Journal of Universal Language 4
September 2003, 61-116

A Survey of Relative Pronouns and their Uses in Natural and Artificial Languages

Alan Libert & Christo Moskovsky
University of Newcastle

Abstract

This paper discusses a variety of questions regarding the form and function of relative pronouns in artificial languages. In the first major part of the paper information about relative pronouns in natural languages is presented. From this part the prospective language designer will get an idea of the range of relevant possibilities occurring in natural languages. The second major part shows some of the choices which languages designers have made with respect to relative pronouns in their languages. Generally these choices have been within a narrow range, that is, not all of the possibilities of natural languages have been selected for artificial languages. Specifically, the morphology and syntax of relative pronouns quite often resembles that of Indo-European languages, which is not surprising, since most language designers have been native speakers of an Indo-European language, and many artificial languages have been based on one or more of these languages. The final section of the paper contains some recommendations for the design and use of relative pronouns in artificial languages.

Keywords: relative pronouns, artificial languages, morphology, syntax

1. Introduction

Relative pronouns, and the clauses in which they occur, are a significant feature of many, although not all, of the world's natural languages.[1] They also occur in every artificial language (AL) which we know of.[2] Their morphology and syntax is more complex than one might think, and in fact it is not even clear what should be considered a relative pronoun or a relative clause. Even within the relatively narrow range of constituents which we have decide to limit the terms *relative pronoun* and *relative clause* to, there is a perhaps surprising degree of variation in natural languages with respect to form and behaviour. We shall attempt to give an idea of this variation (which some AL designers might want to be aware of) before turning to the examination of the form and function of relative pronouns in artificial languages, and then offering some suggestions for the design of this area of grammar.[3]

[1] We use the following abbreviations in glosses of examples:

ABS - absolutive	OPrep - object of preposition
ACC - accusative	PART - participle
ANTIPASS - antipassive	RelP - relative pronoun
DAT - dative	PERF - perfect
DO - direct object	PL - plural
F - feminine	PREP - preposition
FUT - future	PRES - present
GEN - genitive	PST - past
IO - indirect object	PSTPART - past participle
IMPERF-imperfect	REL - relative particle or relative inflection
MASC - masculine	SG - singular
OComp - object of the comparative	SU - subject

Also, "OT" indicates that a translation is by us, rather than being made by the source. We have sometimes not indicated the internal structure of words where this was not relevant. Sometimes we have left out a translation when the word for word gloss was sufficient for understanding the meaning of the example.

[2] We exclude from the set of artificial languages those languages made up in relation to a work of fiction, such as Klingon (which we would label fictional languages). These languages have a different purpose than typical artificial languages, and we shall not consider them here. Also, we have confined our survey to those languages which have traditionally published materials available; that is, we have not considered the large number of languages about which information is only available on the World Wide Web.

2. Natural Languages

Relative constructions in natural languages typically consist of a noun phrase (NP) and a pre-modifying or post-modifying clause containing an element which is coreferential with the modified NP:

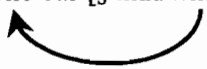
(1) $\left\{ \begin{array}{l} NP_i\ [s \ldots X_i \ldots] \\ [s \ldots X_i \ldots]\ NP_i \end{array} \right\}$

Clausal modification of NPs is a wide-spread phenomenon in natural languages. In fact, linguists involved in the typological description of language (e.g., Downing 1978:385) assert that, in one form or another, NP-modifying clauses exist in all known human languages, which would suggest that clausal modification of NPs is an essential property of natural language.

Relative clauses are by far not the only type of clausal modification of NPs, and, given the objectives of this paper, an essential first step would be to define what is (and what is not) a relative clause. Only then can we address the question of the form and function of relative pronouns (and/or other relative words). Let us consider some of the properties of the relative construction in (2):

[3] We are aware of two previous discussions of relative clauses for an audience of (prospective) AL designers, in Morneau (1994) and Rosenfelder (no date). However, both of these are fairly short.

(2) [$_{NP}$ The car [$_{S}$ that/which/Ø my wife likes]] is too expensive.



(2) is a complex sentence consisting of a main clause *The car is too expensive* and a subordinate relative clause *that/which/Ø my wife likes* modifying the NP *the car*. One of the terms of the relative clause is a relative word (a relative or demonstrative pronoun, a determiner, a complementizer, or some other form of relative marker), which is coreferential with the modified NP, and literature dealing with relative clauses commonly refers to the modified NP as the **antecedent noun phrase**[4] **(ant NP**[5]**)**: a convention that we shall adopt for the purposes of this paper. For ease of presentation we shall refer to relative pronouns as RelPs. One should note that sometimes the RelP is the object of a preposition (as in *the man from whom I bought the book*), and sometimes such a prepositional phrase is contained in an NP (e.g., *the book, the author of which I detest, ...*). The anaphoric relation between the RelP and the ant NP can involve morphological agreement in phi-features (person, number, gender) and sometimes also agreement in case between the two; such agreement is much more obvious in languages which are morphologically richer than English (e.g., the Romance or the Slavic languages). The RelP performs a specific grammatical function within the containing relative clause: a direct object (DO) in this case, but can also be subject (SU), indirect object (IO), prepositional object (OPrep), genitive (GEN), object of the comparative (OComp), or can have the function of adverbial modifier. These grammatical functions are also sometimes referred to as 'positions which can (or

[4] The reader should note that another commonly used term for the antecedent NP is *(relative) Head*; thus the modified NP *the car* in (2) would sometimes be described as the Head of the relative construction *the car that/which/Ø my wife likes*. To avoid confusion with the widely used term *head (of phrase)*, we will avoid using the term *(relative) Head*, but will rather use the alternative **antecedent NP**.

[5] The term 'ant NP' is used by Downing (1978).

cannot) be relativized'.

Syntactically, traditional analyses treat relative constructions like the one in (2) as adjunction structures: it is commonly assumed that in English the relative clause right-adjoins either to some intermediate phrasal level of the ant NP (e.g., [$_{NP}$ [DET] [N' [relative clause]]]) or at the level of the maximal projection (e.g., [$_{NP}$ [NP] [relative clause]]). Also, within the relative clause the RelP is traditionally seen as having undergone fronting (analogous to wh-fronting of interrogative forms) leaving a trace (or a gap (Keenan 1985)) behind: in the case of (2), the RelP would be seen as having originated in the position of the DO of the verb *likes*, and moving, at some derivational stage, to clause-initial position. More recent analyses (especially after Kayne's 1994 'antisymmetry of syntax' proposal) assume that it is the ant NP (or some part of it) that undergoes movement from a position inside the relative clause (in this case, the DO position). Considering the merits (or lack thereof) of this proposal goes well beyond the scope of this paper, but it should be noted that a number of its aspects remain insufficiently clear and a matter of contention (Borsley 1997, Bianchi 2000), therefore we would adopt the more traditional adjunction analysis.

In summary, the construction in (2) involves an antecedent NP and an adjoined subordinate finite clause, which contains an anaphoric element 'relating' the clause to the antecedent NP. Functionally, the relative clause modifies the ant NP (or, much less frequently, another clause).

While within English the picture appears to be reasonably clear, crosslinguistically it is significantly less so, with typological studies attesting to a great degree of diversity in the actual syntactic configuration of NP-modifying clauses: so much so that some have proposed that semantic, rather than syntactic, criteria be used for their

characterization[6]. In the approach adopted by Comrie (1989), we should treat as relative clauses not only constructions which are identifiable as relative clauses according to some formal criteria (such as the construction in (2)), but also constructions such as (3):

(3) a. Passengers leaving on flight 738 should proceed to the departure lounge.
b. The good students all passed the examination. (Comrie 1989:143-4)

because they arguably have "straightforward paraphrases as relative clauses in the traditional sense." (ibid.):

(4) a. Passengers who are leaving
b. The students who are good (ibid.)

The position we take in relation to this is that such an approach would be unproductive for the purposes of this paper, because then any nominal involving a (present or past) participle or an attributive adjective that can yield a relative clause paraphrase, such as in examples (5-6), will count as a relative clause construction:

(5) a. untrustworthy politicians
b. politicians who are untrustworthy

(6) a. the smiling girl
b. the girl who is smiling

[6] Comrie (1989:143) for instance argues that "we need a functional (semantic, cognitive) definition of relative clause, on the basis of which we can then proceed to compare relative clauses across languages, neglecting language-specific syntactic differences ...", while Downing (1978:377) suggests that "we must seek on some other [non-syntactic] level a commonality in terms of which the various syntactic manifestations of what we call relative clauses can be described."

Given that we are interested in a specific form (i.e., the RelP) and its functions, we shall adopt a strictly formal set of criteria for characterizing relative clauses. We shall assume that a relative clause is a finite or a non-finite subordinate clause modifying an NP (or, less frequently, another clause) which contains a clearly identifiable overt, or a recoverable covert, relative word (a pronoun or something else). In this approach, the construction in (7a) will be regarded as a relative clause, because the overt relative word can be recovered:

(7) a. The car Ø my wife likes is too expensive.
b. The car that/which my wife likes is too expensive.

