **Ku**=nulak-ang-**e** tipak perpustakaan

1=return-APPL-3 to library

‘I returned it to the library.’

The corresponding relative clause, however, cannot contain the object clitic, indicating that the relative gap here is an obligatory gap.

(9) Pancor ngeno-ngené

Buku [si ku=tulak-ang-*e/Ø tipak perpustakaan]=no bagus
book REL 1-return-APPL-3 to library=that interesting
'The book that I returned to the library was interesting.'

Even in Japanese, the two gaps seen in the relative clause in (7a) differ in that while the first gap corresponding to the subject nominal can be overtly expressed, the second one corresponding to the object nominal coreferential with the head cannot in any form. This shows that the parallelism Comrie draws between independent sentences and relative clauses does not in fact obtain and that RC gaps are different from those created under discourse conditions.

Just as identifying relative clauses as ordinary sentences is mistaken, labeling markers such as *si* in Pancor ngeno-ngené Sasak and its equivalents in many other languages as relativizers or relative clause markers as we have done above is also misleading. Expressions headed by *si* and its equivalents (*siq*, *saq*, *siq-saq*) in the dialects of Sasak occur in a wide range of modification functions, ranging from both nominal and verbal complements and subordinate adverbial expressions, as seen below:

(10) Pancor ngeno-ngené Sasak

- a. buku [si Ø ne=bace isiq loq Ali]
book NMZ 3=Ø.read by ART Ali
'the book that Ali read'
- b. suara [si ne=ngerontok lawang loq Ali]
sound NMZ 3=knock door by Ali
'sound of Ali knocking on the door'
- c. berita [si angku=n loq Ali beruq merariq]
news NMZ way=3 ART Ali recently marry
'the news that Ali recently got married'
- d. Aku lupa [si angku-n loq Ali wah mbilin kota=no]
I forget NMZ way=3 ART Ali PERF leave town=that
'I forgot that Ali had left the town.'
- e. waktu [si ku=masih sekolah]=no...
time NMZ 1=still school=that
'At the time when I was still going to school.../When I was going to school...'
- f. Ali te-semateq [si=ne lekaq léq rurung]
Ali PASS-kill NMZ=3 walk in street
'Ali was killed when/while he was walking in the street.'

A more appropriate label for *si* in Pancor ngeno-ngené Sasak is "nominalization marker". Then our understanding of Sasak relativization takes a different turn, namely relativization in Sasak—and Austronesian languages and many others, as we shall see below—represents one of the nominal modification functions of nominalized clauses rather than a phenomenon involving full clauses or full sentences subordinated to the head nominal, as in the traditional generative and the typological analysis. Indeed, *si* and its equivalents in other Indonesian languages (such as *yang* in Bahasa Indonesia) mark

nominalizations of even such items as demonstratives and ordinal numbers in the same manner as it marks nominalized clauses, as shown by the nominalization marker *saq* in Puyung meno-mené Sasak below:

(11) Puyung meno-mené Sasak

- a. [saq ino] baru
NMZ that new
'That one is new.'
- b. [saq pertame] mame kance [saq kedue] nine (speaking about one's children)
NMZ first male and NMZ second female
'The first one is male and the second one is female.'
- c. Gitaq [saq Ø nyenke=n tokol leq bucu=no]
look NMZ PROG=3SG sit LOC corner=the
'Look at the one sitting in the corner.'
- d. [saq Ø nyenke=n tokol leq bucu=no] amaq=k
NMZ PROG=3SG sit LOC corner=the father=1SG
'The one sitting in the corner is my father.'
- e. Amaq=k [saq Ø nyenke=n tokol leq bucu=no]
father=1 NMZ PROG=3SG sit LOC corner=the
'My father is the one sitting in the corner/It is my father who is sitting in the corner.'
- f. Kamu amaq=k
you father-1SG
'You are my farther.'

