Prenominal Relative Clauses in Verb-Object Languages[*]

Bernard Comrie

Max Planck Institute for Evolutionary Anthropology
University of California, *Santa Barbara*

The combination of verb-object order with prenominal relative clauses is extremely rare; Dryer (2005e) finds it in only 5 languages in his sample of 756; of these, 3 are Sinitic, one is Sino-Tibetan (probably reflecting either shared ancestry or contact with Chinese), and the fifth is the Formosan language Amis. While it is tempting to suggest contact with Chinese as the explanation for this pattern in Amis, further investigation goes against this hypothesis, since the same pattern is found in some other Formosan languages, does not correlate with degree of Chinese influence, and may be old in the Austronesian family. This similarity between Sinitic and Amis may be due to chance, though it is also conceivable that it could reflect genealogical or contact relations in the more distant past. The rarity of the combination of verb-object order with prenominal relative clauses has led some linguists to suggest a processing explanation that would mitigate against this combination. However, although rare, the combination seems stable in at least some of the languages where it occurs (e.g. Sinitic), and recent work in child language acquisition suggests that the combination may actually provide a processing advantage to the child acquiring the grammatical relations in a relative clause relativizing an object.

Key words: word order typology, relative clauses, Sinitic languages, Formosan languages

One of the typologically unusual features of Chinese, including here almost all of the Sinitic varieties, is the combination of prenominal relative clauses with verb-object order, as illustrated with Mandarin examples (1) and (2).[1]

---

(1) Wǒ   xǐhuān   tā.
    I   like   s/he
    'I like him/her.'

(2) [Zhāngsān   mǎi-de]   qìchē
     Zhangsan   buy-REL   car
    'the car that Zhangsan bought'

Although other constituent orders are possible, as in (3)-(5), the kinds of arguments adduced by Dryer (2005a) suggest strongly that the constituent orders in (1)-(2) are the dominant orders in Mandarin Chinese, and indeed in virtually all other Sinitic varieties.[2] In what follows, I follow the usual convention of the typological literature on constituent order that reference is to dominant order except where otherwise specified, and does not exclude the possibility of other, secondary constituent orders in the language in question, such as Mandarin examples (3)-(5).

(3) Shū   wǒ   mǎi-le.
    book   I   buy-PFV
    'I bought the book.'

(4) Wǒ   bǎ   shū   mǎi-le.
    I   take book   buy-PFV
    'I bought the book.'

(5) Tā   yǒu   yī-ge   mèimei   hěn   xǐhuān   kàn   diànyǐng.
    s/he have one-CLF younger.sister   very   like   see   movie
    'S/he has a younger sister who likes to see movies.'

Likewise, following usual typological practice, I adopt Dryer's functional characterization of relative clauses (Dryer 2005c:366); this approach thus does not contradict claims that so-called relative clauses in some Formosan and other Austronesian languages might be special cases of some more general construction rather than a dedicated syntactic relative clause construction.

---

[1] Mandarin Chinese examples that I assume to be uncontroversial for my illustrative purposes are taken or adapted from Li & Thompson (1981). The following abbreviations are used: CLF: classifier; PFV: perfective; PTCL: particle; REL: relative.

[2] Some Sinitic varieties in heavy contact with Altaic languages have object-verb order, along with prenominal relative clauses; this combination of object-verb order and prenominal relative clause is not infrequent among the languages of the world (109 of the 756 languages in the sample of Dryer (2005e), i.e. 14.4%).

It should be emphasized that what is typologically unusual is the combination of the feature values verb-object and prenominal relative clause. Each of these feature values in isolation is widespread. Dryer (2005b) has 639 languages out of 1370 in his sample with the order verb-object, statistically insignificantly distinct from the 640 languages with the order object-verb. According to Dryer (2005c), prenominal relative clauses are significantly rarer than postnominal relative clauses, by a ratio of 117 to 507 in a sample of 705 languages, but this is still a respectable 16.6% of the languages in the sample.