In contrast, participial structures such as (4a) will not be regarded as relative clauses, because the corresponding paraphrase involves recovery not only of the relative word, but also of the verb:

(8) Passengers **who are** leaving on flight 738

Crosslinguistically, the formal approach to the characterization of relative clauses adopted here will exclude a range of constructions that relevant literature commonly treats as relative clauses, such as (9)–(11) because they are pre-modifying participial constructions which lack a recoverable relative word:

(9) Turkish
[Hasan-ın Sinan-a ver -diğ-i] patates-i yedim
Hasan-of Sinan-to give his potatoes-ACC I-ate
'I ate the potato that Hasan gave to Sinan.' (Comrie 1989:142)

(10) Finnish
Pöydällä tanssinut poika oli sairas
on table having danced boy was sick
'The boy who danced on the table was sick.' (Keenan 1985:144)

(11) German
der in seinem Büro arbeitende Mann
the in his study working man
'the man who is working in his study' (ibid.)

The fact that in some of these languages (e.g., German) there is a corresponding relative clause paraphrase (as in (12)) does not make these constructions relative clauses:

(12) der Mann, der in seinem Büro arbeitet
the man who in his study works
'the man who is working in his study' (ibid.)

On the other hand, constructions which involve either a clearly identifiable relative marker, e.g., a suffix on the verb (as in (13)), or a recoverable relative word (as in (14)) will be treated as relative clauses:

(13) Korean
[Hyənsik-i kɨ kä-lɨl ttäli-n] maktäki
Hyensik-NOM the dog-ACC beat-REL stick
'the stick with which Hyensik beat the dog'
(Comrie 1989:151)

(14) Malagasy
ny vehivavy (izay) manasa ny lamba
the woman that wash the clothes
'the woman who is washing the clothes' (Keenan 1985:157)

Critics of the formal approach adopted here might argue that it would exclude some Japanese constructions (such as (15b)) which some typological studies define as 'prenominal external relatives', because no identifiable overt relative marker is involved:

(15) a. Yamanda-san ga saʃru o kaʃt-te i-ru
Yamanda-Mr SU monkey DO keep-PART be-PRES
'Mr. Yamada keeps a monkey.'
b. Yamanda-san ga kaʃt-te i-ru saʃru
Yamanda-Mr SU keep-PART be-PRES monkey
'the monkey which Mr. Yamada keeps' (ibid.:143)

It could be argued that constructions like (15b) are formally marked as relative clauses with a different word order and a change in the case marking of the ant NP: the DO marker is removed in the relative construction.

Before we complete these introductory notes, a brief word about the distinction between restrictive and non-restrictive relative clauses is in order. Consider the difference in interpretation between (16a) and (16b):

(16) a. Children who are often naughty need discipline.
b. Children, who are often naughty, need discipline. (Collins & Hollo 2000:119)

In (16a) the function of the relative clause *who are often naughty* is not so much to provide new information about the ant NP children but to restrict the set of possible referents denoted by the nominal *children* to only those that are naughty; in effect, the assertion that children need discipline will only hold for those children that are naughty. In contrast, (16b) can be seen as a general statement holding for all referents of the nominal *children*, with the meaning that (all) children are naughty and (therefore) need discipline. While

most relative clauses can have either a restrictive or a non-restrictive reading, some constructions involving relative clause modification of proper nouns (as in (17a)) or of another clause (as in (17b)) can only have the non-restrictive interpretation (for obvious reasons):

(17) a. John, who will shortly leave for Paris, is very happy.
b. John will shortly leave for Paris, which makes him very happy.

For some reason, restrictive relative clauses have attracted a lot more attention and have been investigated substantially more than non-restrictive relative clauses. In the view taken here, however, the distinction between restrictive and non-restrictive relative clauses is essentially semantic[7]: a distinction which very few languages mark formall[8], and it will generally be ignored in the following discussion of natural languages.

2.1. Typology of Relative Pronouns in Natural Languages

Given the topic of this paper, the question of what forms relative pronouns can take crosslinguistically and what their function can be is obviously of central concern. We shall again take the English example (2), repeated here, as a starting point of our discussion.

(2) [$_{NP}$ The car [$_{S}$ that/which/Ø my wife likes]] is too expensive.

[7] Even though it may have certain formal consequences (e.g., English non-restrictive relative clauses do not tolerate *that* as a RelP).

[8] One such language is Persian, in which the ant NP of a restrictive (but not a non-restrictive) relative clause is marked with a specific suffix 'i':

Mardhā-i [ke ketābhārā be ānhā dāde budid] raftand.
men that books to them you-had-given went
'The men that you had given the books to went.' (Comrie 1989:139)

As the example illustrates, in English we can use as a RelP either an interrogative pronoun (*who*, *which*, *what*, etc.), or *that* (whatever *that* is—a pronoun, a complementizer, or something else). There is also the option of using a zero form[9,10]. We shall discuss each of these in turn. In terms of function, it appears that the RelP performs two distinct and unrelated functions: one is anaphoric, i.e., coreferring with the ant NP; the second is conjunctive, i.e., linking the subordinate relative clause to the main clause. This duality of function may, at least partially, have been responsible for an on-going controversy about the morphosyntactic status of RelPs, especially *that* and analogous conjunction- or complementizer-like forms in other languages (to be discussed shortly).

Using interrogative pronouns, or forms derived from them, as RelPs is very common crosslinguistically. This is the case not only in most European languages: e.g., *welcher* (German), *wat* (Dutch), *qui* (French), *qué* (Spanish), *kotory* (Russian), *ktery* (Czech), *koj-to* (Bulgarian), but also in a large number of non-European languages, e.g., *naa* (Tamazight), *siapa* (Indonesian), *wi.n.* (Tunisian Arabic), *ke* (Persian), *te mač'a* (Tzeltal), *min* (Bambara), etc. Typological studies investigating relative clauses readily recognise the fact that in many languages interrogative and RelPs are the same (or similar) not only in form, but also in syntactic behaviour[11], which would suggest that they share a common conceptual core, but relevant literature seems to be remarkably uninformative in this regard.

As mentioned earlier, there could be morphological agreement

[9] Some non-standard varieties of English sometimes use 'as' as a RelP:

him as worked at the Chloride
he who worked at the chloride factory (Oehl 2001)

[10] The reader should note that *that* and Ø are not normally used with non-restrictive relative clauses.

[11] As mentioned earlier, one widely adopted analysis of wh-relatives assumes that wh-relative pronouns undergo fronting the same way as wh-interrogatives do.

in phi-features (person, number and gender) between the RelP and the ant NP, which in view of their coreferential relation is hardly surprising. Consider the Russian relative construction in (27):

(18) muzhčina [s **kotorym**
man-NOM-3SG-MASC with who-DAT-3SG-MASC
ja včera poznakomilsja]
I yesterday got-acquainted
'the man that I yesterday got acquainted with'
(source: http://www.departments.bucknell.edu/russian /language/kotoryj.html)

As the example above illustrates, the RelP *kotorym* agrees in person, number, and gender with the ant NP *muzhčina*. The two have different case markings which is what we can expect given their different grammatical functions. In some rare cases (e.g., Latin and Ancient Greek), though, we may find agreement even in case: a phenomenon that is traditionally described as *attraction* (Comrie 1989:154) in which the RelP takes the case of the ant NP, as in (19) where we would expect to have the RelP *hõn* marked for the accusative rather than the genitive:

(19) ek tõn póleōn [hõn éxei]
from the cities-GEN which-GEN he-has
'from the cities he has' (Comrie 1989:154)

As regards *that* in sentences like (2), traditional descriptive grammars (e.g., Curme 1962, Quirk *et al.* 1985) invariably define it as a RelP. The reasons for assuming that *that* is the same category as *which* (and other wh-relatives) are obvious: the two are interchangeable, but **cannot** co-occur, while functionally, *that* seems to operate in the same way as wh-relatives do. More recent approaches have challenged that position and have proposed that *that* is not a RelP,

but is in fact the same subordinate conjunction (or complementizer) *that* which we find in sentences like (20):

(20) He told her [**that** he was in love with her].