The above examples also present clear evidence that *si*, *saq* and others indeed head nominalized expressions functioning as nominal arguments or nominal predicates in the same manner as simple nouns do as in (f). Constructions such as cleft constructions and wh-questions also involve nominalized expressions in Sasak and other languages. Compare (e) above, which is analogous to the *it*-cleft in English, as in one of the translations for it, and the wh-questions below:

(12) Puyung meno-mené Sasak

- a. Sai [saq Ø bace buku=no]
who NMZ read book=the
'Who read the book?' Lit. 'Who is the one that read the book?'
- b. Ape [saq mu-m bace Ø]
what NMZ PAST=2SG read
'What did you read?' Lit. 'What is the one that you read?'

So-called cleft sentences and wh-questions in Sasak (and many other languages) are nothing but equational sentences consisting of a simple noun phrase or a wh-pronoun and a nominalized clause as a nominal predicate, just as basic equational sentences consist of

two nominals, one functioning as a nominal topic and the other a nominal predicate, as seen below:

(13) Puyung meno-mené Sasak

- a. Amaq=k [dedgan mame ino]
friend=1SG person male that
'My father is that man.'
- b. Sai [dengan mame ino]
who person male that
'Who is that man?'

Comparison between (11e) and (13a), on the one hand, and those in (12) and (13b), on the other, indicates that the gaps in (11e) and (12) are not created by the movement of an NP into an initial position, as an extraction analysis of these constructions have it—there is no gap in the predicate nominals in (13) showing extraction of an argument. Relative clauses, wh-questions and clefts all involve a nominalized clause with a gap, and it is this feature that underlies the unity among these three construction types widely recognized.

To summarize, relative clauses in Austronesian languages is an appositive construction consisting of a nominal (corresponding to a head nominal) and a nominalized clause with a gap (corresponding to a relative clause) juxtaposed as in (14) below. As an appositive construction, the head and the relative clause of this type of RC construction stand in a discourse anaphora relation leading to the entailment relation between expressions like (15a) and (15b).

(14) Pancor ngeno-ngené Sasak

buku [si beng=ku iye Ø]=no
book NMZ give=1 he =that
'the book that I gave him'

- (15) a. Loq Ali mbace buku [si beng=ku iye Ø]=no
ART Ali N.read book NMZ give=1 he =that
'Ali read the book that I gave him.'
- b. Loq Ali mbace [si beng=ku iye Ø]=no
ART Ali N.read NMZ give=1 he =that
'Ali read the one I gave him.'

Cf.

- a. I saw John the butcher in the market.
- b. I saw the butcher in the market.

3. Austronesian nominalizations

The reason that I prefer using the term “nominalization marker” above to the more conventional one of “nominalizer” for Sasak particles such as *si* and *saq* that head nominalized clauses is that Austronesian nominalizations in general take place without any nominalizing morphology, despite the fact a number of languages, e.g. Formosan languages Rukai, Saisiyat, etc., have developed nominalization morphology. For example, Formosan language Mayrinax Atayal nominalizes a sentence without any nominalizing morphology or even a marker.

(16) Mayrinax Atayal (based on Huang 2002)

- a. **yakaat m<in>uwah cuʔhisaʔ** kuʔ naʔakis
 NEG AF<PERF>come yesterday NOM.REF old.man
 ‘The old man didn’t come yesterday.’
- b. **kiaʔ ʔiʔ m-aniq kuʔ [yakaat m<in>uwah cuʔhisaʔ]**
 PROG LIN AF-eat NOM.REF NEG AF<PERF>come yesterday
 ‘The one who didn’t come yesterday is eating (there).’
- c. **kiaʔ ʔiʔ m-aniq kuʔ cuqliq kaʔ [yakaat m<in>uwah cuʔhisaʔ]**
 PROG LIN AF-eat NOM.REF person LIN NEG AF<PERF>come yesterday
 ‘The person who didn’t come yesterday is eating (there).’