And this combination of verb-object with prenominal relative clause really is rare. Dryer (2005e) bases his count on a sample of 756 languages, of which a mere 5 show the combination illustrated in (1)-(2). (By contrast, the opposite combination, object-verb combined with postnominal relative clauses, is found in 96 languages in this sample.)

The make-up of this set of 5 languages is also interesting. Three of them are Sinitic languages—Mandarin, Hakka, and Cantonese, these being moreover the only Sinitic languages in the sample. One of them is Bai—for a survey of the constituent order properties of Bai, see Dryer (forthcoming §4)—a language whose genealogical status is controversial, as either a Tibeto-Burman language in heavy contact with Sinitic or a Sinitic language in heavy contact with Tibeto-Burman; see, for instance, Wang (2006) for a recent discussion. This is not the place to enter into a detailed discussion of the genealogical affiliation of Bai, but it should be noted that under either hypothesis the combination of verb-object order with prenominal relative clauses in Bai could be an instance of a phenomenon that is not independent of the phenomenon in (other) Sinitic languages, reflecting either the common inherited pattern of Sinitic or a pattern adopted through contact with Sinitic. Since dominance of prenominal relative clauses is the only pattern in Sinitic and most of Tibeto-Burman (for the latter, see Dryer forthcoming §§3, 5.2), it is the strongest candidate for reconstruction to both Proto-Tibeto-Burman and Proto-Sino-Tibetan.[3] Since dominance of object-verb order is found in most of Tibeto-Burman (Dryer forthcoming §1), it is likewise the strongest candidate for reconstruction to Proto-Tibeto-Burman. If Bai is a Sinicized Tibeto-Burman language and the combination of verb-object order with prenominal relative clauses is the result of Sinitic influence, then this influence would specifically have caused the change in order from object-verb to verb-object (and perhaps have reinforced the retention of prenominal relative clauses).

The fifth language is more interesting: It is the indigenous Formosan language Amis. Discussion with Elizabeth Zeitoun and her colleagues and students informs me

---

[3] It should, however, be noted that reconstructing the constituent order of proto-languages is often a risky undertaking, since constituent order typology is heavily affected by areal factors (language contact), something which makes it difficult or impossible to construct firm arguments concerning the shared ancestry of constituent order patterns.

that this pattern is by no means isolated among indigenous Formosan languages, although equally it is not universal. (Indeed Dryer (2005e) includes a second indigenous Formosan language, Seediq, which has the cross-linguistically more frequent combination of verb-object and postnominal relative clause.) Given that indigenous Formosan languages have been in contact with Chinese since intense Chinese settlement of Taiwan started in the early seventeenth century, one might be tempted to attribute the combination found in Amis and some other indigenous Formosan languages to contact with Chinese (as is suggested, albeit very tentatively, by Dryer (2005e:390)). However—and here I am again grateful to Elizabeth Zeitoun and her colleagues and students—this seems not to be tenable. First, the incidence of this particular combination in indigenous Formosan languages does not correlate at all well with intensity of contact with Chinese; indeed Amis, the language combining verb-object order with prenominal relative clauses that is cited by Dryer (2005e), is actually one of the indigenous Formosan languages that has had least contact with Chinese, and has moreover had more contact with Japanese.[4]