Some of the arguments put forward in support of that assumption (e.g., Radford 1988) are as follows. Wh-relatives can be used with a preposition, while *that* can't, as in (21), and in contrast to wh-relatives, *that* seems to be barred from occurring in infinitival relative clauses, as in (22), which is what we can expect from a finite complementizer, but not a RelP:

(21) a. the book [PP about which] they were arguing
b. *the book [PP about that] they were arguing

(22) a. She is not a person [which/that you can rely on].
b. She is not a person [which/*that to rely on]. (Radford 1988:483)

Also, while *that* and wh-relatives cannot co-occur in Modern English, this does not necessarily reflect some general restriction on complementizers and relatives co-occurring, because we do find them together in older varieties of English (23) and in other languages (24):

(23) Middle English
a doghter *which that* called was Sophie
(Traugott 1972:156, cited in Radford 1988:486)

(24) Canadian French
la fille *avec qui que* je parle
the girl with who that I speak
'the girl with who I am speaking'

(Lefebvre 1979:80, cited in Radford 1988:486)

These are valid arguments with respect to the morpho-syntactic status of the relative *that*, but we do not see them as conclusive: on the one hand, the fact that in Modern English *that* and wh-relatives cannot occur remains a problem for the approach assuming that *that* is a complementizer in (2), and on the other, English speakers invariably assign an anaphoric value to *that* when it is used in a relative clause (e.g., as in (2)), but never when it is used as a non-relative complementizer. An additional problem for this approach emerges in the analysis of sentences like (25) in which the subject position in the relative clause is relativized.

(25) The man [that told me the news] was my brother-in-law.

English does not allow sentences without an overt subject. If, as traditionally assumed, *that* is a RelP, rather than a complementizer, then the picture in (25) is clear: the pronoun *that* occupies the subject position[12]. If, however, *that* is assumed to be a subordinate conjunction or complementizer, then the absence of a subject in the relative clause in (25) becomes problematic, because conjunctions or complementizers cannot occupy the subject position (and, generally, argument positions). In this approach (25) should be ungrammatical (and obviously it isn't), while (26), which has a proper subject inside the relative clause should be OK, but is, in fact, ungrammatical:

(26) *The man [that **he** told me the news] was my brother-in-law.

Another possibility raised by Downing (1978:385) is that the relative *that* is (at least historically) the demonstrative pronoun, which is not implausible given that in some languages (e.g., Dutch) the demonstrative is used as one of the RelP forms, and there are other languages (Tzeltal) in which the RelP is a combination of the demonstrative and a wh-word (Keenan 1985:151). While this may indeed be the case in terms of origin, in its current usage the relative *that* is more like an invariant particle, as in (27b), than a demonstrative, which in English is sensitive to morphological number (27a):

(27) a. that man → those men
b. the man that came → *the men those came

Overall it appears that while the relative *that* does have some properties which are characteristic of complementizers, it also has properties (more specifically the ability to co-refer with an NP) which are more typical of pronouns. In view of its duality of function (i.e., both as a conjunction and an anaphor), which we mentioned above, the ambiguous status of *that* should not be seen as particularly surprising or exceptional. We therefore propose that the relative *that* and the conjunction *that* are related, but nevertheless different, elements. Some support for such a position comes from languages like Indonesian in which there are two distinct complementizers which seem to be in complementary distribution: a relative complementizer *yang* (alongside the wh-relative *siapa*) used only in relative clauses, and a non-relative complementizer *bahwa* used elsewhere, but not in relative clauses:

(28) a. Dua ekor harimau [**yang** berasal dari Sumatera]
two tail tiger [which come from Sumatera]
mati kemarin
die yesterday
'Two tigers which came from Sumatera died yesterday.'

[12] Or, in a more recent approach, *that* has undergone movement to Spec CP leaving a trace in the position of the subject (the same way wh-relatives would).

b. Mereka percaya **[bahwa** Mary akan menikah
they believe [that Mary will marry
dengan Joko]
with Joko]
'They believe that Mary will get married to Joko.'

The reader should note that conjunction-like relative words are quite common crosslinguistically and we find them in a range of typologically distinct languages e.g., Akan *áà*, Bulgarian *deto*, Hausa *da*, Hebrew *šɛ-*, Tunisian Arabic *illi*, Vietnamese *má*, etc. (Downing 1978). It should also be noted that crosslinguistically it is not unusual to have both a wh-relative and a conjunction-like relative (as is the case in English) even though the two might involve subtle differences in usage; it is also not unusual in some languages (e.g., Indonesian, Hungarian, Hebrew) to have a combination of the two in the same relative clause.

Crosslinguistically, interrogative and conjunction-like relatives, which are typical of postnominal relatives, are not the only forms that can be used to mark relative clauses. The definite determiner/demostrative is sometimes used as a RelP:

(29) German
der Mann, der in seinem Büro arbeitet
the man the in his study works
'the man who is working in his study' (Keenan 1985:144)

(30) Keʃkchi
li kwi:nq li s-ø-sakʃ-ok re
the man the PAST-3S ABS-hit-ANTIPASS PREP
li isq
the woman
'the man who hit the woman' (ibid.:160)

Some languages (e.g., Chinese) use relative particles for that purpose, as in (31):

(31) Mandarin Chinese
[wo de-la (ta) yidim de] neige nanhaizi
I hit-PERF him once REL that boy
'the boy that I hit once' (ibid.:149)

Other languages may use a relative prefix (32) or a relative suffix (e.g., (13)) on the verb of the relative clause:

(32) Luganda
omusajja omukazi gwe-ya-kuba
man woman REL-she-hit
'the man who the woman hit' (ibid.:149)

Some languages use personal pronouns inside the relative clause to mark the position which has been relativized: a phenomenon commonly referred to as *pronoun retention*. Chinese is one such language (see (31)); Persian is another:

(33) Man zan-i râ ke John be u sibe zamini
I woman-the DO that John to her potato
dâd mishenasam
gave know
'I know the woman that John gave the potato to (her).'
(ibid.:147)

The 'retained' pronoun is also commonly referred to as *resumptive* pronoun. In most languages allowing this, pronoun retention is optional, and the possibility to retain the pronoun is at least partially related to which position in the relative clause is relativized (SU, DO, PrepO, etc.). It seems that very few languages allow pronoun

retention in relative clauses in which the subject is relativized, and generally pronoun retention is more likely to occur in positions occurring lower on the Accessibility Hierarchy (to be discussed shortly).

Let us now briefly consider the 'zero' option and also the possibility to use the RelP as a relative adjective; as regards the former, this is when the position of the RelP remains unfilled as in (2), repeated here:

(2) [$_{NP}$ The car [$_{S}$ Ø my wife likes]] is too expensive.

That there is a true relative clause is shown by the fact that one can have a RelP (namely *who(m)*) in the Ø position. It seems that not many languages which have overt relative words also allow the use of zero forms as RelPs. Among the languages with postnominal relative clauses, Downing (1978) lists only English, Danish, and Vietnamese as such languages. These languages do not normally allow the zero form when the SU position in the relative clause is relativized:

(34) *[$_{NP}$ The man [$_{S}$ Ø married my sister Mary]] is a famous architect.

even though sentences analogous to (34) might be permissible in some non-standard varieties of English, e.g., Black American English (Downing 1978).

RelPs are also sometimes used as relative adjectives, which attributively modify a nominal in the relativized position of the relative clause:

(35) John hasn't returned my book, [**which** fact] surprises me.

This type of use of the RelP can be found in various languages, e.g.,

(36), but does not seem to be very common.

(36) Serbo-Croatian
roman o ratu, koje delo prevodim
novel about war which work I-translate
'a novel about the war, which work I am translating'
(Keenan 1985:153)

The reader will note that it is the whole relative complex *which fact*, *koje delo*, not the RelP alone, that refers back to an ant NP (as in (36)) or to the whole clause (as in (35)). It is also noteworthy that this type of use of the RelP seems to be limited to non-restrictive relative clauses, which would explain why the relative *that* would be unacceptable in the position of *which* in (35).

2.2. Typology of Relative Clauses in Natural Languages

As mentioned in the introductory notes, crosslinguistically a major typological division is between **prenominal** and **postnominal** relative clauses, reflecting the relative position of the relative clause to the ant NP (i.e., the NP that the relative clause modifies). The most obvious example of a language with postnominal relative clauses is English:

(2) [$_{NP}$ The car [$_{S}$ that/which/Ø my wife likes]] is too expensive.

and this is the pattern for most European languages, e.g., Germanic (as in (12), repeated here), Romance (as in (37)), Slavic (as in (38)):

(12) German
der Mann, der in seinem Büro arbeitet
the man who in his study works
'the man who is working in his study' (Keenan 1985:144)

(37) French
J'ai lu un roman [qui m'a beaucoup amusé].
'I read a novel that entertained me a great deal.'
(source: http://www.orbilat.com/Modern_Romance/Gallo-Romance/French/Grammar/Syntax/Pronouns/French-Syntax-Pronouns_Determiners-Relative_Pronouns.html)

(38) Russian
Ja ne videl spektakl o kotorom vy govorite
I not see-PAST show about which you talk
'I haven't seen the show you are talking about.'
(source: http://www.departments.bucknell.edu/russian/ language/kotoryj.html)

Prenominal relative clauses are the dominant relative clause type in languages such as Korean (in (13)) and Japanese (in (15b), repeated here:

(13) Korean
[Hyənsik-i ki kä-lil ttäli-n] maktäki
Hyensik-NOM the dog-ACC beat-REL sick
'the stick with which Hyensik beat the dog' (Comrie 1989:151)