As the above examples show, a nominalized clause without any nominalization morphology or a marker in Mayrinax Atayal functions both as an argument marked by the nominative particle (16b) and as a relative clause linked to the head nominal (16c). The parallel pattern obtains in Tagalog, as shown below:

(17) Tagalog (courtesy of Naonori Nagaya)

- a. **Hindi d<um>ating** ang matanda-ng lalaki **kahapon**
 NEG come<AF> TOP old-LIN man yesterday
 ‘The old man didn’t come yesterday.’
- b. **K<um>a-kain doon ang [hindi d<um>ating kahapon]**
 DUP<AF>-eat there TOP NEG come<AF> yesterday
 ‘The one who didn’t come yesterday is eating there.’
- c. **K<um>a-kain doon ang tao-ng [hindi d<um>ating kahapon]**
 DUP<AF>-eat there TOP person-LIN NEG come<AF> yesterday
 ‘The person who didn’t come yesterday is eating there.’

The nominalization markers in Sasak dialects, which seem to be a later development, mark what has been nominalized as such, much like Chinese *de* and Japanese *no*. Such nominalization markers may not occur in certain contexts, as in the case of Chinese *de* and Japanese *no*, or may be optional, as in Sasak *si*, *siq*, etc.

3.1 The role of the Austronesian focus morphology in argument nominalizations

While the Austronesian focus morphology may occur in both action nominalization and argument nominalization, it plays a very important role in profiling the grammatical role of the argument nominalized in the latter process, and it is likely that this is the original role of the focus morphology in Proto Austronesian. Thus, the so-called Actor focus (AF) affix derives an Actor nominalization, which typically refers to an agent that performs some action. Similarly, a PF affixes derives a nominal expression typically designating what is affected. In the same vein, Locative (LF) and Instrumental (IF) affixes derive nominals designating a place and instrument used to perform some action, respectively.

(18) Mayrinax Atayal (based on Huang 2002)

- a. **m**-aquwas kuʔ irawaiŋ=mu (AF construction)
 AF-sing NOM.REF friend=1SG.GEN
 ‘My friend is singing.’
- a’. Baq-un=mu kuʔ [**m**-aquwas] kaʔ hacaʔ (Actor nominalization)
 know-PF=1SG.GEN NOM.REF AF-sing LIN that
 ‘I know that singer/one who is singing there.’
- b. ma-hnuq kuʔ [β-in-ainay nukʔ naβakis] (Patient)
 AF-cheap NOM.REF buy<PF.REALIS>buy GEN.REF old.man
 ‘What the old man bought was cheap.’
- c. yaɣhapuyan kuʔ [naniq-**an** cuʔ βuŋaʔ nkuʔ ʔulaqiʔ] (Locative)
 kitchen NOM-REF eat-LF ACC.NONREF yam GEN.REF child
 ‘The kitchen is (the place) where the child eats yam.’
- d. kaa ptiq-ani kuʔ [**pa**-patiq=mu] (Instrumental)
 NEG.IMP write-IF.IMP NOM.REF IF-write=1SG.GEN
 ‘Don’t write with my pen/thing to write with.’

In Sasak and other languages in which the focus morphology is reduced, there can be only two or three types of argument nominalizations, as we saw in Sasak above. In any event, it is clear that the gap contained in a relative clause in Austronesian languages is the one that has been created in the process of argument nominalization rather than in the process of relativization, which simply juxtaposes a (head) nominal and an argument nominalized clause with a gap in the appositive syntagm.

3.2 Argument nominalizations in other languages

The pattern of argument nominalization and the role of the focus morphology in Austronesian languages seen above actually are not entirely unique to this language family. At least two language families have similar mechanisms of argument nominalization. Many Tibeto-Burman languages have morphology distinguishing among subject nominalization, object nominalization, and instrument nominalization, as shown in (a) forms below:

(19) Northern Qiang (Ronghong) (Huang 2007: 192-194; pc)

Subject nominalization

- a. **fa** **ɕupu guə-m** (equivalent of Austronesian AF)
 clothes red wear-NMZ
 ‘one wearing red clothes’
- b. [**fa** **ɕupu guə-m**] tɕymi the: (Appositive RC)
 clothes red wear-NMZ child that.CL
 ‘that child who wears red clothes’

(20) Object nominalization

- a. [**qa (-wu) khe**]-tɕ (equivalent of Austronesian PF)
 1SG-AGT cut-GEN
 ‘one I am cutting’

- b. [qa (-wu) khe]-tɕ səf tho-zgu (Appositive RC)

1SG-AG cut-GEN tree that-CL
 ‘the tree that I am cutting’

(21) Instrumental nominalization

- a. pies khukhu-s (equivalent of Austronesian IF)

meat slice-NMZ
 ‘what (is used) to slice meat.’