It should be noted that unequivocal statements on the order of relative clause and head noun are infrequent in grammars of Formosan languages, and this may reflect a real difficulty in deciding whether, in languages with alternative orders, one is indeed dominant over the other; if no such predominance of one order can be established, then the language would fall under Dryer's (2005c:367) category of "Mixed types of relative clause with none dominant", which may reflect balanced presence of any of the six individual constituent order types identified by Dryer (2005c:367), although the four types other than prenominal and postnominal relative clauses are each rare (ranging down from just over 2.55% of the sample). Dryer's (2005e:390) claim that Amis has prenominal relative clauses is based on a personal communication from Joy Wu, who works at first hand on the language, and indeed all relative clauses in Wu (2007) are prenominal. Some grammars and other sources for some languages give only examples of prenominal relative clauses, such as Chang & Lee (2002) for Kavalan, Egli (1990) for Paiwan, Li & Tsuchida (2002) for Pazih, and Zeitoun (2002) for Mantauran Rukai. In the case of Pazih, the texts included in Li & Tsuchida (2002) suggest that the dominant order is indeed prenominal; for Mantauran Rukai, Zeitoun (2007:310-311) notes that both orders

---

[4] Japanese has prenominal relative clauses, but like nearly all languages with this feature value has object-verb as its dominant order. If one were to attribute prenominal relative clauses in Amis (and presumably other Formosan languages with this constituent order) to Japanese influence, then one would be left with the problem that while there are other instances of the borrowing of prenominal relative clauses by originally verb-object languages from object-verb languages, e.g. in Semitic languages of Ethiopia under Cushitic influence, in all other known cases this is accompanied by the shift from verb-object to object-verb, i.e. does not lead to the typologically unusual combination we find in Amis.

are possible; while Tang (2002:312) shows that in Paiwan postnominal relative clauses also occur. I would thus be prepared to add Pazih to the list of Formosan languages with clear dominance of prenominal relatives, and suspect that some other languages would also be added to this list, although further work clearly needs to be done.

An added complication is that at least some of the instances identified as prenominal relative clauses might actually be head-internal, as is argued for the Paiwan apparent NRel type by Tang (2002:312), although Aldridge (2004) argues that Seediq has all three of postnominal, prenominal, and internally headed relative clauses. Again, further work is clearly needed. It should, incidentally, be noted that combinations of verb-object order with either internally headed relative clauses or mixed types are also rare in Dryer's (2005e) sample: Application of the interactive research tool shows only 1 language—Kutenai, the subject of Dryer's own field research—as being verb-object with internally headed relative clauses, and only 8 languages as being verb-object with mixed relative clause types, out of a total sample of 756 languages. (By comparison, 17 languages are object-verb with internally headed relative clauses, 30 object-verb and mixed with regard to relative clauses.)

In addition, although it is indeed the case that few non-Formosan Austronesian languages have prenominal relative clauses as their dominant constituent order—indeed consultation of the interactive research tool that accompanies Haspelmath et al. (2005) reveals only one Austronesian language in the sample of Dryer (2005c) with prenominal relative clauses, namely Amis—there are Malayo-Polynesian (i.e. non-Formosan Austronesian) languages that vary between dominant postnominal and recessive prenominal relative clauses, such as Tagalog (Schachter & Otanes 1972:123-124) and Yami (Rau & Dong 2006:125-126), the latter the only Austronesian language of Taiwan that is Malayo-Polynesian. Languages like Tagalog show up in Dryer (2005e) as having postnominal relative clauses, since in cases of clear dominance Dryer, following usual typological practice, only counts the dominant constituent order. Tagalog also has dominant verb-object order, as indeed do most Austronesian languages in Dryer's (2005b) sample (134 out of 153, using the interactive research tool). The fact that prenominal relative clauses are found at least as an option both in the Malayo-Polynesian branch of Austronesian and in several indigenous Formosan languages, currently taken to belong to different primary branches of the family, suggests (but does not of course prove) that prenominal relative clauses might actually be a very old phenomenon in the language family, and in particular that in indigenous Formosan languages with prenominal relative clauses this phenomenon is not to be attributed to contact with Chinese.