(15) b. Yamanda-san ga kaʃt-te i-ru saʃru
Yamanda-Mr SU keep-PART be-PRES monkey
'the monkey which Mr. Yamada keeps' (Keenan 1985:143)

Typological studies investigating relative clauses have established that there is a strong correlation between the basic word order in a language and the position of the relative clause in relation to the ant NP (Downing 1978, Keenan 1985). Thus SVO and VOS languages (i.e., languages in which the verb precedes the object) predominantly use postnominal relative clauses, while in verb-final languages the predominant type of relative clause is the prenominal one. It should be noted, however, that this correlation is not an absolute universal, and there are languages which break away from this pattern. For instance, Mandarin Chinese uses prenominal relative clauses even though it is an SVO language, while Persian, a verb-final language, uses postnominal relative clauses (Downing 1978:383):

(39) Mandarin Chinese
[wo dale (ta) de] neige ren laile
I hit him REL that man came
'The man that I hit came.' (Downing 1978:395)

(40) Persian
Hasan mard-i-rā [ke zan (u –rā) zad] mišenāsad
Hasan man-ACC that woman he-ACC hit] knows
'Hasan knows the man that the woman hit.' (Comrie 1989:148)

Prenominal and postnominal relative clauses are both defined as **external** (Keenan 1985:143) in the sense that the ant NP is external to the relative clause that modifies it. It has been proposed that another basic typological division is between **external** and **internal** relative clauses (e.g., Keenan & Comrie 1977), where an internal relative clause construction is defined as one in which the ant NP is actually inside the relative clause. Tibetan (41) and Diegueño (42) would be examples of languages involving internal relative clauses:

(41) Tibetan
Peemɛ thep khii-pa the nee yin
Peem-ERG book-ABS carry-PART the-ABS I-GEN be
'The book Peem carried is mine.' (Keenan 1989:161)

(42) Diegueño

Tʏnay	ʔʏwa:ø	ʔʏwu:w-pu-L^{y}	ʔciyawx
yesterday	house-DO	I saw-DEF-in	I will sing

'I will sing in the house that I saw yesterday.' (ibid.:162)

A closer examination of these examples will reveal that they involve nominalized or participial constructions used as NP-modifiers that are not marked with an identifiable relative word (pronoun or some other element). In this regard, the two languages cited here are not exceptional: crosslinguistically, constructions defined as internal relative clauses normally lack a distinct relative marker (Keenan 1985:163). This being the case, such constructions would not meet the strictly formal criteria adopted here and therefore will not be treated as relative clauses in the current analysis. One exception to this pattern is Bambara:

(43) tye ye [ne ye so min ye] san
man PAST I PAST horse REL see buy
'The man bought the horse which I saw.' (ibid.:162)

where we find a distinct optional relative marker *min* inside what appears to be an embedded subordinate finite clause. Clearly the construction in (43) meets the formal criteria we adopted for the characterization of relative clauses. The ant NP *so* 'horse' occurs inside the relative clause, which would suggest that it is indeed an internal relative clause. Internal relatives like the one in (43) can be seen as evidence in support of the ant NP moving analysis[13] proposed in Kayne (1994) and elsewhere: it could be argued that, while external relatives involve overt movement of the ant NP from clause-internal to clause-external position, ant NPs in internal relatives remain overtly *in situ*. It could also be claimed that internal ant

[13] Head-raising analysis in their terms.

NPs undergo non-overt movement to a clause-external position at some other level of representation, e.g., at the level of Logical Form (LF)[14]. If this is indeed what happens, then at the level of LF the construction in (43) will presumably have the form of (44):

(44) tye ye [ne ye t_i min ye] so$_i$ san

and will be no different from other premodifying relative clauses.

It is interesting to note that Downing (1978:397) calls constructions like (43) *replacive relative clauses* and assumes that they belong to the category of free (headless) relatives, as in (45), in which there is no (overt) ant NP:

(45) a. I know [what you think].
b. [Whoever thinks that] must be crazy.
c. Put it [wherever you can find a spot]. (Collins & Hollo 2000:120)

It is unclear to us if this is the correct approach in view of the fact that in (43) there is an ant NP even if it is located inside the relative clause which is supposed to be modifying it, while in constructions like (45) there is no such (overt) NP either internal or external to the (bracketed) relative clause.

2.3. The Accessibility Hierarchy in Natural Languages

In this final section on natural languages we shall briefly consider an aspect of relative clause formation which the artificial language constructor may find interesting. It concerns the question

[14] It is now a standard assumption in generative literature that Chinese wh-words undergo such non-overt movement.

which syntactic positions in the relative clause can be relativized. It has been established (Keenan & Comrie 1977, 1979) that natural languages differ in relation to which and how many positions can be relativized, and that such crosslinguistic differences as there are are subject to an implicational universal known as the Accessibility Hierarchy (AH).

AH states the relative order of the positions that can be relativized in a natural language. These are subject (SU); direct object (DO); indirect object (IO); prepositional object (PrepO); Genitive (GEN); and object of the comparative (OComp) (see table 1 below). It could be argued that this ordering reflects level of markedness: from less marked to more marked.

Table 1

Subject (SU)	(a) SU → *That's the man* [**who** *ran away*]
⇩	
Direct Object (DO)	(b) DO → *That's the man* [**who(m)** *I saw yesterday*]
⇩	
Indirect Object (IO)	(c) IO → *That's the man* [**who(m)** *I gave the letter*]
⇩	
Object of Preposition (OP)	(d) OP → *That's the man* [**about who(m)** *I told you*]
⇩	
Genitive (GEN)	(e) GEN → That's the man [**whose sister** I know]
⇩	
Object of Comparative (OC)	(f) OC → That's the man [**who(m)** I am taller **than**]

Not all natural languages relativize all these positions, and the AH also states that if a language can only relativize a single position in the relative clause, it would be the position of the subject (SU); if a language can relativize two positions, these will be SU and DO, and so on. Another aspect of the AH is that if a language relativizes a position lower in the hierarchy it will also relativize all positions that are higher. For instance, if a language relativizes IO, it will also

relativize SU and DO.

The AH is one natural language universal which has also been investigated in second languages (see e.g., Gass 1979, and the references cited there). It has been established that, when second language (L2) learning takes place in a naturalistic setting, the relative clauses of the L2 are acquired in the order specified in the AH: in other words, naturalistic L2 learners first acquire SU relative clauses, then DO relative clauses, and so on. It has, however, also been established that formal instruction in relative clause formation may actually enable the L2 learner to skip stages in acquiring the relative clauses of the L2. Experimental studies have shown that L2 learners only need to be taught how to form relative clauses lower on the AH, and they generalize this knowledge to all positions that are higher. For instance, if L2 learners are only taught how to form IO relative clauses (but not any other types), they will also learn 'by default' how to form SU and DO relative clauses, but not OPrep or OComp. We believe that this has important implications about artificial languages in view of the fact that they are invariably acquired as a second language.

3. Artificial Languages

3.1. The Source of Relative Pronouns

The source of the RelP(s) of an AL, like that of other elements, depends largely on the type of AL. For example, a pan-Romance AL will probably use RelPs which are those of one or more Romance languages, and a simplified version of a natural language will probably take its RelPs from that language, while a mainly *a priori* AL might be expected to have completely new RelPs. Thus for example American Speech, a simplified version of English, has *HU*, *HWITSH*, *HWaT*, and *THCT* as its RelPs.

Most of the many ALs based largely on Latin have RelPs which are at least somewhat similar, if not identical to RelPs of Latin; in particular most of these pronouns have [k] as their initial segment, and sometimes [w] as their second one. Latin-based ALs which have at least one RelP beginning with [kw] include Latinulus (*quis*), and Weltsprache (Eichhorn) (*kwo*).

In Kosmos, which is also based on Latin, the RelPs *qui* and *quicunque* begin with a similar sequence, [kv] (which is what the letters <qu> are supposed to stand for in this language). The same is true of the Novilatin RelP *qvo*. It is probably not a coincidence that the creators of both Kosmos and Novilatin were speakers of German, in which the orthographic sequence <qu> is pronounced [kv].

Examples from Latin-based ALs of RelPs which begin with [k], but which do not have [w] or [v] as their next segment, are *ki* (from Carpophorophilus' Language), *kelo* (from Mundelingva), and *cujo* (from Universal-Latein). However, there are many ALs which are not based on Latin with RelPs which also begin with [k] (or [kw]). These include Esperanto (*kiu*), American (*ce*), Arulo (qua), Eurolengo (*kie*), Panamane (*ki*), pan-kel (*ki*), and Volapük (*kel*). The fact that pan-kel is in this group is interesting, since it was supposed to be a Germanic-based language.

Balta RelPs do not have [k] as their first segment, but they do contain it, as in *oka*. This fact may perhaps be explained by the fact that pronouns generally begin with a vowel in the language (although interrogative pronouns do not, as we shall see).

