

Elements of complex structures, where recursion isn't

The case of relativization*

Masayoshi Shibatani
Rice University

In their recent work, Hauser, Chomsky and Fitch (2002:1569) suggest that recursion "is the only uniquely human component of the faculty of language". In both generative and typological studies, the relativization site has been considered to be one of the places where recursion of sentences takes place. This paper examines a number of wide-spread patterns of relativization around the globe and argues that what have been identified as relative clauses/sentences are in fact nominalized entities, lacking some crucial properties of both full clauses and sentences. It is furthermore shown that these nominalized forms are neither syntactically nor semantically subordinate to, or dependent on, the nominal head they modify.

*Work reported here was supported in part by the National Science Foundation grant BCS-0617198. I am grateful to the Institute for Advanced Study and the Research Centre for Linguistic Typology at La Trobe University in Melbourne, Australia for the hospitality extended to me while I was revising this paper. Special thanks are due to Sasha Aikhenvald, Noritaka Fukushima, Albert Alvarez, Hiroshi Hayashi, and Roberto Zariquiey for the helpful discussions on the topics dealt with in this paper and for their providing me with relevant examples from their extensive knowledge of the languages identified below. The following list includes other colleagues who provided data on the languages of their expertise, for which I am very grateful. Examples not identified for their sources are either due to the courtesy of the following individuals or based on my personal knowledge: Alexandra Aikhenvald (Russian), Michael Colley (Brazilian Portuguese), Jaime Daza (Bolivian Quechua), Bjarke Frellesvig (Danish), Telma Kaan (K'ichee'), Noritaka Fukushima (Spanish), Albert Alvarez (Yaqui, Spanish, French), Hiroshi Hayashi (French), Martin Hilpert (Swedish), Chenglong Huang (Qiang), František Kratochvíl (Czech), Yu Kuribayashi (Turkish), Thomas Larsen (K'ichee'), Naonori Nagaya (Tagalog), Akio Ogawa (German), Stella Telles (Brazilian Portuguese), Vera Podlesskaya (Russian), Roberto Zariquiey (Spanish, Quechua).

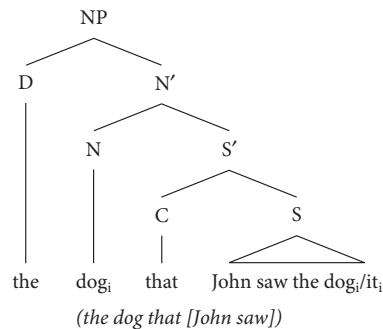
1. Chomskyan notation for recursion and syntactic structures

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

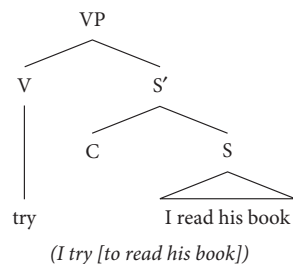
- (1) a. $S \rightarrow NP VP$
 b. $NP \rightarrow (D)N'$
 c. $N' \rightarrow N(S')$
 d. $VP \rightarrow V(S')$
 e. $S' \rightarrow (C) S$

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 on 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



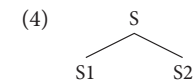
- b. Verb complement



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 (2a) 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 formulated in terms of the grammatical relations of the NPs relativized on. The achievements made in this work, however, crucially depend on the assumption noted above. Once this assumption and other attendant assumptions are abandoned, the achievements would turn out to be no more than a sandcastle. We will see that this is the case below.

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) $S_1 \Leftrightarrow S_2$



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. $S [NP] > S_1 [S_2]$ Expansion (complements and adverbial clauses)
 b. $[S_1 + S_2] > S_1 [S_2]$ 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 as shown in (5). 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, evidential- and 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 nominal 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) or the schematic representation such as (5b) (see Deutscher in this volume for a very similar view). The discussion below is basically synchronic but the plausibility of diachronic speculations based on comparative data from diverse languages such as Heine and Kuteva (2007) and others rests on the correct understandings of the synchronic structures in question.

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 Austronesian languages in maintaining the structural contrast between so-called Actor-focus (AF) and Patient-focus (PF) constructions.¹ 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 its weaker version: “all languages can relativize Subjects.” (Comrie & Keenan 1979:652)

- (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)

1. In this paper the term “Western Austronesian languages” is used as a cover term embracing both Western Malayo-Polynesian languages (of the Philippines, Indonesia, Malaysia, and Madagascar) and Formosan languages (of Taiwan). See Shibatani (2008) for the details of Sasak relativization.

- 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 positions of the relativized NPs in the Keenan-Comrie approach. Comrie and Horie (1995) and Comrie (1998) argue, on the basis of the apparent parallelism between relative clauses with gaps and ordinary sentences with similar gaps in Japanese seen below, that relative clauses (in Japanese) are no different from ordinary sentences with anaphoric 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
 I=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 Sasak
 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 anaphoric gaps 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 Sasak dialects occur in a wide range of modification functions shown below, ranging from both nominal and verbal complements and subordinate adverbial expressions, which do not necessarily have a gap like the one found in a relative clause.

- (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 ART 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 I=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 LOCstreet
 '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" glossed as NMZ in the examples above. Then our understanding of Sasak relativization takes a different turn, namely relativization in Sasak — and many others, as we shall see below — represents one of the modification functions of nominalized expressions rather than a phenomenon involving full clauses or full sentences subordinated to the head nominal, as in the traditional generative and typological analysis. Indeed, *si* and its equivalents in other Austronesian 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]
 NMZ first male and NMZ second
 nine (speaking about one's children)
 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. It is my father who is sitting in the corner.'
- 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.'

- f. *Kamu amaq=k*
 you father-1SG
 ‘You are my father.’

The above examples also present clear evidence that *si*, *saq* and others indeed head nominalized expressions functioning as arguments or nominal predicates in the same manner as simple nouns in (f).

To summarize, the relative clause in Western Austronesian languages is an appositive construction consisting of a nominal (corresponding to a head nominal) and a nominalized expression with a gap (corresponding to a relative clause) juxtaposed as in (12) below. As an appositive construction, the head nominal identifies the entity referred to by the nominalized expression functioning as a relative clause (see further discussion on this point in section 8 below). The nominalized expression, in turn, functions as a modifier of the head nominal.

- (12) Pancor Ngeno-Ngené Sasak
buku [si beng=ku iye Ø]=no
 book NMZ give=1 he =that
 ‘the book that I gave him’

Modification of a noun by another nominal element is not at all rare, as seen in noun compounds (e.g., *gold watch*, *songbird*) and in the use of participle forms as modifiers (e.g., *singing bird*, *a bird singing in the bush*).

3. Austronesian nominalizations

The reason that I prefer using the term “nominalization marker” above to the more conventional “nominalizer” for Sasak particles such as *si* and *saq* heading nominalized forms is that Austronesian nominalizations are in general formed without any such marker. For example, Formosan language Mayrinax Atayal nominalizes (part of) a sentence without any additional marker.

- (13) 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-anıq 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-anıq 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 form without any nominalization marker in Mayrinax Atayal functions both as an argument marked by the nominative particle (13b) and as a nominal modifier (or RC) linked to the head nominal (13c). The parallel pattern obtains in Tagalog, as shown below:

- (14) Tagalog
- 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 like Sasak *si*, *siq*, etc.