Given the rarity of the combination of verb-object order and prenominal relative clauses, it would nonetheless be worth examining whether there are reasons further back in history for the parallelism between Sinitic and some indigenous Formosan languages

in this respect. Proponents of a genealogical link between Sinitic and Austronesian might wish to seek a genealogical explanation, projecting this combination back to the proto-language, although it should be noted in this connection that this becomes somewhat more difficult if one accepts the usual assumption that Sinitic is part of Sino-Tibetan. Most Tibeto-Burman (i.e. non-Sinitic Sino-Tibetan) languages have dominant object-verb order, so one would have to assume that Tibeto-Burman innovated its basic constituent order within the clause in this direction. (Application of the interactive reference tool to Dryer (2005b) shows that 94 of the 102 Sino-Tibetan languages in Dryer's sample have object-verb order; of the 8 that do not, 3 are Sinitic, while 1 is Bai and 4 are Karen. The same procedure applied to Dryer (2005c), incidentally, reveals that in his sample 42 of 57 Sino-Tibetan languages have prenominal relative clauses, including of course all 3 Sinitic languages, while 7 have postnominal relative clauses, including the 4 Karen languages, which have verb-object order.) Alternatively, one might hypothesize that the parallelism between Sinitic and Austronesian languages could have arisen during a period of contact on the Asian mainland before Austronesian moved to its currently assumed point of dispersal, namely the island of Taiwan. It is, of course, also possible that the Sinitic (plus Bai) and Formosan phenomena are historically quite unrelated, i.e. that only chance would be responsible for the fact that the world's only two known potentially independent instances of this phenomenon are spoken in close geographical proximity. (If one looks at enough features and enough languages, then the probability increases that even statistically rare phenomena, such as a coincidence of this kind, will be found in one's sample.)

Given that the combination of verb-object order and prenominal relative clauses is so rare cross-linguistically, one might be led to ask whether there is any motivation for this rarity. One suggestion that has been made is that this particular combination provides problems for processing, in that a relative clause occurring at the beginning of the object in a verb-object sequence will mean that the processing of the main clause will invariably have to be interrupted in order to process such a prenominal relative clause attached to any object noun phrase. This general idea goes back at least to Ross (1967:251) and was developed further, with special reference to relative clauses, by Kuno (1974); for a more general processing-based account of constituent order preferences, including their relevance for cross-linguistic frequency patterns, see Hawkins (1994). Of course, one should also bear in mind the contra-indications, in particular that this combination, if indeed disadvantageous from a processing point of view, has been remarkably stable in the Sinitic languages, and perhaps also shows a reasonably high level of stability in some indigenous Formosan languages. Moreover, the strictures against prenominal relative clauses in verb-object languages should apply equally to postnominal verbal clauses in object-verb languages, but the combination of object-verb order and postnominal relative

clauses is far from rare (96 of 756 languages in Dryer (2005e)).

In addition, some more recent work on the processing, more specifically the child language acquisition of Chinese (Yip & Matthews 2007), suggests that there may actually be advantages to the combination of SVO word order and prenominal relative clauses, in particular in relativizing on objects, as in examples like (6), a Cantonese example cited by Yip & Matthews (2007:282).

(6) [Po_4po_4 $maai_5$] di_1 $tong_4$-$tong_2$ ne_1?
      grandmother buy   CLF   candy   PTCL
     'What about the candy that Grandmother bought?'

In this example, 'grandmother' is interpreted as subject of 'buy', while 'candy' is interpreted both as head of the relative clause and as object of 'buy', which means that the sequence 'grandmother-buy-candy' appears in exactly the same subject-verb-object order as one would find in a simple clause. Yip & Matthews argue that this may well give the child a useful crutch in coming to terms with the identification of grammatical relations in relative clause constructions, especially those relativizing objects. (Indeed, they go on to suggest that children may actually analyze such constructions as head-internal relative clauses, i.e. as [*Po_4po_4 $maai_5$ di_1 $tong_4$-$tong_2$*] *ne_1?*)