- b. tse: [pies khukhu-s] xtɕepi ŋuə (Appositive RC)

that:CL meat slice-NMZ knife COP
 ‘That is the knife that is used to slice meat.’

The Rhongon dialect of Qiantg marks subject nominalization by the suffix *-m* and instrumental nominalization by *-s*. Object nominalization, on the other hand, involves no nominalizer; instead it requires a genitive marking as in (20). These nominalized clauses also function as arguments, as illustrated below:

(22) Northern Qiang (Courtesy of Chenglong Huang)

qa [ləyz tse-m] e: u-tɕu-a
 1SG book read-NMZ one.CL DIR-see-1SG
 ‘I see one who is reading books.’

Unto-Aztecan is another language family that display the pattern of argument nominalizations similar to Austronesian and Tibeto-Burman. Here the distinction made is among subject nominalization (*-me*), object nominalization (*-‘u*) and locative nominalization (*-‘Vpol*), and possibly some others.

(23) Yaqui (González 2007; pc)

Subject nominalization

- a. jú-me usí-m yeéwe
 DET-PL child-PL play
 ‘The children are playing.’

- b. jú-me [yeéwe-me] (equivalent of Austronesian AF)
 DET-PL play-NMZ
 ‘the ones playing’

- c. jú‘u yoéme jú-me usí-m [yeéwe-me] kákam máka-k (RC)
 DET man DET-PL child-PL play-NMZ candy-PL give-PERF
 ‘The man gave candies to the children who were playing.’

(24) Object nominalization

- a. inepo uka chu‘u-ta tea-k
 1SG DET.AC dog-ACC find-PERF
 ‘I found the dog.’

- b. in uka tea-ka-‘u (equivalent of Austronesian PF)
 1SG DET.AC find-PERF-NMZ
 ‘what I found’

- c. U chu'u [in tea-ka-'u] chukuli (RC)
 DET dog 1SG.GEN find-PERF-NMZ black
 'The dog that I found is black.'

(25) Locative nominalization

- Wa kari [nim be'e-pea-'apol] ujyooli (equivalent of Austronesian LF RC)
 DEM house 1SG.GEN sleep-DES-NMZ pretty
 'The house where I want to sleep is pretty.'

The following are examples in which argument nominals function as clausal arguments and as a nominal predicate paralleling some Austronesian and Northern Qiang examples above.

(26) Yaqui

- a. [U nim pu'akta-'u] bette
 DET 1SG.GEN bear-NMZ heavy
 'What I bear is heavy.' 'My burden is heavy.'
- b. Jabesa [wa jiosam noktua-me]
 who that book read-NMZ
 'Who is the one that read that book?'

Turkish makes a distinction between subject nominalization and object nominalization in terms of different participial forms of verbs. Present participle ending *-en* marks subject nominalization, while in the case of (one of) the future participle(s), the nominalized form is identical with the base form.