The majority of the RelPs which we know of that do not start with [k] are, not surprisingly, from *a priori languages*, such as *xu* (aUI), *lr* (Babm),[15] *han* (Oz), and *av* (Ro), *safa* (Sotos Ochando's Language), and *sio* (Suma). It is also not a surprise that the RelPs of Tutonish/nu teutonish, e.g., *hu* and *vat*, do not being with [k]. The RelPs of the Blue Language, a mixed language, also do not have [k]

[15] Each letter in Babm stands for a syllable, and this word is pronounced [lerɑ].

as their first segment, e.g., *ra*,[16] nor do those of another mixed language, Vela, *co* and *ic*[17]. This is true of the RelP of Interglossa, which is an *a posteriori* language; the fact that this pronoun *su* can only act as a subject might make one think that it was derived from the English word *subject*.

3.2. Formal Identity and/or Relatedness of Relative Pronouns and Other Words

3.2.1. Formal Identity of Relative Pronouns and Relative Adjectives

In English relative adjectives are not frequently used; perhaps this is true of languages in general.[18] Some ALs clearly do have relative adjectives (in the sense in which we are using the term) which are the same in form as their pronominal counterparts. This is true of Esperanto: the *PMEG* (in the section called "Rilata *kiu*") states "one can ... repeat the noun in question after *kiu*" (rather than having *tiu* 'that' at the beginning of the antecedent), and gives the following example:

(46)	La	sola	infano	de	mia	amiko,	kiun	infanon
	the	only	child	of	my	friend	which	child

[16] However, Bollack (1900:40) says, "The English relative *that* is generally translated in [the Blue Language] by the conjunction: **ko**. If, however, *that* involves the sense of «whose» or «whom» one shall employ the relatives ...". The word *ska* 'what one, what people' also seems to be used as a RelP in one sentence in Bollack (ibid.:57)

[17] The Vela letter <c> is pronounced [s].

[18] By "relative adjective" we mean the relative word in such situations as the Latin "erant intinera duo, quibus itineribus ..." 'There were two ways, by which ways ...' (Caesar, *Gallic War* I.6, OT, cited by Lewis 1879:1510 as an example of when "the antecedent is repeated after the *rel.*"). Unlike some authors we do not use the term *relative adjective* to refer to possessive forms related to RelPs, e.g., *whose*. We briefly shall discuss these possessive forms below.

vi vidis ĉe mi antaŭ nelonge, mortis hieraŭ
you saw at I before not.long died yesterday
'The only child of my friend, which child you saw at my house not long ago, died yesterday.' (OT)

The use of the relative adjective and the repeated noun eliminates the potential ambiguity with respect to the antecedent, which could otherwise be either the child or the friend. *Kiu* can also be used as an adjective in a free relative clause, although this very seldom happens. An example is given in (47).

(47) Kiu hundo bojas, tiu ne mordas
which dog barks that not bites
'Which dog barks, that one does not bite.' (ibid., OT)

Gode and Blair (1951:27) say about the Interlingua relative words *que* and *qual* that "The former is primarily a relative pronoun, the latter is a relative adjective." The sequence *le qual* (*le* being the definite article) "is pronominal and can be pluralized" (ibid.), as in the following example:

(48) Le cavallo e le asine le quales
the horse and the donkey the which
non essava sellate ...
were not saddled (ibid.)

3.2.2. Formal (Non-)Identity and (Non-)Relatedness of Relatives and Interrogative Pronouns

In a large number of ALs at least some forms can be either RelPs or interrogative pronouns. This is true of American Speech, Anti-Volapük, Dutton Speedwords, Esperanto, Eurolengo, INTAL, Novilatin, pan-kel, SPL, Universal-Latein, and Vela, among others.

Such languages can differ in the degree of overlap: all forms can be used as either relative or interrogative pronouns, or only one or more of them can be. For example, Esperanto has a complete overlap of forms, while pan-kel seems to have a partial overlap. The non-completeness of the overlap in pan-kel is due to fact that there are different interrogative human and non-human pronoun forms, while there is a single RelP. Such discrepancies in agreement/inflectional marking between the two categories may be the cause of a large proportion of partial overlaps. A language in which they are not is SPL, which has an almost complete overlap, both types of pronouns being marked for case number and gender. The RelP "has exactly the same forms as the Interrogative Pronouns ... except the Singular Nominative" (Dominicus 1982:54). The nominative singular forms of the RelP are *qui*, *quae*, and *quod*, for masculine, feminine, and neuter respectively, while those of the interrogative pronoun are *quis* for masculine and feminine and *quid* for the neuter.[19]

When relative and interrogative forms are not formally identical there are two possibilities: either they can be related in form or they can be quite different, and when they are related in form, there are again two possibilities: one type of word can be derived from another, or not. Interestingly, while we have found some languages in which some interrogatives are built from relatives, we know of no examples in which the derivation goes in the other direction. The Panamane pronoun *velchy* 'which' can be a relative or an interrogative (and there is another RelP meaning 'which', *lekel*), but the interrogative pronouns meaning 'who', *kih*, and *kis'keh*, seem to be derived from the RelP *ki* 'who', and the interrogative pronouns *keh* and *kes'keh* apparently are built from the RelP *ke* 'that'.

[19] Dominicus (ibid.) does not state it, but one of his examples sentences (ibid.) shows that the neuter accusative singular forms also differ, which is what one would expect, given the identity of form of neuter nominative and accusative forms. Although *qui* is not used as an interrogative pronoun, it is used as an interrogative adjective in SPL.

In American some interrogatives can be seen as derived from relatives, as the latter begin with <c->, while some of the former begin with <cw->, e.g., the masculine gender RelP *ce*, and its interrogative counterpart *cwe*. However, other interrogatives do not start with <c->, but with <hw->, <tw->, or just <w->. The interrogative pronouns of Pantos-Dimou-Glossa seem to be based on, or at least to contain, its RelPs. For example, the nominative masculine nominative singular RelP is *ke*, while a corresponding interrogative pronoun is *ke-ly*. In Balta two interrogatives are derived from RelPs: *li-oka* 'qui, quel',[20] and *li-okea* 'que, quoi' from *oka* 'qui' and *okea* 'quoi'. In Zahlensprache the same marker is attached to relatives to make them into interrogatives, although without a hyphen, e.g., *li ki* 'qui' from the RelP *ki* 'qui'. This Zahlensprache marker was taken from Volupük, and I assume the same is true of Balta's *li-*, though in Volupük not all interrogatives contain *li* (which is also attached to words to form *yes-no* questions).

Volupük relatives and interrogatives both start with or contain [k], but at least in most cases one could not say that one has been derived from another, e.g., the basic RelP is *kel*, while *kim* is the interrogative 'who'.[21] Similarly in Mundelingva both relatives and interrogatives are formally similar in that both have [k] as their initial segment, but they are different, e.g., the relative *kelo* 'qui' and the interrogative *kiso* 'qui'. The same general situation holds in Weltsprache (Volk and Fuchs): both relative and interrogative pronouns start with *v*, which is followed by *e* in the RelPs, and *i* in the interrogatives. Thus the masculine/feminine RelP is *vel*, while its

[20] Given the mismatch of English and French relative and interrogative forms (e.g., that *que* can be the human or nonhuman accusative RelP, while English has distinct forms for these), we have retained French glosses which our sources give, rather than giving translations of them.

[21] At least two Volupük relative adverbs and their corresponding interrogative adverbs share some phonetic material, e.g., the relative *kelöp* 'where' and the interrogative *kiöp* 'where'.

interrogative counterpart is *vil*.[22]

In a minority of the ALs which we have looked at relatives and interrogatives cannot be seen as particularly related in form (i.e., more similar to one another than to other kinds of words), and as far as we know, with one exception, all such ALs are *a priori* languages. For example, the basic Babm RelP is *lr* 'some (object) which' while its basic interrogative pronoun is *qw* 'what' (SG). To our knowledge, the only non-*a priori* language in which this situation holds is Eurasto.[23]

3.2.3. Formal (Non-)Identity of RelPs and Demonstratives and/or Definite Articles

We do not know of any ALs in which RelPs are formally identical to definite articles, but see the next section for an AL in which a RelP may contain a definite article. The American Speech word *THCT*, like *that* in English, can be a demonstrative pronoun, a demonstrative adjective, or a RelP (as well as a conjunction).

3.2.4. Formal (Non-)Identity of RelPs and Personal Pronouns

We know of no examples from ALs in which RelPs and personal pronouns are formally identical, and if we did find such a case, we might find that the words in question were never RelPs, but simply always personal pronouns, i.e., that the language did not have relative clauses in the strict sense.

[22] Interestingly in this language relative and interrogative adverbs, unlike the pronouns, have the same form, e.g., *vo* 'where' is both a relative and an interrogative

[23] Suma has the interrogative pronouns *pavo* 'who' and *pivo* 'what' while its RelP, as already mentioned, is *sio*, so these are different in form. However, according to Russell (1966:4), "The interrogative words may be used as relative pronouns or conjunctions similar to English", although the example he then gives seems to involve an indirect question rather than a relative clause, casting doubt on whether he really means what he seems to be saying.