What is responsible for argument nominalization in Western Malayo-Polynesian and Formosan languages is the so-called Austronesian focus morphology, which profiles the grammatical role of the argument nominalized. While the same morphology is seen in the predicate of a sentence (cf. (14a) and (14b) above), it is likely that the original function of Austronesian focus morphology lied in argument nominalization (see below on the development of focus-marked sentence predicates). Thus, the so-called Actor Focus (AF) affix derives an Actor nominalization, which typically refers to an agent that performs some action. Similarly, the PF affix derives a nominal expression typically referring to what is affected. In the same vein, Locative Focus (LF) and Instrumental Focus (IF) affixes derive nominals designating a place where a certain action takes place and an instrument used to perform the named action, respectively.

- (15) Mayrinax Atayal (based on Huang 2002)
- a. *m-aquwas ku irawaiŋ=mu (AF construction)*
 AF-sing NOM.REF friend=1SG.GEN
 ‘My friend is singing.’

- a. β aq-un=mu ku? [m-aquwas] ka?
 know-PF=1SG.GEN NOM.REF AF-sing LIN
 haca? (Actor nominalization)
 that
 ‘I know that singer/one who is singing there.’
- b. ma-hnuq ku? [β -in-ainay nuk? na β akis] (Patient)
 AF-cheap NOM.REF buy(PF.REALIS)buy GEN.REF old.man
 ‘What the old man bought was cheap.’
- c. γ ayhapuyan ku? [naniq-an cu? β u η a? nku?
 kitchen NOM-REF eat-LF ACC.NONREF yam GEN.REF
 γ ulaqi?] (Locative)
 child
 ‘The kitchen is (the place) where the child eats yam.’
- d. kaa ptiq-ani ku? [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 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 noun modifier (or RC) in Austronesian languages is the one that has been created in the process of argument nominalization rather than in the relativization process, which, as noted above, simply juxtaposes a (head) nominal and an argument nominalized expression with a gap in appositive syntagm.²

2. Our position regarding the relative clause constructions in Austronesian differs from that of Foley’s (1976), who analyzes RC and participial constructions as instances of the [Adjunct + N] construction, which is defined as “**non-nominal** modifier of a head noun within a noun phrase” (13; emphasis added) distinguishing them from the [Noun + Noun] construction type, which involves “nominalizations, gerunds and possessive phrases” (69). LaPolla with Huang (2003:225) seems to take a similar position (see also Huang 2008). Though the exact types of construction that fall under their [NP + N(P)] are unclear (nominalized RCs in Qiang are included in this type), they seem to exclude finite relative clauses from this construction type. Our position is that all these modification structures, including finite relative clauses, are basically of the [Noun + Noun] type.

5. Argument nominalizations in other languages

The pattern of argument nominalization and the role of the focus morphology in Western Austronesian languages seen above actually are not entirely unique to this language group. Indeed, a fairly large number of languages around the world have argument nominalizations of similar type. Many Tibeto-Burman languages have morphology distinguishing between agent nominalization, patient nominalization, and instrument nominalization, as shown in (a) forms below.³

- (16) Northern Qiang (Ronghong) (Huang 2008:194–196)
- Agent nominalization
- a. fa ζ upu guə-m (equivalent to Austronesian AF)
 clothes red wear-NMZ
 ‘one wearing red clothes’
- b. [fa ζ upu guə-m] t ζ ymi the: (Appositive RC)
 clothes red wear-NMZ child that.CL
 ‘that child who wears red clothes’
- (17) Patient nominalization
- a. [qa (-wu) khe]-t ζ (equivalent to Austronesian PF)
 1SG-AGT cut-GEN
 ‘one I am cutting’
- b. [qa (-wu) khe]-t ζ səf tho-zgu (Appositive RC)
 1SG-AG cut-GEN tree that-CL
 ‘the tree that I am cutting’
- (18) Instrumental nominalization
- a. pies khukhu-s (equivalent to Austronesian IF)
 meat slice-NMZ
 ‘what (is used) to slice meat’
- b. tse: [pies khukhu-s] xt ζ epi η uə (Appositive RC)
 that:CL meat slice-NMZ knife COP
 ‘That is the knife that is used to slice meat.’

The Rhonghon dialect of Qiang marks agent nominalization by the suffix *-m* and instrument nominalization by *-s*. Patient nominalization, on the other hand, involves

3. Whether these processes should be labeled “agent/patient nominalization” or “subject/object nominalization” for Tibeto-Burman and others discussed below is immaterial for our immediate purposes.

no nominalizer; instead it requires a genitive marking as in (17) above (see section 8 on the connection between genitives and nominalizations). These nominalized forms also function as arguments, as illustrated below:

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

Uto-Aztecan is another language family that displays the pattern of argument nominalization similar to Western Austronesian and Tibeto-Burman. In Yaqui the distinction is made between subject nominalization (*-me*), object nominalization (*-ʼu*) and locative nominalization (*-ʼVpo*), and possibly some others.

- (20) Yaqui (Alvarez 2007)
 Subject nominalization
- a. ju-me usi-m yeewe
 DET-PL child-PL play
 ‘The children are playing.’
- b. ju-me [yeewe-me] (equivalent to Austronesian AF)
 DET-PL play-NMZ
 ‘the ones playing’
- c. juʼu yoeme ju-me usi-m [yeewe-me] kakam maka-k (RC)
 DET man DET-PL child-PL play-NMZ candy-PL give-PERF
 ‘The man gave candies to the children who were playing.’
- (21) Object nominalization
- a. inepo uka chuʼu-ta tea-k
 1SG DET.AC dog-ACC find-PERF
 ‘I found the dog.’
- b. in uka tea-ka-ʼu (equivalent to Austronesian PF)
 1SG DET.AC find-PERF-NMZ
 ‘what I found’
- c. U chuʼu [in tea-ka-ʼu] chukuli (RC)⁴
 DET dog 1SG.GEN find-PERF-NMZ black
 ‘The dog that I found is black.’

4. Heath (1972:235) mentions that *-ʼu* appears “not possible to use ...as a modifying RC after a head noun.” Either the information in his source is incomplete or the use of object nominalization in RCs in Yaqui is a recent innovation.

- (22) Locative nominalization
 Wa kari [nim boʼo-pea-ʼapo] ujoyooli (equivalent to Austronesian LF RC)
 DEM house 1SG.GEN sleep-DES-NMZ pretty
 ‘That house where I want to sleep is pretty.’

The following are examples in which argument-nominalized forms function as an argument and as a nominal predicate paralleling some Austronesian and Northern Qiang examples above.

- (23) Yaqui
- a. [U nim puʼakta-ʼu] bette
 DET 1SG.GEN bear-NMZ be.heavy
 ‘What I bear is heavy.’ ‘My burden is heavy.’
- b. Jabesa [wa jiosam noktua-me]
 who DEM 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.

- (24) Turkish (Lewis 1967: 158ff)
 Subject nominalization
- a. beklıy-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’
- bʼ. [gelecek] haber (appositive RC)
 will.come news
 ‘news which will come’

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

- a. bir tani-diğ-im
 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')

6. 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 forms juxtaposed to the head nominal. Despite this similarity, none of the specialists of Tibetan or Uto-Aztecan languages speaks of the “subjects-only” constraint similar to the one noted by Keenan and Comrie (1977) on the Austronesian relativization. Ross (1995:729–730), while opting for the less charged term “pivot” than “subject”, also 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 Reid’s (1982/83) hypothesis, modern Austronesian clause structures evolved from equational predicate-topic constructions, where an argument-nominalized form functioned as a nominal predicate, as illustrated by the Tagalog examples below:

- (26) Tagalog
- | | | | |
|----|---|---|--------------|
| | 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 will cut' | | 'the meat' |
| | 'The meat is the one the man will cut' → 'The man will cut the meat.' | | |
| | (PF construction) | | |

Starosta, Pawley and Reid’s (1982/83) 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 predicate nominal, thereby creating a situation where the integrated topic nominal is understood to trigger the focus marking in the verb.⁵

Thus, argument-nominalized forms 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 expressions that engenders the “subjects-only” effect, if relative clause formation is believed to involve a full sentence as a modifier.