(27) Turkish (Lewis 1967: 158ff)

Subject nominalization

- a. bekliy-en-ler
 wait-PTCPL-PL
 'those who are waiting'
- a'. [bekliy-en] misafir-ler (appositive RC)
 wait-PTCPL guest-PL
 'guests who are waiting'
- b. haber gelecek (future sentence)
 news will.come
 'The news will come.'
- b'. gelecek (subject nominalization)
 'who/which will come, the future'
- b''. [gelecek] haber (appositive RC)
 will.come news
 'news which will come'

(28) Object nominalization (*-dik*, *-cek* plus a personal suffix)

- a. bir tanı-dığ-ım
 one know-P.PTCPL-1SG
 'one I know, an acquaintance of mine'

- b. [tani-diğ-im] bir adam
know-P.PTCPL-1SG one man
'a man I know' ('a man characterized-by-my knowing')

3.3 The “subjects-only” constraint as an Austronesian epiphenomenon

As is clear from the above exposition, many languages from different language families seem to have a relative clause formation similar to the Austronesian RC pattern, all making use of nominalized clauses juxtaposed to the head nominal. Despite this similarity, none of the specialists of Tibetan languages and Uto-Aztecan languages speaks of the “subjects-only” constraint similar to the one noted by Keenan and Comrie (1977) and by some Austronesian specialists on the Austronesian relativization. For example, Ross (1995: 729-730), while using the less charged term “pivot” than “subject”, tells us that “in a PAN [Proto Austronesian] relative clause the (deleted) noun phrase coreferential with its head noun had to be its pivot.” This is a curious fact in view of the clear parallelism in the relativization pattern across these different language families. The answer to this puzzle that I offer is that the “subjects-only” constraint is an epiphenomenon seen only in Austronesian, where predicate formation in Proto Austronesian also involved argument nominalizations.

According to Starosta, Pawley and Ried's (1981) hypothesis, modern Austronesian clause structures evolved from equational predicate-topic constructions, where an argument nominalized clause functioned as a nominal predicate, as illustrated by the Tagalog examples below:

(29) Tagalog (courtesy of Naonori Nagaya)

AF NOMINAL PREDICATE + TOPIC

- a. [H<um>i-hiwa ng=karne] + [ang lalaki]
 RED<AF>-cut GEN=meat TOP man
 ‘one cutting meat’ ‘the man’
 ‘The man is the one cutting meat.’ → ‘**The man** cut meat.’
 (AF construction)

PF NOMINAL PREDICATE + TOPIC

- b. [Hi-hiwa-in ng=lalaki] + [ang=karne]
RED-cut-PF GEN=man TOP=meat
‘one the man is cutting’ ‘the meat’
‘The meat is the one the man is cutting’ → ‘The man cut **the meat**.’
(PF construction)

Starosta, Pawley and Reid's (1981) account posits a reanalysis of equational predicate-topic construction into more tightly integrated structures in which the topic nominals have been reanalyzed as arguments of the verb of a nominalized predicate nominal, thereby creating a situation where the integrated topic nominal is understood to trigger the focus marking in the verb.

Thus, argument-nominalized clauses with a gap in Austronesian are involved in both relative clause formation and in the formation of one-place predicates, and it is this dual function of Austronesian nominalized clauses that engenders the “subjects-only” effect, if relative clause formation is believed to involve a full clause as a modifying clause.

It is possible to talk about the grammatical relation of the gap in the nominalized clause juxtaposed to a head nominal, though it is presumptuous to do so since the structures of nominalized clauses are different from those of sentences and the grammatical status of the arguments in nominalized clauses is not entirely clear at this stage of research; e.g. are nominalized agentive nominalized clauses in the dative case really subjects like the nominative subjects of main clauses? One might characterize the following relative clause constructions in Yaqui as cases of subject relativization and object relativization on the basis of the presumed grammatical roles of the gaps in the nominalized clauses.

(30) Yaqui

Subject relativization

- a. jú'u yoéme jú-me usí-m [Ø yeéwe-me] kákam máka-k
 DET man DET-PL child-PL play-NMZ candy-P give-PERF
 'The man gave candies to the children who were playing.'

Object relativization

- b. U chu'u [in Ø tea-ka-'u] chukuli
 DET dog 1SG.GEN find-PERF-NMZ black
 'The dog that I found is black.'

One could also describe Austronesian relativization patterns in the same way, but as soon as one did so, the "subjects-only" constraint would disappear. In both (30a) and (31a), the gap occurs where a subject is expected, and in both (30b and (31b), what are missing can be construed as objects since what look like subjects occur in the genitive form, as agentive nominals generally do in nominalized clauses in Japanese, Turkish and many others.