In Uropa what seems to be equivalent to a RelP may be derived from a personal pronoun and a suffix which sometimes functions as an article, though the facts are not clear to us and it is far from certain that the word in question is indeed a RelP. The word *y* means 'a certain, a particular', and *sa* occurs in some sentences with the meaning 'it' (and once apparently with the meaning 'he'), although Donisthorpe (1913:3) states that there are "only three" personal pronouns in Uropa, and does not give *sa* as one of them. When *y* is used as a suffix it sometimes adds a meaning of definiteness to a phrase, e.g., *ruby doma* 'the red house' (ibid.:9) (c.f., *rub doma* 'a red house' (ibid.)), and sometimes it creates a possessive from a noun or pronoun, e.g., *may* 'my' (from *ma* 'I'). Donisthorpe (ibid.:5) says, "What is called the possessive 'case' of a noun is not really a noun at all, but a demonstrative or Pointer singling one particular thing out of a crowd. [...] There is very little difference between 'a mother's love' and 'maternal love'." This may mean that Donisthorpe views *-y* as, at least in some situations, an adjective-forming suffix.

In Uropa relative clauses begin with *y sa* (or *say*, or with *y sam/samy* if the RelP is the object of the relative clause); Donisthorpe (ibid.:7) says, "Sometimes a whole sentence is used as an adjective, a particularizing or demonstrative adjective, and for this reason ... we add the word *y* to the sentence." He gives (ibid.) what we might call the underlying structure of a sentence containing a relative clause, *Ha e Jaka-kondi-sam-y doma* 'this is the Jack-built-it house', in which apparently *-y* marks the whole sentence as adjectival.[24] The surface form of this sentence is *Ha e doma y sam/samy Jaka kondi* 'This is the house which Jack built' (ibid.). One might think then that *say* is not a RelP but simply the combination of a

[24] The derivation of sentences in which the subject of the relative clause was coreferential with something in the main clause might be more complex, since *sa* and *y* would not be adjacent in the underlying structure.

personal pronoun and an adjectival suffix, which in these instances marks sentences as relative clauses, i.e., as functioning as adjectives. However, consider the sentence below.

(49) Oma say oc dane dwen dane
man say quickly gives twice gives
'He who gives quickly gives twice.' (ibid.:6)

Here we could not say that *say* is simply a combination of the personal pronoun of the relative clause and *-y*, since if the relative clause had been a simple main clause, presumably *la* 'he' would have occurred, and we should then expect *lay* to be in the place of *say* here. *Lay* is a word of Uropa, but it means his (parallel to *may*, given above); as far as we know it does not occur as the first element of relative clauses. One might then argue, assuming that this sentence does not represent an error, that *say* has progressed beyond being a simple combination of a personal pronoun and *-y*, and has become a (more general) RelP. However, we would hesitate to assert that it is indeed a RelP without further evidence.

3.2.5. Formal (Non-)Identity of RelPs and Conjunctions

In some ALs there is a RelP which is formally identical to a conjunction, namely the conjunction which introduces an indirect statement, equivalent to the English conjunction *that* or the conjunction used in comparatives, corresponding to English *than*. We shall not consider cases where the RelP in question is also formally identical to an interrogative pronoun; in such cases this cannot be said to be a fact about the RelP alone. For example in Universal-Sprache, *ke* is the (singular) RelP and the (singular) interrogative pronoun, as well as being a conjunction at least partly equivalent to French 'que'. In Esperanto the conjunction in question begins with the same segment as the relatives/interrogatives, but is different, *ke*, and there are

probably a good number of other ALs with this pattern.

The languages in which a RelP has the same form as the conjunction meaning 'that' but in which there is no interrogative pronoun with this form may all have the same pattern as the English word *that*. The Dutton Speedwords word *k* 'that' is not cognate with that, but follows this pattern, i.e., it can function as a RelP or a conjunction, but not as an interrogative. The same general holds in Hom-Idyomo, although described in a different way by Cárdenas (1923:39): "... ***ke***, which is a conjunction, but which is sometimes used as a relative pronoun replacing *kio* when preceded by another pronoun having the same ending. Thus *tio ke mi me diray te*,[25] 'this which I tell you', instead of *tio kio mi diray*. However, this substitution is optional, not necessary." In fact, Cárdenas uses *ke* as a RelP when it does not immediately follow a word "having the same ending". The same sort of situation also holds in the Blue Language, where, as we have already said, *ko*, which is basically a conjunction, can act as a RelP.

The American Speech word *THCT* can be a RelP or a conjunction, as well as a demonstrative, as we have already mentioned. That is, it acts in the same way (in this respect) as *that*, not surprisingly.

The RelP of Europal is *as*.[26] The interrogative pronouns are different in form, (and the equivalent of the English conjunction *that*, *dat* is also different). However, *as* is also used in comparative type constructions, and would also be used to translate the second occurrence of *as* in phrases of the type *as big as*.

[25] We are using the symbols <ɾ> here for a letter of Hom-Idyomo which we cannot reproduce here; it seems to be a script letter <ɾ>; the alphabet of this language also has a "regular" letter <r>.

[26] Weisbart (1912:7) calls it a reflexive pronoun, but this is apparently an error, given how it is used in texts in this work, and at another point he does call it a relative pronoun.

3.3. The Possibility of Omitting Relative Pronouns

We have seen that some natural languages allow RelPs to be dropped. Gode and Blair (1951: 28) explicitly say that one cannot omit a RelP in Interlingua: "In contrast to English usage there are no relative constructions without a relative pronoun." One of the examples which they then give is the Interlingua translation of 'The onions you ate smell to high heaven': *Le cibollas que tu ha mangiate odora al alte celo*. The *PMEG* (in the section "Rilata *kiu*") says, "**Do not leave out the relative *kiu***. In some languages this is possible, but not in Esperanto."

At least one language, Latinulus, is explicitly said to allow relative clauses without a RelP, under certain conditions: 1) "The relative pronoun can be omitted when it is the object complement of the verb which follows it" (Martellotta 1909:102), 2) "The omission can also be made when the relative follows a superlative" (ibid.). Examples provided by Martellotta to illustrate these conditions are given in (50a) and (50b) respectively.

(50) a. pueris-f go ama-bam
boy-F I love-IMPERF
'The girl I loved' (Martellotta 1919:102, OT)

b. los-f esseo maximam pulchra pueris-e-f go
they-F are most beautiful boy-PL-F I
umquam vide-vi
ever see-PERF
'They are the most beautiful girls I have ever seen.' (ibid., OT)

(50b) is the only example given for the second condition, but it is not a good example for it, since it would also fall under the first condition. There is another situation in which the RelP need not be present, namely when it would have been the object of a preposition:

(51) bebis-f, tu-f eme-vi ica tois-e per,
baby-F you-F buy-PERF this toy-PL for
esseo mori-tum
is die-PSTPART
'The baby girl you bought the toy for has died.' (ibid.:103, OT)

Martellotta also states two circumstances in which an (overt) RelP must not be dropped, one of which is when there are one or more words intervening between the antecedent and the RelP. It seems rare for ALs to allow relative clauses without overt relative pronouns; the fact that Latinulus does permit it may well be due to the fact that its syntax, at least to a large extent, may be based on that of English.

Apparently in Dutton Speedwords relative clauses do not (always) require overt RelPs:

(52) C e l uv iku fy w h ce ...
this is the first satisfactory reason we have receive
'This is the first satisfactory reason we have received ...'
(Dutton 1951:121, translation p. 44)

This is also presumably due to the influence of English. We have also seen a relative clause in American without a RelP, although in such cases one may wonder whether it is due to an error, rather than something permitted in the language.

There is at least one AL, Interglossa, in which it is not only permissible, but required, for a relative clause to lack a RelP under some conditions. Its RelP is *su*, which can be the subject, but not a verbal or prepositional object in a relative clause. Hogben (1943:35) says, 'When the relative pronoun is *not* the subject, no equivalent takes its place. We proceed precisely as in conversational English.' One of the two examples he gives is below.

(53) Un anthropi; mi pre vise; non habe bio
SG man I before see not have life
'The man I saw is dead.'(ibid.)

3.4. Inflectional Marking

ALs differ widely in the inflectional marking which they allow or require on RelPs. Of course this is often or usually related to the degree of inflectional marking in the language in general, or on pronouns in general, although there are some interesting exceptions. (Of course different relative pronouns in the same language could differ in the degree of marking they can/must have.) Morphological marking is one possible way of eliminating or reducing ambiguity with respect to the ant NP of a RelP, a matter of concern to some AL creators. We include as inflectional marking the existence of different forms for human and non-human, or animate and inanimate antecedents (which will often be covered by natural gender agreement, if this is marked on the RelPs). We are focusing on RelPs which are different in form from interrogative pronouns, or where the two kinds of pronouns may be the same in form, but differ in possibilities for inflection, i.e., where we can make statements about RelPs which do not necessarily apply to interrogatives in the language in question.