It is still possible to talk about the grammatical relation of the gap in the nominalized form juxtaposed to a head nominal, though it is presumptuous to do so since the structures of nominalized forms are different from those of sentences and the grammatical status of the arguments in nominalized expressions is not entirely clear at this stage of research; e.g., is the agentive nominal marked genitive in a nominalized form (see (27b) below) really a subject like the nominative subject in an independent clause? 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 form.

- (27) Yaqui
- Subject relativization
- a. ju'u yoeme ju-me usi-m [Ø yeewe-me] kakam maka-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 (28a) and (28b), gaps occur where a subject and an object are expected, and in (28b) what looks like a subject occurs in the genitive form, as agentive nominals generally do in nominalized forms as in Yaqui (above), Turkish, Japanese, and many others.

5. See Naylor (1973) and Himmelmann (1991) for the analyses of Tagalog sentence structures in terms of equational predicate-topic constructions in which nominalizations function as a nominal predicate.

- (28) Tagalog
 Subject relativization?
 a. mga bata-ng [nag-la-laro Ø]
 PL child-LIN AF.PRF-RED-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’

7. *Wh*-relatives and their ilk

Perhaps the idea that relative clauses are full clauses or sentences comes from the observation on relative clauses in English and other European languages where interrogative pronouns and other forms are used as relative pronouns standing for the gap in a relative clause, thereby insuring that all the clausal arguments are somehow preserved despite replacement in form and positional displacement of their occurrence. In this regard Bolivian Quechua provides an interesting case and a good introduction to the issues surrounding *wh*-relatives. Like Turkish discussed above, Quechua has two types of participial form representing agent and patient nominalizations.

- (29) Bolivian Quechua (Bills et al. 1971:200)
 a. hasut'i-q
 whip-NMZ
 ‘the whipper, the one who whipped someone’
 b. hasut'i-sqa
 whip-NMZ
 ‘the whipped one, the one whom someone whipped’

As in the other languages examined above, these nominalized forms can function as noun modifiers similar to relative clauses.

- (30) a. [warmi-ta hasut'i-q] runa
 woman-ACC whip-NMZ man
 ‘the man who beat the woman’
 b. [runa(-q) hasut'i-sqa=n] warmi
 man(-GEN) whip-NMZ-3 woman
 ‘the woman whom the man beat’

In addition to these participial nominalization forms, Bolivian Quechua has *wh*-relatives as in European languages. Interestingly, however, *wh*-forms also function as arguments as in (31c) and (32).

- (31) a. pichus mikhu=n? (*pichus* ‘who’ = *pi-chus* ‘person-Dubitative’)
 who eat=3
 ‘Who ate?’
 b. Ni-wa-y [pichus mikhu=n]
 tell-1OBJ-IMP who eat=3
 ‘Tell me who ate.’
 c. rikhu=ni [pichus mikhu=n]
 see=1SG who eat=3
 ‘I saw (the one) who ate it.’
 d. rikhu=ni runa-ta [pichus mikhu=n]
 see=1SG man-ACC who eat=3
 ‘I saw the man who ate it.’
- (32) rikhu=ni [imatachus Maria ruwa=n] (*imatachus* ‘what (ACC)’ = *ima-ta-chus*
 see=1SG what Maria do=3 ‘thing-ACC-Dubitative’)
 ‘I saw what Maria did (the thing that she did).’

It may very well be the case that these *wh*-nominalized forms and their use in noun modification are a calque of Spanish *wh*-expressions, but what is interesting is that in other dialects of Quechua, especially Peruvian varieties, they do not seem to have developed the modification function of these *wh*-forms. For example, Huallaga Quechua has the following *wh*-nominalized forms, but apparently no headed relatives making use of these forms (David Weber & Roberto Zariquiey p.c.).

- (33) Huallaga Quechua (Weber 1983:62)
 a. [ima-wan wallpa-ta wanych-shayki-ta-pis] apamu-y
 what-COM chicken-ACC kill-SUB-ACC-INDEF bring-IMP
 ‘Bring me whatever you killed the chicken with.’
 b. rika-y may-pa away-shayki-ta-pis
 look-IMP where-GEN go-SUB-ACC-INDEF
 ‘Look wherever it is that you may be going.’

English *wh*-relatives and *wh*-nominalized forms (commonly known as “free relatives”) show almost complementary distribution, except for the adverbial forms *where*, *why* and *how*, where the two uses overlap. That is, *wh*-forms usable as relative clauses are not usable as free relatives, and free relatives are not generally usable as noun modifiers, as seen in the following pattern.

- (34) a. I saw the man who/*what won the first prize.
 b. I read the book which/*what you recommended to me.
 c. I will meet the man who/*whoever comes tomorrow.
 d. I read what/*which you recommended to me.

- e. I will meet whoever/*who comes tomorrow.
- f. I will visit the place where you live.
- g. I will visit where you live.

The pattern like this has prevented us from considering free relatives as nominalized expressions in which interrogative pronouns simply stand for the gaps created by the argument nominalization process and from analyzing *wh*-relatives as simply making use of these *wh*-nominalized forms as modifiers. If we look at other European languages, such a possibility emerges more clearly as we see a great deal of overlap between *wh*-relatives and free relatives or between other types of relative clause and their independent nominal use outside the relativization context.

According to Adams (1972:9), Ancient Greek relative clauses are traditionally divided into two types according to their use: (a) “adjectival (if the clause modifies another noun in some way)” or (b) “substantival (if the clause modifies itself)”. In traditional grammar, the term “substantive” is used to refer to nouns and their equivalents, and the examples below show that *wh*-forms in Ancient Greek are no different in function from those nominalized expressions examined above.

- (35) Ancient Greek (Adams 1972: 9, 13)
- a. [hòn gàr theòì philoúsìn], apothnéískei
‘whom the god loves, dies young’
 - b. [hós ou lambáneì tòñ stauròn autoù kai akoloutheì opíso mou] ouk éstin mou áksios
‘(He) who does not take up his cross and follow me is not worthy of me’

Just like the nominalized forms studied above, these Ancient Greek forms are used to modify a noun, forming relative clause constructions as below:

- (36) Ancient Greek (Adams 1972:9, 14)
- a. Lázaros, [hòn égeiren ek nekròn Iēsoús]
‘Lazarus, whom Jesus had raised from the dead’
 - b. Teúkros, [hòs áristos Akhaiòn]
‘Teukros, the best of the Achaeans’

Latin is similar to Greek in that *wh*-nominalized forms can function both as arguments and noun modifiers as seen below:

- (37) Latin (Ehrenkranz & Hirschland 1972: 24, 28)
- a. [qui mentiri solet], peierare consuevit
‘Whoever is in the habit of lying, is accustomed to swear falsely’
 - b. at sunt [qui dicant]
‘but there are those who might say’
- (38) Latin (Gildersleeve & Lodge 1895: 395, 396)
- a. Iūsta glōria, [qui est frūctus virtūtis]
‘real glory, which is the fruit of virtue’

- b. Uxor contenta est [quae bona est] ūnō virō
‘A wife who is good is contented with one husband.’

In modern Romance languages, we again see a great deal of overlap between the argument and the modification function of *wh*-forms.