It should be noted that while Sinitic languages have the basic clausal constituent order subject-verb-object, this is not the case with most of the indigenous Formosan languages under discussion, which are verb-initial, i.e. have both subject and object after the verb.[5] From the viewpoint of the processing difficulty caused by having to interrupt the processing of a main clause in order to process a relative clause, prenominal relative clauses in a language where the clause order is V-NP-NP would seem to be the absolutely worst scenario, since the processing of every single relative clause will require interruption of the processing of the main clause. (The same would, of course, apply equally to prenominal clauses in a verb-final language, but as noted in footnote 2 the combination of verb-final order and prenominal relative clauses is not particularly rare.)[6] But from the viewpoint of the hypothesis advanced by Yip & Matthews (2007), an advantage similar to that gained in Chinese would hold in these languages too, since a structure like [V

---

[5] It is often difficult or impossible to determine whether they are verb-subject-object or verb-object-subject, in part because of variability in constituent order of nominal arguments, in part because of different voice/focus patterns, which can make identification of syntactic subjects difficult, impossible, or at least controversial.

[6] Dryer (2005d:330) shows, incidentally, that object-verb languages are verb-final by a ratio of 497 SOV languages to only 9 OVS languages in a sample of 1228 languages, with only 4 OSV languages, i.e. the sets of verb-final and object-verb languages are almost identical.

NP$_1$] NP$_2$ would parallel that of a main clause [V NP$_1$ NP$_2$]. Of course, if one accepts that such processing advantages accrue to the combination of verb-object order and prenominal relative clauses, then one is left with the problem of explaining why this combination is so rare—assuming that this rarity is not just the result of chance. (I emphasize again that as one considers more cases, then one increases the probability of occurrence of even statistically rare phenomena, such as an unprincipled skewing between logically parallel phenomena, like postnominal and prenominal relative clauses.)

In conclusion, one can say that the combination of verb-object order with prenominal relative clauses is indeed a typological rarity, although not quite confined to Sinitic and other languages where the pattern is or might be the result of close contact with Sinitic. The existence of this pattern in some indigenous languages of Formosa seems not to be attributable to recent contact with Chinese, although whether these are independent instances of the same rare phenomenon or a single instance resulting from processes further back in history remains an open question. This typological parallel between Sinitic and (some) Formosan could be a coincidence. While claims that the combination gives rise to processing complications may well be true, there may also be processing considerations that operate in the opposite direction, in particular with regard to child language acquisition, which may account for the stability of the combination where it is attested. I acknowledge that my conclusion is inconclusive. But I hope that setting out the problem in this way will encourage those with specific hypotheses to test them against a wider range of relevant data than has hitherto been the case.

References

Aldridge, Edith. 2004. Internally headed relative clauses in Austronesian languages. *Language and Linguistics* 5.1:99-129.

Chang, Henry Yungli, and Amy Pei-jung Lee. 2002. Nominalization in Kavalan. *Language and Linguistics* 3.2:349-368.

Dryer, Matthew S. 2005a. Determining dominant word order. *The World Atlas of Language Structures*, ed. by Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie, 371. Oxford & New York: Oxford University Press.

Dryer, Matthew S. 2005b. Order of object and verb. *The World Atlas of Language Structures*, ed. by Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie, 338-341. Oxford & New York: Oxford University Press.

Dryer, Matthew S. 2005c. Order of relative clause and noun. *The World Atlas of Language Structures*, ed. by Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie, 366-367. Oxford & New York: Oxford University Press.

Dryer, Matthew S. 2005d. Order of subject, object and verb. *The World Atlas of Language Structures*, ed. by Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie, 330-333. Oxford & New York: Oxford University Press.

Dryer, Matthew S. 2005e. Relationship between the order of object and verb and the order of relative clause and noun. *The World Atlas of Language Structures*, ed. by Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie, 366-369. Oxford & New York: Oxford University Press.

Dryer, Matthew S. (forthcoming). Word order in Tibeto-Burman languages. *Linguistics of the Tibeto-Burman Area*. [Draft manuscript downloaded on 2008 June 23 from http://linguistics.buffalo.edu/people/faculty/dryer/dryer/dryer.htm.]