Some ALs have a RelP which has no inflectional marking of any kind. The pan-kel RelP *ki* seems to be of this type: Wald (1909:6) glosses it as 'welcher, welche, welches (der[,] die, das)', i.e., it is used for all three genders, and in two pan-kel sentences (ibid.) it has the same form when it is a subject and when it is a verbal object.[27]

[27] Couturat and Leau (1907/1979:20) state, "The interrogative-relative pronoun is: ki = qui, ka = quoi", i.e., that different pronouns are used for humans and non-humans, but they may be confused here. Wald (1909:6) does give different interrogative pronouns for humans and non-humans, *ki* for the former and *ka* for the latter, and the fact that *ki* is both an interrogative and a relative pronoun might

Likewise Anti-Volapük's RelP, *ke*, does not change for any inflectional category (although, as in pan-kel, there are different interrogative pronouns depending on humanness, and one of these forms is homonymous with the RelP).[28]

The RelP of American is marked for several inflectional categories. In fact it consists of nothing more than <c>/[k] and these inflectional markings: O'Conner (1917:27) says, "The relative pronoun is formed by placing after c the letters indicating genero, case and number, in the order named." (Animacy is included in gender, since American has five (natural) genders, masculine, feminine, common, immaterial inanimate, and material inanimate.) Thus, for example, *coṅs* is the dative plural immaterial form of the relative pronouns. However, all the inflectional marking seems to be optional. O'Connor (ibid.:28) explicitly states this with respect to personal pronouns: "It is usually unnecessary to write the personal pronoun in full, especially when the noun to which it relates is in the nominative case ... The omitted letters are indicated by an apostrophe", and the form *c'* occurs twice in his book, one instance of which is given below:

(54) Ult ı modo c' ı posido translatıo-ṅ.
last is form which is placed translation-DAT
'The last is the form which is placed in the translation.'
(ibid.:46)

have misled Couturat and Leau into thinking that *ka* also had both uses. It is an interesting point that the pan-kel interrogative distinguishes humanness while the relative does not. The discrepancy may also be due to the fact that Couturat and Leau's source for pan-kel is the second edition of Wald's book, while ours is the fourth edition; Wald apparently made some changes in his language, and perhaps this is one of the areas where a change was made.

[28] In this and the following two sections we are only considering RelPs which are not indefinite RelPs (or sequences equivalent to them). As we shall see below, in Anti-Volapük, gender is shown in free relative clauses, since the article, which is marked for gender, together with the RelP, are at the beginning of such clauses.

3.4.1. Agreement Marking (Number, Natural Gender, Humanness, Animacy)

The RelP of Pantos-Dimou-Glossa is marked for natural gender (and through it humanness or animacy) and number. The nominative singular masculine, feminine, and neuter forms of it are *ke*, *ka*, and *ko* respectively, and the corresponding plural forms are *keci*, *kaci*, and *koci*.

aUI has separate forms for human, non-human animate and inanimate ant NPs, *xu* 'who', *xo* 'which', and *xE* 'what' respectively.[29] Both of these can take the plural affix, *-n-*, e.g., *xnu*, and *xu* can also take the feminine marker, *-yv-*, yielding *xyvu*. In fact it can also take the masculine marker, *-v-*, yielding *xvu*, although Weilgart (1979: 49) uses *xu* in a sentence when a male antecedent is invovled, so the use of *xvu* rather than *xu* seems to be optional, perhaps being used when the maleness of the antecedent is to be emphasized, or when it is not obvious from the expression of the antecedent.

In Oz, distinguishing both gender and humanness/animacy in the RelP is optional: "**han** is the usual relative pronoun" (Elam 1932:18), used with both human and inanimate antecedents, "unless it is desired to call attention to the sex by means of the relative, in which case **h** may be used with any of the gender roots" (ibid.). These "gender roots" are *ap*, *at*, *ak*, *af*, *an*, and *as* for masculine, feminine, common, neuter, indefinite, and abstract, respectively, and at least some of them can occur alone as personal pronouns. It can be seen that the "usual relative pronoun" contains the indefinite root.

The situation in Arulo is somewhat similar to that of Oz: Talmey (1925:9) states, "**Il el, ol** [masculine, feminine, and neuter 3rd person singular personal pronouns] may be attached to the relative pro-

[29] However, in Weilgart (1979:185) *xnu* is used when the ant NP refers to animals. We assume that *xo* is a pronoun; Weilgart (ibid.:237) glosses it as 'which (of animals)'.

nouns to avoid ambiguity; **la filiino di mea amikulo, elquan vu vidis hiere en mea domo, recevis un premio,** the daughter of my friend whom you saw yesterday in my house received a prize." Among the relative-interrogative pronouns of Arulo is *quo*; it "refers to a fact or to an indefinable thing" (ibid.:8). In the example below, it is used for an antecedent which is a clause, not an NP.[30]

(55)	el	departis	nerekupereble,	quo	dezolis
	she	has.departed	irrevocably	which	has.desolated
	mea	kordio			
	my	heart (ibid.)			

In contrast, Russell (1966:4) states that in Suma "[t]he relative pronoun cannot be used to refer to a clause". In the following example, tu 'this, it' appears instead of a RelP:

(56)	mu	te	soka.	tu	baki	mu	to	lusi
	he	be	hungry	this	make	he	to	steal
	'He was hungry, which caused him to steal.'							
	(ibid., including the word for word gloss)							

As far as we can determine, in Communicationssprache the RelP, *wia*, takes number marking but not gender/humanness/animacy marking.[31] The RelP of Mundolinco, *ci*, does not seem to distinguish humanness/animacy, although its interrogative, which is built from it does; in fact, the interrogative pronoun forms consist of *ci* and *homo*, *bestio*, or *resso*, e.g., *cibestio* means something like 'which animal', and *ciresso*, 'what'.

[30] It is not only used for clauses, consider the following example, which might make one wonder what Talmey meant by "an indefinable thing": *Quon el trovis* 'What did she find?".

[31] Our source for this language, Couturat and Leau (1903/1979) states (p. 243) that it is "invariable in gender, but declinable like a noun".

3.4.2. Case Marking

The discussion on case marking here relates to case assigned within the relative clause; to our knowledge in no AL does a RelP receive case marking to agree with the case of its ant NP, as can happen in through case attraction in Ancient Greek (see (19)).

There are some ALs which have possessive relative words, equivalent to English *whose*. It is not clear to us whether these should be treated as forms bearing genitive case, or as adjectives (since it is difficult or impossible to perform the kinds of syntactic tests which would distinguish the possessive adjective *my* from the genitive pronoun *mine*), and so it is not obvious whether these should be treated as case-marked forms. For example, this is true of the aUI word *xum* 'whose'. On the other hand, in Suma there is no possessive form of a RelP; in order to express possession by the antecedent one must use a preposition:

(57)	pamo,	dea	sio	taso	te	tila
	man	of	RelP	son	is	sick
	'the man whose son is sick' (Russell 1966:4)					

Given the problem of determining the nature of forms with the meaning 'whose', we shall not count a language which has only such a form, and not a distinct accusative RelP, as having case marking on its RelP(s).

The RelP of Pantos-Dimou-Glossa has separate forms for nominative and accusative, the latter being marked by the suffix *-m*, e.g., the masculine singular nominative form is *ke*, while its accusative counterpart is *kem*. Pronouns in general in this language only seem to morphologically distinguish two cases; with noun phrases headed by nouns case is marked on the article, and five cases are distinguished.

The RelP *ra* of the Blue language is inflected for case, but its

paradigm involves a different inventory of cases than the personal pronouns of the language. Only three case forms are distinguished, the nominative-accusative, *ra* (and *re* in the plural), the genitive-ablative, *era*, and the dative, *ara*, while the personal pronouns have, in addition to genitive-ablative and dative forms, distinct nominative and accusative forms, and what is called a vocative form, although its functions may be different from what one generally thinks of as a vocative.[32]

The RelP of Communicationssprache, *wia*, distinguishes five cases: nominative, genitive, dative, accusative, and ablative, while that of Mundelingva, *kelo,* distinguishes all of these except the ablative, as do the interrogative pronouns of the language.

3.4.3. Marking for Restrictive vs. Non-Restrictive Clauses

We know of no AL in which the fact that a relative clause is restrictive or non-restrictive is consistently obligatorily marked on the RelP in that clause, or completely determines the choice of RelP.[33] In American *THCT* is restricted to restrictive clauses (if this is why Foulk 1937 means by 'NXSXSI RELXETI KLAZ') while *HU* and *HWITSH* can occur in either type of relative clause, but this is what one would expect, since this is how the English versions of these words behave.[34]

3.5. Free Relative Clauses

There are three basic possibilities for RelPs and free relative

[32] The interrogative pronouns of the Blue Language, although different in form, also seem to have only a single form for the nominative and accusative.