- (39) Spanish (SP), Brazilian Portuguese (PO), French (FR)
- a. Veré al hombre [que viene mañana] (SP)
Verei o homem [que vem amanhã] (PO)
Je verrai l' homme [qui viendra demain] (FR)
I will see the man that comes tomorrow
‘I will see the man who comes tomorrow.’
 - b. Veré al [que viene mañana] (SP)
Veré a [quien viene mañana] (SP)
Verei [quem vem amanhã] (PO)
Je verrai (celui) [qui viendra demain] (FR)
I will see (to the) that comes tomorrow
‘I will see the one who comes tomorrow.’
- (40) a. Leeré el libro [que usted recomienda] (SP)
Lerei o livro [que qual você recomenda] (PO)
Je lirai le livre [que vous recommandez] (FR)
‘I will read the book that you recommend.’
- b. Leeré lo [que usted recomienda] (SP)
Lerei o [que você recomenda] (PO)
Je lirai ce [que vous recommandez] (FR)
‘I will read the one that you recommend.’

Closer matches between the relative clauses and the argument forms like the above cases are in fact a rarer pattern. In Slavic, nominalized forms headed by Russian *kto* ‘who’ and Czech *kdo* ‘who’ cannot be used as relative clauses, but those headed by *čto/cto* ‘what’ and *kotoryj/kterej* ‘which’ can modify a noun.

- (41) Russian
- a. (tot,) [kto vymyl ruki], mozhët nachatj jestj
(that) who washed hands can start eating
‘The one who has washed his hands can start eating.’
 - b. *Maljčik, [kto vymyl ruki], mozhët nachatj jestj
boy who washed (his) hands can start eating
‘The boy who has washed his hands can start eating.’

- (42) Czech
- a. (ten,) [kdo stojí tám-hle], ještě
 (that) who stand.3SG.PRES there-PART yet
 ne-měl dort
 NEG-have.3S.MASC.PAST cake.ACC
 ‘The one who stands over there hasn’t got the cake yet.’
- b. *kluk [kdo stojí tám-hle], ještě
 boy who stand.3SG.PRES there-PART yet
 ne-měl dort
 NEG-have.3SG.MASC.PAST cake.ACC
 ‘The boy who stands over there hasn’t got the cake yet.’
- (43) Russian
- a. tot, [kotoryj/chto stoit tam], eto drug otca
 that which/what stand there it friend father
 ‘The one who is standing there is my father’s friend.’
- b. tot chelovek, [kotoryj/chto stoit tam], eto drug otca
 that man which/what stand there it friend father
 ‘That man who is standing there is my father’s friend.’
- (44) Czech
- a. ten, [kterej/co sto-jí]
 that.MASC.NOM which.SG.NOM/what.SG.NOM stand-3SG.PRES
 tám-hle], to je táty kamarád
 there-PART it.SG.NOM is dad.SG.GEN buddy.SG.NOM
 ‘The one standing there is my dad’s buddy.’
- b. ten muž, [kter-ý/co sto-jí tám-hle], je táty kamarád
 that.MASC.SG.NOM man.SG.NOM which-SG.NOM/what.SG.NOM stand-3SG.PRES there-PART is dad.SG.GEN buddy.SG.NOM
 ‘The man who is standing there is my dad’s buddy.’

Germanic languages seem to use interrogative pronouns more sparingly in nominalization and relative clauses than in the other languages examined above. Danish does not use *hvem* ‘who’ and *hvad* ‘what’ in relative clauses at all, and they head nominalized expressions rather marginally.

- (45) Danish (the grammaticality rating courtesy of Bjarke Frellesvig)
- a. *[hvem står der] er min fars ven⁶
 [Who is standing there] is my father’s friend.
- b. ???[hvem jeg så igår] er min fars ven
 [Whom I saw yesterday] is my father’s friend.
- c. ???[hvad er på bordet] er min fars
 [What is on the table] is my father’s.
- d. ?[hvad jeg læser nu] er meget interessant
 [What I am reading now] is very interesting.
- e. [hvad jeg læser nu] er avisen (“OK but not very natural”)
 [What I am reading now] is the newspaper

It is significant that those *wh*-elements that do not head nominalized expressions here do not occur as relative pronouns (see also the Swedish forms below). What Scandinavian languages use as a filler for the gap in argument nominalization are adverbial forms *som* ‘as/like’ (for both subject and object nominalizations) and *der* ‘there’ (for subject nominalizations), and the nominalizations headed by these can function as modifiers.

- (46) a. Den [som/der står der] er min fars ven.
 DEF.ART AS/THERE stands there is my father’s friend
 ‘The one who is standing there is my father’s friend.’
- b. Mande-n [som/der står der] er min fars ven
 man-DEF.ART AS/THERE stands there is my father’s friend
 ‘The man who is standing there is my father’s friend.’
- (47) a. Den [som jeg så igår] er min fars ven.
 DEF.ART AS I saw yesterday is my father’s friend
 ‘The one whom I saw yesterday is my father’s friend.’
- b. Mand-en [som jeg så igår] er min fars ven.
 man-DEF.ART AS I saw yesterday is my father’s friend
 ‘The man I saw yesterday is my father’s friend.’

6. Notice that these forms, according to Bjarke Frellesvig, cannot be improved even if the definite article *den* precedes them.

Swedish shows a pattern similar to Danish. Interesting is the fact that *vad* ‘what’ in Swedish, unlike the Danish counterpart, allows an argument nominalization, but its use as a nominal modifier appears not fully established.

- (48) Swedish
- a. Jag läser [vad ni rekommenderar]
I read what you recommend
‘I read what you recommend.’
 - b. ?Jag läser boken [vad ni rekommenderar]
I read book what you recommend
‘I read a book that you recommend.’

German divides gap fillers into two groups. Ones (the article series) used for argument nominalization and relative clauses, and those (interrogative pronouns) used primarily for the former.

- (49) German
- a. Ich empfangе den, [der morgen kommt]
I receive ART ART tomorrow comes
‘I receive the one who comes tomorrow.’
 - b. Ich empfangе den Mann, [der morgen kommt]
I receive ART man ART morning comes
‘I receive the man who comes tomorrow.’
- (50) a. Ich lese (das), [was Sie empfehlen]
I read ART what you recommend
‘I read what you recommend.’
- b. %Ich lese das Buch, [was Sie empfehlen]
I read ART book what you recommend
‘I read the book which you recommend.’
- (51) a. Ich empfangе, [wer (auch) morgen kommt]
I receive who (also) tomorrow comes
‘I receive who(ever) comes in tomorrow.’
- b. *Ich empfangе den Mann, [wer morgen kommt]
I receive ART man who tomorrow comes
‘I receive the man who comes tomorrow.’

Apparently there is a dialectal difference over the use of the *was* ‘what’ relative seen in (50b). Sentences like this seem acceptable to the speakers of certain southern German dialects and Swiss German, but not to High German speakers. In any event, German, like Scandinavian and Slavic languages, shows the pattern where the relative clause

usage of argument nominalizations constitutes a subset of the available argument nominalizations. That is, there are *wh*-type argument nominalizations that are not used as noun modifiers.

Compared to her other Germanic sisters, English has developed the usage of *wh*-forms more extensively. Thus both *who*- and *which*-forms are used in relative clauses, whereas *what*-forms and *wh.ever* forms are used only outside the relativization context. But this division of labor seems a recent development, as older forms of English used *which*- and *who*-forms as arguments. *Middle English Dictionary* (2000) offers the following description about one of the uses of *which*: “[a]s an independent relative with indefinite or generalized force referring to a thing, an abstraction, et., introducing noun clauses” (p. 492) (see example (52a) below). Similarly *whō* was also used “[a]s an independent relative referring to a person or persons used as subj. or pred. nom. introducing noun clauses” (p. 538), as illustrated in (52b). Setting aside the oddity of calling these *wh*-forms as “independent relatives” when there are no “antecedents” for them to relate to, these descriptions are revealing and consistent with our observation and analysis, especially in the recognition that these *wh*-forms head nominalized clauses and that they are referring expressions by themselves (see below).