(31) Tagalog (courtesy of Naonori Nagaya)

Subject relativization?

- a. mga bata-ng [nag-la-laro Ø]
 PL child-LIN AF.PRF-DUP-play
 'children who are playing'

Object relativization?

- b. aso-ng [na-kita ko Ø]
 dog-LIN PF.PERF-see 1SG.GEN
 'dog that I saw'

3.4 Genitives as nominalizations

Many languages show formal similarity between the genitive construction and nominalizations, as noted by Matisoff (1972) for Sino-Tibetan and others (see also Deutscher's symposium paper). A clear case of this is Chinese, in which the particle *de* marks genitive adnominal modification (32a), clausal nominalization (32b), and adnominal modification by a nominalized clause (32c), as well as nominalization of a noun itself (32d).

(32) Chinese (based on Li and Thompson 1081:113, 116, 576; courtesy of Chris Schmidt)

- a. [wǒ] **de** chènshān
I NMZ shirt
'my shirt'
- b. nǐ méi yǒu [wǒ xǐhuān] **de**
you not exist I like NMZ
'You don't have what I like.'
- c. [Zhāngsān mǎi] **de** qìchē hěn guì
Zhangsan buy NMZ car very expensive
'The car that Zhangsan bought is very expensive.'
- d. Zhè shì [wǒ] **de**, nà shì [fùqīn] **de**
this is I NMZ that is father NMZ
'This is mine, and that's the father's.'

In Modern Japanese the particle *no* marks a noun standing in the genitive relation to its head as well as a nominalized clause as a whole, but not when a modifying nominalized clause has a head, as in RC construction (33c).

(33) Japanese

- a. [Taroo]=**no** hon
Taro=NMZ book
'Taro's book'
- b. [Taroo=ga katta]=**no**=o boku=mo hosii.
Taro=NOM bought=NMZ=ACC I=too want
'I too want what Taro bought.'
- c. [Taroo=ga katta] **hon**=o boku=mo hosii.
Taro=NOM bought book=ACC I=too want
'I too want the book that Taro bought.'
- d. Kore=ga [boku]=**no** de, are=ga [otoosan]=**no** da.
this=NOM I=NMZ COP that=NOM father=NMZ COP
'This is mine and that's the father's.'

In Old Japanese verbs and adjectives had special inflectional endings in their nominalized function distinct from finite predicate forms. When nominalization was indicated by inflection, the nominalized clause by itself functioned as an argument without *no*, and at that stage what corresponded to (33b) could take the accusative case marker without the nominalization marker. In other words, the nominalization marker *no* occurred when the form did not inflect for nominalization, namely a noun. (There is a possibility that this *no* was an inflected form of the copula.) Middle Japanese lost the inflectional distinction between the finite and the nominalizing endings, which presumably prompted the use of *no* where, as in (33b), the predicate showed no sign of nominalization inflection any longer, but not where the nominalization function is clear as in the pre-head position, as in (33c).

Genitives as nominalizations of nouns may sound odd, but we saw earlier that Indonesian languages and others can nominalize demonstratives and ordinal numbers. The nominalization of nouns derives forms with the meaning of "a thing pertaining to N",

and this is the meaning that the genitive case in general expresses, including that of possession, of creation (Picasso's painting), of professional association (John's colleague), etc. etc. Indeed, the meaning of "a thing pertaining to N" is a general meaning of the attributive function of a nominal expression in the appositive construction with a head, as in the relative clauses in Austronesian, Sino-Tibetan, Uto-Aztecan, Turkish and Japanese that we have seen above. The relative expression "the man that I know" in these languages literally means something like "the man, the one pertaining to my knowing" just as the genitive construction "John's friend" means "the friend, the one pertaining to John". (Li and Thompson's (1981:113ff) labeling of the Mandarin genitive *de* phrases as "associative phrases"; Lewis' (1967:164.) literal translation of the personal participle RC form [*kardeş-im-in bekle-diğ-i*] *misafir* 'the guest whom my brother is/was waiting' ('pertaining to my brother's waiting'.)