Egli, Hans. 1990. *Paiwangrammatik*. Wiesbaden: Otto Harrassowitz.

Haspelmath, Martin, Matthew S. Dryer, David Gil, and Bernard Comrie. (eds.) 2005. *The World Atlas of Language Structures*. Oxford & New York: Oxford University Press. [Available online at http://wals.info/]

Hawkins, John A. 1994. *A Performance Theory of Order and Constituency*. Cambridge & New York: Cambridge University Press.

Kuno, Susumu. 1974. The position of relative clauses and conjunctions. *Linguistic Inquiry* 5.1:117-136.

Li, Charles, and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.

Li, Paul Jen-kuei, and Shigeru Tsuchida. 2002. *Pazih Dictionary*. Language and Linguistics Monograph Series A2-1. Taipei: Institute of Linguistics, Academia Sinica.

Rau, D. Victoria, and Maa-Neu Dong. 2006. *Yami Texts with Reference Grammar and Dictionary*. Language and Linguistics Monograph Series A10. Taipei: Institute of Linguistics, Academia Sinica.

Ross, John R. 1967. *Constraints on Variables in Syntax*. Cambridge: MIT dissertation.

Schachter, Paul, and Fe T. Otanes. 1972. *Tagalog Reference Grammar*. Berkeley: University of California Press.

Tang, Chih-Chen Jane. 2002. On nominalization in Paiwan. *Language and Linguistics* 3.2:283-333.

Wang, Feng. 2006. *Comparison of Languages in Contact: The Distillation Method and the Case of Bai*. Language and Linguistics Monograph Series B3. Taipei: Institute of Linguistics, Academia Sinica.

Wu, Joy. 2007. "Voice" markers in Amis: a Role and Reference Grammar analysis. *Language and Linguistics* 8.1:95-142.

Yip, Virginia, and Stephen Matthews. 2007. Relative clauses in Cantonese-English bilingual children: typological challenges and processing motivations. *Studies in Second Language Acquisition* 29.2:277-300.

Zeitoun, Elizabeth. 2002. Nominalization in Mantauran (Rukai). *Language and Linguistics* 3.2:241-282.

Zeitoun, Elizabeth. 2007. *A Grammar of Mantauran* (*Rukai*). Language and Linguistics Monograph Series A4-2. Taipei: Institute of Linguistics, Academia Sinica.

Department of Linguistics
Max Planck Institute for Evolutionary Anthropology
Deutscher Platz 6
D-04103 Leipzig
Germany
comrie@eva.mpg.de

VO 型語言中的名前關係子句

Bernard Comrie

德國萊比錫普郎克演化人類學研究所

加州大學聖塔芭芭拉分校

　　VO 語序與名前關係子句的組合是相當少見的；Dryer (2005e) 在他的七百五十六個語言的樣本中只找到五個語言；其中三個是漢語系語言，一個是漢藏語（可能反映有共同的來源或與漢語有所接觸），第五個則是台灣南島語阿美語 (Amis)。雖然我們可以把阿美語的現象歸因於和漢語的接觸，但是深入的調查分析並不支持這個假設；因為在一些其他的台灣南島語中發現的相同現象與漢語的影響程度並無相關，而可能是南島語系中古老的現象。漢語系語言和阿美語間的相似性可能只是偶然（不過也不能完全排除兩者在遠古時代有親屬關係或是彼此有密切接觸之可能性）。有些學者認為 VO 語序與名前關係子句之所以不容易組合在一起是因為這會使語句處理相當困難。不過，雖然罕見，這種組合在某些語言（如漢語系語言）似乎非常穩定，最近的兒童語言習得研究也發現這種組合其實有一些處理上的優勢。

關鍵詞：語序類型，關係子句，漢語系語言，台灣南島語