The use of *who*- and *which*-forms as arguments in Modern English is seen in what Huddleston and Pullum (2002:1076) call “free choice construction” (53a). And finally *which*-forms are also used together with *that*, similar to the Slavic pattern (41)–(44), in older expressions or as archaisms in Modern English, as in (53b).

- (52) Middle English (*Middle English Dictionary*)
- a. And [which falleþ on þat furste flur] schal beo Quene.
 - b. [Who aske this] Leseth his asking trewely.
- (53) Modern English
- a. Invite [who/whom you like]. (Huddleston & Pullum 2002:1076)
 - b. After silence, [that which comes nearest to expressing the inexpressible] is music. (Aldous Huxley *Music at Night* 1931)

The pattern of development we see in English then is specialization of form for different functions, *who*- and *which*-forms for the modification function, and *what*- and *wh.ever*-forms for the argument function, which is consistent with the evolutionary pattern of adaptation and specialization. That is, English *wh*-argument nominalizations were once more widely used both as arguments and modifiers, where the latter were extensions of the former. They then have started to specialize in the specific functions ending up in the more-or-less complementary distribution pattern we saw earlier in (34).

8. What is nominalization?

When I asked a couple of experienced Mayan specialists if the K'ichee' relative clause in (54a) was nominalized, their immediate answer was "No!" There is no nominalization morphology, and the verb form in the relative clause is finite with its arguments fully instantiated in clitic form. Indeed, the relative clause can stand as a complete sentence as in (54b).⁷

- (54) K'ichee' (Larsen & Norman 1979:357)
- 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 took some convincing to change their opinion, as I tried with one of the Mayan specialists I spoke to, by showing that the relative clause in (54a) functions as a nominal argument paralleling a simple noun, as below:

- (55) K'ichee'
- 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.⁸

"Nominalization" is a functional (not a morphological or formal) notion referring to creation of a referring expression. Some (e.g., Weber 1983) prefer the term

7. Independent sentences and relative clauses, however, crucially differ in K'ichee' in that while the former may contain full noun phrases identifying all cliticized arguments, the latter cannot; in a relative clause one of them must be gapped (see example (54a)) as in argument nominalizations in other languages. Thus, *x-Ø-u-ch'ay lee achih lee ixoq* (ASP-3SG.ABS-3SG.ERG-hit the man the woman) 'The woman hit the man' is a possible sentence, but it cannot be used as a relative clause.

8. By this statement, no disrespect is intended to my colleagues in Mayan linguistics referred to above. Also, neither Tom Larsen nor Telma Kaan, whom I also consulted on K'ichee', said that the relative clause in (54a) was not nominalized.

"substantivization" as in traditional grammars of Greek and Latin. Even traditional terms "gerund" and "participle" do not refer to formal properties; "gerund" refers to words derived from verbs that function as nouns, and "participle" to the property of a word "partaking" of the nature of (or sharing the properties of) both a noun and an adjective. Indeed, there are lexical nominalizations that do not show any morphological change but which are recognizable as nominalizations; e.g., nouns *drive, play, walk*, etc. derived from the corresponding verbs in English. Thus, whether or not a form in question has a finite verb form is to a large extent irrelevant, just as isolating languages like Chinese do not show a difference in verb form in either independent sentences or nominalized forms. As long as such forms function as a referring expression like nouns and behave like nouns syntactically, they are nominalized, just as nouns *drive, play, walk*, etc. are.

What we have been identifying as nominalized forms above may be termed "grammatical nominalizations" in opposition to "lexical nominalizations" of the *employer- / employee*-type. While lexical nominalization creates new lexical items belonging to the noun class of the language, grammatical nominalization creates new referring expressions that have no lexical status (see the following discussion on other properties of grammatical nominalizations). Grammatical nominalizations, especially those that show a clausal character, have often been considered a type of relativization and are called "headless relatives" or "free relatives", as if they were derivatives of relative clauses. There is no basis for this other than the fact that they show formal resemblances to relative clauses (for the good reason clear from the discussion above) and the skewed perspective many linguistics have had about grammatical nominalizations, namely viewing them from the perspective of relative clauses. Many indeed attempt to derive these nominalized forms in question from relative clause constructions by deleting a head noun (see Adams (1972) on Ancient Greek and other contributions on different languages in the same volume in which Adams's article is found, Weber (1983) on Quechua, Sneddon (1996) on Indonesian, Huang (2008) on Qiang). Matisoff (1973) also entertains such a possibility for Lahu nominalization but abandons it by noting the following:

Overweighing this consideration is the simple fact that an underlying N_{th} [relative head nominal] is not at all necessary to explain how these sentences work. To interpret the *ve*-clauses as ordinary nominalizations does not distort the meaning, and has the crucial advantage of avoiding multiplication of covert entities... Once we admit deleted N_{th} 's after some *ve*'s, consistency would demand that we stick them in after every post-verbal *ve*, even in the (very numerous) cases where the only semantically possible N_{th} would be an empty one like *š-cə* 'thing' or *š-lə* 'matter'. (Matisoff 1973:484–485)

Another favorite way of analyzing the relevant nominalized expressions as relative clauses is by positing a PRO that functions as a relative head, eschewing the deletion issue. For example,

- (56) Japanese
- a. [Taroo=ga tabe-ta]=no=wa...
Taroo=NOM eat-PAST=NMZ=TOP
'the one that Taro ate is...'
- b. [[Taroo=ga tabe-ta](=no) PRO]=wa⁹...
- (57) Spanish
- a. lo [que usted recomienda]
ART that you recommend
'what you recommend'
- b. lo [PRO [que usted recomienda]
ART that you recommend
'what you recommend']

Indeed, these nominalized forms are typically used in the context where an "antecedent" is available for identifying what they refer to. However, analyzing these nominalizations as relative clauses by positing a fictitious relative head or abstract category such as PRO only reveals lack of understanding of what grammatical nominalization is all about. Notice also that there are so-called free relatives that do not function as relative clauses (see (41), (42), (48), (50), (51)), which preclude the deletion analysis of these.

While Matisoff's (1973: 484) point that the purpose of nominalization (in using the particle *ve* in Lahu) "like an English *-ing* and *to* nominalization is...to reify a clause abstractly, without committing it to the modification of anything else in the sentence" is correct, its thrust is not strong enough. The limitations of Matisoff's understanding of nominalization are revealed by the weakness in his explanation (Matisoff 1972) for the connection between grammatical nominalization, relativization and genitivization, which he recognizes in a fair number of Asian languages. He attributes this pattern of sharing nominalization morphology (e.g., Lahu *ve*, Japanese *no*, Chinese *de*) in the three functional domains to "the nominalizing power of *ve* [and others]" and its connection to the subordinating function (Matisoff 1972: 251, 254). That is, Matisoff's explanation, together with the assumption that relative and genitive constructions involve subordination, is that these nominalization morphemes have a subordinating function and therefore are used in those constructions. Setting aside the real issue of why nominalization has a subordinating function, which Matisoff does not address, his assumption that both relative and genitive constructions involve subordination is questionable in many cases.

9. See, for example, Wrona (2008). There is also a proposal that treats the nominalization marker *no* itself as a pronoun.