4. Sentences and clauses

When I asked a couple of experienced Mayan specialists if the K'ichee' relative clause in (34a) was nominalized, their immediate answer was "No!"—there is no nominalization morphology, and it can moreover stand as a complete sentence, as in (34b).

(34) K'ichee' (Larsen and Norman 1979)

- a. lee ixoq lee [x-Ø-u-ch'ay lee achih]
the woman REL ASP-3SG.ABS-3SG.ERG-hit the man
'the woman whom the man hit.' (Or 'the woman who hit the man')
- b. x-Ø-u-ch'ay lee achih
ASP-3SG.ABS-3SG.ERG-hit the man
'He/she hit the man.'

It takes some convincing to do to change their opinion, as I have tried with one of the Mayan specialists I spoke to, by showing that the relative clause in (34a) functions as a nominal element paralleling simple noun phrases, as below:

- (35) a. x-Ø-inw-il lee [ixoq]
ASP-3SG.ABS-1SG.ERG-see the woman
'I saw the woman.'
- b. x-Ø-inw-il lee [x-Ø-u-ch'ay lee achih]
ASP-3SG.ABS-1SG.ERG-see the ASP-3SG.ABS-3SG.ERG-hit the man
'I saw the one whom the man hit.' (Or 'I saw the one who hit the man.')

This simple anecdote shows how much linguists are (still?) preoccupied with form. "Nominalization" is a functional (not a formal) notion referring to conversion of a form into an element that functions as a nominal element of a sentence structure. Languages may have morphological or other formal means indicating this functional change, but they may not, as we saw earlier that Austronesian nominalizations may not involve any formal mechanism (see (16) and (17)). Along the similar line, I have also heard a fairly experienced linguist claim that her language has no subordinate or complement clauses because all such clauses are finite. (See Tom Givon's "Multiple routes" symposium paper.) But here again, a distinction between the formal and functional aspect of the

finiteness phenomenon has escaped attention. Forms may have finiteness features but they may lack the functional correlates of these features, and this is precisely the distinction between the nominalized clause seen in (34a) and (35b) and sentence (34b). While they are identical in form and are presumably both finite in form, only sentence (34b) plays the pragmatic predication function, e.g., making an assertion.

I opened this paper by noting that the distinction between sentences and clauses is not made in the Chomskyan recursive phrase structure rules, which imply that sentences may recur in clause internal positions, and linguists have often failed to make this important distinction. Though an intuitive understanding of the distinction between the two seems to be there, linguists have generally tended either to gloss over the distinction or to be confused about the relationship between the formal finiteness features characterizing sentences and the pragmatic functions they perform. This confusion is understandable because certain embedded clauses do display some formal finiteness features such as tense marking and agreement features in the verb. While the distribution of formal finiteness features in different types of clauses deserves close studies, finite embedded clauses are by no means functionally finite—and are accordingly not sentences—in that they fail to make a predication that sentences in their use do.

Unfortunately the term “predication” itself is used both vaguely and ambiguously. I want to recognize at least two senses of this term. One is a syntactic sense reflecting the role verbs play in organizing and giving coherent structure to a proposition. Enumeration of nouns does not form a coherent proposition, since nouns only have a referring function. Verbs are said to play a predication function in their role of unifying thoughts by describing the properties and relations that nouns may have. Simply pointing out objects in the world does not lead to a coherent thought. It is only by describing their nature by the use of verbs that a coherent thought, a proposition, is formed. The term prediction is also used to refer to the formal relation manifesting this function of the verb, namely the subject-verb syntagm. All (full) clauses and sentences display this formal predication relation, and here morphological trappings such as tense marking and agreement features are irrelevant.