In most cases studied above, relative clauses are not syntactically dependent upon their head and they can stand on their own, as shown below, although some languages require overt nominalization markers (as in Japanese *no*) or noun-marking articles (e.g., Spanish *lo*; see below) when the nominalized forms are used as arguments.

- (58) Lahu (Matisoff 1973: 483, 484)
- a. [chu]ve vəʔ qhâʔ jâ lâ
fat NMZ pig expensive very Q
'Is the fat pig very expensive?'
- b. [chu] ve qhâʔ jâ lâ
fat NMZ expensive very Q
'Is the fat one very expensive?'
- (59) Pancor Ngeno-Ngené Sasak
- a. Loq Ali mbace buku [si beng=ku iye]=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]=no
ART Ali N.read NMZ give-1 he=that
'Ali read what I gave him.'
- (60) Yaqui
- a. ju'u yoeme ju-me usi-m [yeewe-me] kaka-m maka-k
DET man DET-PL child-PL play-NMZ candy-PL give-PERF
'The man gave candies to the children who were playing.'
- b. ju'u yoeme ju-me [yeewe-me] kaka-m maka-k
DET man DET-PL play-NMZ candy-PL give-PERF
'The man gave candies to the ones who were playing.'
- (61) Chinese
- a. nǐ méi yǒu [wǒ xīhuān] de chènshān
you not have I like NMZ shirt
'You don't have a shirt that I like.'
- b. nǐ méi yǒu [wǒ xīhuān] de
you not have I like NMZ
'You don't have what I like.'
- (62) Japanese
- a. Taroo=wa [boku=ga yatta] hon=o yonde-i-ru.
Taroo=TOP I=NOM gave book=ACC read-PROG-PRES
'Taro is reading the book that I gave him.'

- b. Taroo=wa [boku=ga yatta]=no=o yonde-i-ru.
Taro=TOP I=NOM gave=NMZ=ACC read-PROG-PRES
'Taro is reading what I gave him.'

(63) Spanish

- a. Leeré el libro [que usted recomienda].
I.will.read the book that you recommend
'I will read the book that you recommend.'
- b. Leeré lo [que usted recomienda].
I.will.read the that you recommend
'I will read what you recommend.'

Contrary to Matisoff's (1972:242) assumption, these relative clauses do not seem to be "semantically subordinate to a noun head" either. To the extent that the (a) forms above entail the (b) sentences and to the extent that the understanding of these entailments does not depend on the understanding of the head noun, the relevant relative expressions here are not semantically subordinate to the head.¹⁰ These facts, both syntactic and semantic, accord better with our earlier suggestion that these constructions are basically appositive.

The same argument applies to the genitive constructions that involve nominalization morphology. Matisoff (1972:242) believes that possessive structures in (64) below are semantically subordinate to a noun head, but the syntactic and semantic relations between the (a) and the (b) forms in (64)–(66) below suggest otherwise.

(64) Lahu (Matisoff 1973:483)

- a. [yɔː] ve vâ? qhâ? jâ lâ
he NMZ pig expensive very Q
'Is his pig very expensive?'
- b. [yɔː] ve qhâ? jâ lâ
he NMZ expensive very Q
'Is his very expensive?'

(65) Chinese

- a. Zhè shì [wǒ]de shū nà shì [fùqīn] de
this is I NMZ book that is father NMZ
'This is my book, and that's the father's.'

10. Compare these relative clauses with noun compounds where a noun modifies another noun. The sentence *I bought a bookcase* does not entail *I bought a book*.

- b. Zhè shì [wǒ]de nà hì [fùqīn] de
this is I NMZ that is father NMZ
'This is mine, and that's the father's.'

(66) Japanese

- a. Kore=wa [boku]=no hon de, are=wa [otoosan]=no da.
this=TOP I=NMZ book COP that=TOP father=NMZ COP
'This is my book and that is the father's.'
- b. Kore=wa [boku]=no de, are=wa [otoosan]=no da.
this=TOP I=NMZ COP that=TOP father=NMZ COP
'This is mine and that is the father's.'

In other words, the genitive forms in these languages are nothing but nominalized forms along side similar nominalized forms that can be juxtaposed to a nominal head as in relative expressions. Remember that Sasak (and other Austronesian languages in Indonesia) can nominalize demonstratives to yield forms like Puyung Meno-Mené *siq ino* 'that one' or Bahasa Indonesian form *yang itu* 'that one' with the nominalization markers *siq* and *yang*. In Lahu, Chinese, Japanese, and others, nouns can be further nominalized to give forms referring to an entity pertaining to the noun that is nominalized. This is precisely what Japanese form *Taroo=no e* (Taro=NMZ painting) 'Taro's picture', for example, means, namely the juxtaposition of 'what pertains to Taro' and 'painting', where the latter identifies "what pertains to Taro". The specific meaning of "pertaining to Taro" can vary. The interpretation of "possessed by Taro" is likely to be the most prevalent one, but other possibilities such as "drawn by Taro" and "Taro being the subject of" exist.¹¹ The proposed nominalization analysis of these genitive constructions also renders superfluous Matisoff's (1973:140ff) deletion analysis of the (b) form in (64); the (b) form in (64) parallels the (b) form in (58), to which Matisoff does not apply a deletion analysis. This parallelism is accounted for straightforwardly by our analysis of genitives as nominalizations.

The essence of nominalization, then, is creation of a referring expression, hence its sharing this essential nominal function with nouns, which refers to a state of affairs characterized by an event denoted by the clause (event nominalization), to an entity characterized in terms of the denoted event in which it has crucial relevance (argument nominalization), or to an entity having crucial relevance to the referent of a noun ("genitive" nominalization). Grammatically nominalized forms are referring expressions

11. Cf. Li and Thompson's (1981:113ff) labeling of the Mandarin "genitive" *de* phrases as "associative phrases", and Lewis's (1967:164.) literal translation of the Turkish 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').

by themselves without any head nominal or pronoun, either abstract or concrete, and they fill the syntactic noun function as arguments or predicate nominals just as lexical nominalizations of the type *employer* and *employee* are referring expressions by themselves and play the nominal syntactic role. Grammatical argument nominalizations referring to entities characterized in terms of events, however, differ from these lexical nominalizations in that they lack identifications, specifications or names; the lexical nominalizations above name the entities referred to. Thus, the English expression *what I bought yesterday*, for example, characterizes an entity referred to in terms of an event of my buying it yesterday, but its identity is not specified — it could be a book, a newspaper, a hamburger, an umbrella, or any other things that could be bought.

The properties and their distribution of grammatical argument nominalizations follow from this process of creating a variety of new referring expressions pertaining to a limitless range of everyday events that vary in their content. First, an argument nominalization is associated with a presupposition that an event characterizing the entity referred to has taken place. The expression *what I bought yesterday* presupposes that I bought something yesterday. This property of an argument nominalization carries over to its relativization function such that relative clauses represent presuppositions, as *the book which I bought yesterday* presupposes that I bought a book yesterday. Thus, neither in the nominalized form nor in its relative clause function, is there assertion of the state of affairs denoted in the presupposition even if the form may contain a finite verb form as in the English examples here. This is an important distinction between sentences and nominalized clauses/forms, and which allows the latter to be embedded or subordinated into a main assertive clause (see below).

Since argument nominalizations are typically created for the nonce, they are often indefinite, and this accounts for the use of indefinite interrogative pronouns such as *what* and *who* as a place holder for the gap created by the nominalization process. Though many languages require such nominalized forms to be preceded by an article or demonstrative, such an article, even if a definite one is used, does not necessarily definitize what follows. This kind of article is becoming grammaticalized and is functioning much like the nominalization markers *no* (Japanese) and *de* (Chinese) seen above. For example, the following Spanish nominalized forms headed by the definite articles do not refer to definite referents identifiable to the hearer; in (b) and (c), with the subjunctive verb forms, they are not even specific — the speaker has no specific referent in mind (see the additional relevant use of *lo* marking what appear to be nominalized in (67d) provided by Albert Alvarez).

- (67) a. María está leyendo su libro y yo voy a leer
 María is reading her book and I FUT to read.INF
 lo [que Juan me dio].
 ART that Juan me gave
 ‘Maria is reading her book and I am going to read what Juan gave me.’

- b. La [que hable inglés] ganará el certamen.
 ART that speak.3SG.SBJNCT English win.3SG.FUT ART contest
 ‘The (female) one who speaks English will win the contest.’
- c. El [que diga que es mentira] es el asesino.
 ART that say.3SG.SBJNCT that is false is ART killer
 ‘The (male) one who says that it is false is the killer.’
- d. Voy a leer lo indispensable/de Cervantes/elegido por Juan.
 FUT to read.INF ART indispensable/of Cervantes/chosen by Juan
 ‘I am going to read the indispensable/the one from Cervantes/the one chosen by Juan.’

Because the identity of what are referred to by argument nominalizations is unspecified, their distribution is constrained in such a way that they typically occur in those contexts where the entity identification is provided or sought — in (a) relative clause constructions, where the head supplies the identification, (b) *wh*-questions that seek the identity of the entity referred to in the equation format, (c) cleft-type constructions, in which the entity identification is made in the equation format, (d) where the entity identification can be made from the context (see also example (65) and (66) above), or (e) the entity referred to is generic (see the earlier Greek and Latin examples in (35) and (37)).

- (68) Puyung Meno-Mené Sasak
- a. dengan [saq nyenke=n tokol leq bucu]=no
 man NMZ PROG=3SG sit LOC corner=the
 ‘the man who is sitting in the corner’
- b. Sai [saq nyenke=n tokol leq bucu]=no
 who NMZ PROG=3SG sit LOC corner=the
 ‘Who is the one sitting in the corner?’
- c. [saq 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.’
- d. Gitaq [saq nyenke=n tokol leq bucu]=no
 look NMZ PROG=3SG sit LOC corner=the
 ‘Look at the one sitting in the corner’

Among these constructions, the *wh*-question (68b) and what looks like a pseudo-cleft construction (68c) are often analyzed as cases of extraction in analogy with the extraction analysis of relative clauses, where the *wh*-element and the identifier nominal in the pseudo-cleft are somehow extracted from the nominalized construction

(see, for example, Larsen & Norman 1979). There is no real motivation for such an analysis. These constructions are the same as simple equational constructions of the following type, which no one would analyze as involving extraction of an element.

- (69) Puyung Meno-Mené Sasak
 a. Sai ino?
 who that
 ‘Who is that?’
 b. Kamu amaq=k
 you father-1sg
 ‘You are my father.’

The only difference between these simple equational sentences and the ones involving nominalized forms is that the latter have presuppositions associated with nominalization, such that (68b–c) presuppose that there is someone sitting in the corner (see above). While in Western Austronesian languages questions involving nominalized forms as in (68b) are normal question forms, some languages have both simple and complex *wh*-question formation, the latter of which involves nominalized forms. For example, in Yaqui either (70a) or (70b) below can be used, the difference being that the latter has a presupposition associated with the nominalized form that someone read that book.

- (70) Yaqui
 a. Jabesa jiosam noktua
 who book read
 ‘Who read the book?’
 b. Jabesa [wa jiosam noktua-me]
 who that book read-NMZ
 ‘Who is the one that read that book?’

9. Conclusion: Sentence and clause

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. 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 predication function 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 subordinate clauses deserves close study, finite subordinate clauses are by no means functionally finite—and are accordingly not sentences—in that they do not assert, order, warn, promise, or express the speaker's ideas/desires/surprises, lacking this central function of sentential predication. Formal finiteness features such as tense inflection in a main clause mark elements that support assertion (e.g., by locating a proposition at a specific time) but their presence in a subordinate clause does not mean that an assertion, for example, is made, as is apparent in the indicative relative clause. As pointed out above, relative clauses, whether they are subordinated to the head or appositive, do not make an assertion, as is clear from the well-known negation test: *I didn't read the book which John recommended to me* does not negate the presupposition that John recommended something to me. Likewise, the Latin subjunctive found in a nominalized form like *Habeō [quae velim]* ‘I have what I should like’ does not express the speaker's opinion “I should like/wish something” as the subjunctive **sentential** form *Velim* does.¹² Thus, even if relative clauses and other types of modification elements are finite and clausal, they are never sentences. Indeed, the fact that a single sentence makes only one assertion seems to reflect a strong cognitive constraint on linguistic structures, and thus failure to make a clear distinction between clauses and sentences misses this very important aspect of human linguistic ability.¹³

While our argument is that many of the so-called relative clauses are not full clauses, let alone full sentences, and while it is not easy to characterize what is and what is not a clause (e.g., is a nominalization based on a transitive verb with a genitive agent and a gapped patient such as Japanese [*Taroo=no Ō katta*]=no ‘what Taro bought’ a clause?), we can still talk about the degree of nominalization in terms of how similar the internal syntax of a given nominalization is to a full clause with a finite verb and the full array of its arguments realized.¹⁴ However, a categorical division between participial nominalizations and those containing a finite verb, for example, is not possible.¹⁵ Participial forms may inflect for tense, as in Palauan, where participles are a separate form class inflecting for the past and the future tense (e.g., *a ?elat el ngikel* (ART smoked LIN fish) ‘smoked fish’; *a lelu kl el babier* (ART to.be.read LIN

12. Significantly, even in Latin, relative clauses do not admit the imperative mood.

13. See Chafe (1987) and Hurford (2003) on the cognitive limitations on a sentence.

14. See Lehmann (1986) for a relevant discussion and his attempt to identify typological correlates of the degree of nominalization.

15. See footnote 2 for proposals to distinguish between, for example, gerundive modifiers and relative clauses structurally. These proposals suffer from their failure to recognize the possibility of nominalizations with finite verb forms. See Noonan (2007) and Koptjevskaja I.-Tamm (1993) on this and related issues in nominalizations.

book) ‘a/the book to be read’ (Foley 1976: 16)). (Also see the Turkish future participle in (24) above.) Participial nominalizations may contain a nominative or an accusative argument, while others turn them (optionally) into genitive (see Quechua examples (30)). Some nominalizations with a finite verb contain a subject in nominative form, while others turn such an argument into genitive (optionally in some case; see the Japanese example immediately above in this paragraph). In still some other forms, e.g., so-called gerunds, arguments may be missing altogether (e.g., *[Reading] keeps your brain going*). Nominalizations of different internal syntax may coexist in a single language, and it is legitimate to ask what such formal differences correlate with. Gołąb (1972: 30–31) offers a ready answer for this as follows:

The question is whether the two categories under consideration, relative clauses and corresponding participles, are functionally identical... The replacement by a participle of a relative clause with a predicative verb preserves the basic external function of the relative clause with respect to its governing noun or pronoun, i.e., the function of an “adnominal” modifier. This replacement, however, deprives the relative clause of **grammatical predication, thus destroying its status as a sentence**. This ultimately means that a process presented primarily as something concrete, taking place in time and space, after the elimination of the relative clause, is conceived as something abstract, beyond time and space. Thus nominalization of a verbal clause deprives it of very important semantic features connected with the “verbality” itself. So, relative clauses and corresponding participles... are not identical semantically, and it is easy to see why many languages contain in their systems two different ways involving different semantic connotations. (Emphasis added)

Everything that Gołąb says above is correct, except he makes the cardinal mistake of confusing sentences and clauses; relative clauses are not sentences, as argued above. They are at best nominalized clauses lacking the predication function of making assertion, commanding, etc. It is this lack of the most essential function of predication that makes nominalized expressions incorporable into a sentence allowing them to function as modifying elements in various constructions such as relative clauses, noun and verb complements, and other types of subordinate modifier.¹⁶ This is the power of nominalization and its connection to the subordination function (Matisoff 1972:251–254).