These morphological trappings correlate with the pragmatic notion of predication, namely the act of anchoring a proposition in specific speech context, which is exactly what the speaker does in uttering a sentence. A sentence, in other words, has the pragmatic predication function of grounding a proposition in speech context by establishing connections between a proposition (and its elements) and the elements of speech context, namely the speaker, the hearer, deictic (both spatial and temporal) orientations, etc. etc. It is by virtue of this predication function of a sentence that it can have truth value. In addition to tense marking and locative adverbials, all kinds of discourse particles, honorifics, evidentials and modality are related to this predication function. Sentences are associated with illocutions such as assertion, questioning and promising precisely because of the connection that a sentence establishes itself with the speaker. Sentences, in other words, provide interface between the speaker’s inner thought and the outside world.

Clauses are different from sentences in that they do not connect directly to the outside world even if they have certain formal finiteness features. For example, the clause *Mary is kind* in the sentence *John believes that Mary is kind*, does not assert that Mary is kind, and negating the sentence as *John doesn’t believe that Mary is kind* does not lead to the

negative assertion that Mary is not kind. The clause *Mary is kind*, despite its finiteness features, does not perform the pragmatic predication function. The same is true with the type of relative clauses, restrictive relative clauses, which we have examined in this paper, even though most of them display formal finiteness features, such as tense, aspect, and focus marking.

A similar confusion over the distinction between formal finiteness and functional finiteness or predication function of a sentence is wide-spread in another domain of syntactic complexity, namely verb serialization. The widely held definitions of serial verb constructions, on the one hand, talk about the ability of serialized verbs to occur independently—as a main verb in non-serial context—and about the fact that serialized verbs allow a single predication or single tense, aspect, and polarity value.

(26) Bril (2004:2-3)

- Verbs and Verb phrases (or predicates or nuclei) constitute one single predication referring to aspects of a single event;
- Lexical autonomy is a prerequisite for serialization, excluding non-autonomous coverbs and nonfinite forms, as well as co-lexicalized compounds.

(27) Aikhenvald (2006: 1)

- [SVCs] are monoclausal... and they have just one tense, aspect, and polarity value.
- Each component of an SVC must be able to occur on its own right.

Clearly these definitions contain internal contradictions. If serial verbs together make a single predication, then it is the case either that one of the verbs lacks predication function or dependent on the other, or that the two verbs jointly make a single predication, so that in either case, the two verbs are not autonomous and cannot make individual predications on their own when they are taken out of the serial context. Or if each component of an SVC is autonomous and can make an individual predication, then they do not constitute a single predication. By the same token, when there is only one tense, aspect and polarity value in SVCs, each component should not be able to occur on its own, for in order for a verb to constitute a sentence on its own, it must have its own tense, polarity value, etc. to ground itself in the speech context, as described above.

In this paper I have tried to show that recursion of sentences do not occur in (restrictive) relative clauses or in other embedded constructions. Thus, diachronic processes should not produce structures like the ones sketched for Heine and Kuteva's (2007) hypothesis regarding the rise of subordinate structures, where a sentence is embedded in another sentence, as in the Chomskyan recursive phrase structure rules. Or is it a matter of sloppy use of the symbol S, by which they simply mean a clause rather than a sentence? If synchronic grammar is concerned only with clauses rather than sentences, the Chomskyan recursive rules may be valid, but not for clauses such as relative clauses, many of which are not full clauses, as has been argued above. On the other hand, it is more troublesome to consider S to represent a clause rather than a sentence, for the diachronic scenario. The notion of a clause is only relevant when there is a distinction between a sentence and a clause, and at the stage where there was not distinction of this sort, it is a sentence, not a clause, which existed.

The prohibition of sentence recursion reflects the cognitive limitations associated with the unit of sentence, namely two or more predications cannot be easily made in the space of a single sentence, or an intonation unit of several seconds, due to the kind of severe limitations that Chafe, Pawley (symposium paper), and others say the human mind has in processing information. Then formation of complex structures requires a process of converting a sentence into an abstract object that is stripped of the predication function of a sentence but still has the propositional content of a sentence. Nominalization is one such cognitive process that must be acquired before complex syntactic structures of the embedding type can emerge.