16. See (10) for the variety of context in which nominalized forms are used in Pancor Ngeno-Ngené Sasak.

References

- Adams, D.Q. 1972. Ancient Greek. In *The Chicago Which Hunt: Papers from the Relative Clause Festival*, P.M. Peranteau, J.N. Levi & G.C. Phares (Eds), 9–22. Chicago IL: Chicago Linguistic Society.
- Alvarez, A. 2007. Relative clauses and nominalization in Yaqui. Paper presented to the Seminario de Complejidad Sintáctica, Universidad de Sonora, Hermosillo, México. November, 2007.
- Bills, G.D., Bernardo Vallejo C. & Troike, R.C. 1969. *An Introduction to Spoken Bolivian Quechua*. Austin TX: University of Texas Press.
- Chafe, W. 1987. Cognitive constraints on information flow. In *Coherence and Grounding in Discourse*, R. Tomlin (Ed.), 21–51. Amsterdam: John Benjamins.
- Comrie, B. 1998. Rethinking the typology of relative clauses. *Language Design* 1(1): 59–86.
- Comrie, B. & Keenan, E.L. 1979. Noun phrase accessibility revisited. *Language* 55(3): 649–664.
- Comrie, B. & Horie, K. 1995. Complement clauses versus relative clauses: Some Khmer evidence. In *Discourse Grammar and Typology: Papers in Honor of John W.M. Verhaar*, W. Abraham, T. Givón & S.A. Thompson (Eds), 65–75. Amsterdam: John Benjamins.
- Ehrenkranz, J. & E.C. Hirschland. 1972. Latin. In *The Chicago Which Hunt: Papers from the Relative Clause Festival*, P.M. Peranteau, J.N. Levi & G.C. Phares (Eds), 23–29. Chicago IL: Chicago Linguistic Society.
- Foley, W.A. 1976. *Comparative Syntax in Austronesian*. Ph.D. dissertation, University of California, Berkeley.
- Gildersleeve, B.L. & Lodge, G. 1895. *Gildersleeve's Latin Grammar*, 3rd Edn. Hampshire: Macmillan.
- Gołąb, Z. 1972. The relative clauses in Slavic, Part 1: Common Slavic and North Slavic. In *The Chicago Which Hunt: Papers from the Relative Clause Festival*, P.M. Peranteau, J.N. Levi & G.C. Phares (Eds), 30–39. Chicago IL: Chicago Linguistic Society.
- Hauser, M.D., Chomsky, N. & Fitch, W.T. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298(22): 1569–1579.
- Heath, J. 1972. Uto-Aztecan relative clauses. In *The Chicago Which Hunt: Papers from the Relative Clause Festival*, P.M. Peranteau, J.N. Levi & G.C. Phares (Eds), 230–245. Chicago IL: Chicago Linguistic Society.
- Heine, B. & Kuteva, T. 2007. *The Genesis of Grammar: A Reconstruction*. Oxford: OUP.
- Himmelman, N.P. 1991. “The Philippine challenge to universal grammar”. *Arbeitspapier Nr. 15* (Neue Folge) Institut für Sprachwissenschaft, Universität zu Köln.
- Hopper, P. & Traugott, E.C. 1993. *Grammaticalization*. Cambridge: CUP.
- Huang, C. 2008. Relativization in Qiang. *Language and Linguistics* 9(4): 735–768.
- Huang, L. 2002. *A Study of Mayrinax Syntax*. Taipei: Crain.
- Huddleston, R. & Pullum, G.K. 2002. *The Cambridge Grammar of the English Language*. Cambridge: CUP.
- Hurford, J. 2003. The neural basis of predicate-argument structure. *Behavioral and Brain Sciences* 26(3): 261–283.
- Keenan, E.L. 1985. Relative clauses. In *Language Typology and Syntactic Description*, Vol. 2: *Complex Constructions*, T. Shopen (Ed.), 141–170. Cambridge: CUP.
- Keenan, E.L. & Comrie, B. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1): 63–99.
- Koptjevskaja-Tamm, M. 1993. *Nominalizations*. London: Routledge.

- LaPolla, R.J. with Huang, C. 2003. *A Grammar of Qiang with Annotated Texts and Glossary*. Berlin: Mouton de Gruyter.
- Larsen, T.W. & Norman, W.N. 1979. Correlates of ergativity in Mayan grammar. In *Ergativity: Toward a Theory of Grammatical Relations*, F. Plank (Ed.), 347–370. London: Academic Press.
- Lehmann, C. 1986. On the typology of relative clauses. *Linguistics* 24: 663–680.
- Lewis, G.L. 1967. *Turkish Grammar*. Oxford: Clarendon Press.
- Li, C.N. & Thompson, S.A. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley CA: University of California Press.
- Middle English Dictionary*. 2000. Part W4, Part W5. Ann Arbor MI: University of Michigan Press.
- Noonan, M. 2007. Complementation. In *Language Typology and Syntactic Description*, 2nd Edn. Vol. 2, T. Shopen (Ed.), 52–150. Cambridge: CUP.
- Matisoff, J.A. 1972. Lahu nominalization, relativization, and genitivation. In *Syntax and Semantics*, Vol. 1, J.P. Kimball (Ed.), 237–257. New York NY: Academic Press.
- Matisoff, J.A. 1973. *The Grammar of Lahu*. Berkeley CA: University of California Press.
- Naylor, P.B. 1973. Topic, focus, and emphasis in the Tagalog verb clause. Ph.D. dissertation, University of Michigan.
- Ross, M. 1995. Reconstructing Proto-Austronesian verbal morphology: Evidence from Taiwan. In *Austronesian Studies Relating to Taiwan*, P. Jen-kuei Li, Cheng-hwa Tsang, Ying-kuei Huang, Dah-an Ho and Chiu-yu Tseng (Eds), 727–791. Taipei: Institute of History and Philology, Academia Sinica.
- Shibatani, M. 2008. Relativization in Sasak and Sumbawa, Eastern Indonesia. *Language and Linguistics* 9(4): 865–916.
- Sneddon, J.N. 1996. *Indonesian: A Comprehensive Grammar*. New York NY: Routledge.
- Starosta, S., Pawley, A.K. & Reid, L.A. 1982/83. The evolution of focus in Austronesian. In *Papers from the Third International Conference on Austronesian Linguistics, Vol. 2: Tracking the Travellers* [Pacific Linguistics C-75], 145–170. Canberra: The Australian National University.
- Weber, J.D. 1983. *Relativization and Nominalized Clauses in Huallaga (Huanuco) Quechua*. Berkeley CA: University of California Press.
- Wrona, J. 2008. The nominal and adnominal forms in Old Japanese: Consequences for a reconstruction of pre-Old Japanese syntax. In *Proto-Japanese*, B. Frellesvig & J.B. Whitman (Eds), Amsterdam: John Benjamins. 199–221.