[33] In Lojban there are different relative clause markers for restrictive and non-restrictive clauses, but the pronoun itself is the same.

[34] However, in at least one other respect the behaviour of *THCT* does not seem to be identical to that of its (standard) English counterpart.

clauses in language: 1) the language does not allow them, 2) the language allows them and uses the same set of RelPs in them as in other, "non-free" (i.e., "headed"), relative clauses, and 3) the language allows them, but the set of RelPs used in the two types of relative clauses is not the same.[35] In situation 3) the set could partially overlap or be completely different, and here there are various possibilities. Some RelPs might be restricted to free relative clauses, or to non-free ones, RelPs of the latter type often being called indefinite relative pronouns, and some might be able to occur in either type; the inventory of RelPs in an AL could consist of any combination of these. Explicit information about this is rarely given in works on ALs, and one often has to look at example sentences and texts to get an idea of what is possible. However, if a RelP is glossed as 'whoever' or 'whatever', one could probably take it to be an indefinite RelP, and thus limited to occurring in non-free relative clauses.

We know of no ALs in which free relative clauses are explictly not permitted. (However, if one considers sequences such as the English *he who* or French *celui qui* to be not indefinite RelPs, but sequences involving personal pronouns or demonstratives followed by a RelP, with the personal pronoun or demonstrative not being in the relative clause, then there could be ALs without free relative clauses, if these are the only kinds of supposedly indefinite RelPs in the language.) Languages in which possibility 2) apparently holds

[35] Free relative clauses must be distinguished from clauses in which the relative clause has been moved in front of the main clause, as in the following example:

> Kiu-n malĝojo ne turment-is, tiu ĝojo-n ne sent-as.
> who-ACC sadness not torment-PST that.one joy-ACC not feel-PRES
> 'Whom sadness did not torment, he does not feel joy.' (*PMEG*:Rilata *kiu*, OT)

From looking at the relative clause itself one cannot tell whether it is headless, but if one looks at the main clause one of this sentence will see an antecedent. The *PMEG* gives *Tiu, kiun malĝojon ne turmentis, ne sentas ĝojon* as the equivalent of this sentence , and we might say that this represents the basic form of the sentence, before a (stylistic?) movement operation has taken place.

include American, that is, there are no special RelPs for free relative clauses, and the "regular" relative can be used in them.

Languages which have indefinite RelPs probably represent possibility 3). In some ALs a demonstrative is followed by a RelP, as in French, e.g., Arulo *ta, qua* 'he, the one who'. In such constructions, it is not clear that we are actually dealing with a free relative, since the first element is arguably a part of the main clause, and thus is the ant NP of the RelP. In fact, this appears to be so for Arulo, given the following sentence:[36]

(58) La historio laudas ti, qui bonfacis al homaro
The history praises those, who do.good to humanity
(Talmey 1925:10, OT)

In Anti-Volapük the equivalent constructions involve an article followed by a RelP, e.g., *lo ke* 'he who', *la ke* 'she who'.

In some ALs it may be clearer that indefinite RelPs are involved. Babm forms words which seem to be indefinite RelPs with the suffix *-u*, e.g., *lru* 'whatever'; the same suffix is also used with interrogatives, e.g., *qwu* 'whatever'. Panamane has some "compound" (Amador 1936:17) relatives, e.g., *kieby* 'whoever', *keseby* 'whatever', which, we assume, would be used in free relative clauses.

In Esperanto some apparent indefinite RelPs involve *ajn*, preceded by a RelP, e.g., *kio ajn* 'whatever'. The *ajn* is orthograpically separate "[i]n order to avoid confusion with the accusative plural ending" (Kellerman 1910:181-2). It is not clear that forms with *ajn* are always free relative clauses; consider the following sentence, in which *kiu ajn* could not be translated by *she who*.

[36] In any case, apparently it is posssible to have headless relative clauses in Arulo with a simple RelP, e.g.,:

Quu serchas danjero, perisas en ol
who seeks danger perishes in it (Talmey 1925:10, OT)

(59) Kiu ajn ŝi est-os, mi dezir-as al ŝi feliĉon!
who ever she be-FUT I desire-PRES to she happiness
'Whoever she will be, I desire happiness for her!' (*PMEG*: *Ajn*, OT)

In any case, *ajn* is not required in free relative clauses: the *PMEG* has a section entitled "Rilataj subfrazoj – forlaso de TI-vorto" ("Relative subordinate clauses – omission of TI-word"), in which it is stated, "There do not exist absolute rules, when one can leave out a TI-word [i.e., correlative demonstrative]. Most serious is the clarity: If the sentence will become too unclear, one should not leave out the TI-word." Among the examples given there is the following:

(60) Kiu hav-as forto-n, havas rajton
who have-PRES strength-ACC have-PRES right-ACC
'He who has strength has the right.' (OT)

This sentence is equivalent to *Tiu,* kiu havas forton, *havas rajton* (ib id.), 'He who has ...'

3.6. Restrictions on Ant NPs and Marking for Different Ant NPs

In some sentences there is more than one possible ant NP for a RelP, leading to ambiguity. One way of diminishing the potential for ambiguity is to have agreement marking for number and/or humanness/gender/animacy on RelPs. Novilatin has such agreement, but it also has another means of reducing ambiguity. What we might call the basic RelP is *qve*. This can only be used to refer to the closest possible ant NP (i.e., the first one which is found when counting backwards from the RelP). Otherwise a different relP, *ilqve*, must be used, as shown in the following example:

(61) I kruc-igne dei batteries de Hattoy e dei penisle
the cross-fire of.the batteries of Hattoy and of.the peninsula
Iges, ilqve dupla-b se en violencie ...
Iges which double-PST self in violence
'The cross-fire of the batteries of Hattoy and of the peninsula Iges, which were doubling in violence ...' (Beerman 1907:60, OT from the French)

Here it is the cross-fire, and not Igne (or Hattoy or, we believe, the batteries), which were doubling in violence. One might note that the way the nearest antecedent is counted seems to be to take the first smallest NP one finds; in some sense 'the cross-fire ...' is as close as 'the peninsula Igne', since 'the cross-fire ...' contains 'the peninsula Igne', i.e., the left edge of both NPs ends in the same place.

Suma has a different solution the problem, that of not using a RelP in such contexts. Russell (1966:4) says:

> If there is ambiguity, repeat the antecedent in place of *sio*:
> Ambiguous: *papo dea kato, sio te tila* father of the boy who is sick? the father or the boy?)
> Unambiguous: *papo dea kato, tu kato te tila* (father of the boy, this boy is sick);
> *papo dea kato, tu papo te tila* (father of the boy, this father is sick)

3.7. Prenominal/Postnominal and External/Internal Relative Clauses

We have never seen a prenominal relative clause in any ALs, nor have we seen any internal relative clauses in them. We exclude here sentences of the sort discussed in note 35, which seem to represent a stylistic (and marked) option, and not the basic order of the language.

3.8. The Accessibility Hierarchy

To our knowledge, no work describing an AL discusses the accessibility hierarchy in reference to that language.[37] Therefore we cannot describe limitations on relativization, we can only state what we have found to be relativized in different languages. However, if a RelP in a language distinguishes several case forms, as does American, this is evidence that the language allows relativization at least fairly far down the hierarchy, even if we see no examples of it.

In Esperanto objects of prepositions can be RelPs, e.g.,:

(62) Mi memoras tiun aferon, pri kiun vi parolas
I remember that matter about which you speak
(Kellerman 1901:106)

The same is true of Hom-Idyomo and Interlingua, and presumably many other ALs. We do not recall ever having seen a RelP as an object of comparison.

3.9. Wh-Movement

In aUI and pan-kel, wh-movement is not required, as shown in the following pan-kel example.

(63) son, fat gasto ki
son, father visited who
'the son whom the father visited' (Wald 1909:6)

However, this is a fact about wh-movement in general in these languages, i.e., it applies to both interrogative and relative clauses. To

[37] The situation in Interglossa does not involve the accessibility hierarchy, but rather what can be represented by an overt RelP, which is a different question.

our knowledge there is no AL in which wh-movement is required in one type of clause but not the other.

3.10. Resumptive Pronouns

Resumptive pronouns are not a common feature of ALs; we have noticed only two occurrences of them. However, both of these (which occur in translations of the "Lord's Prayer" are unusual in that it is a subject pronoun which is retained. Below is the occurrence in Carpophorphilus' Language.

(64) O baderus noderus, ki du esso in seluma ...
oh father our who thou art in heaven
'Our father, who art in heaven ...' (Couturat & Leau 1979:24)

One might think that the reason for the presence of the resumptive pronoun that the ant NP is meant to refer to the 2nd person, but is not a 2nd person pronoun; the presence of the personal pronoun in the relative clause makes it clearer that the subject of the clause is 2nd person.

4. Recommendations

There are various aspects of the form and function of RelPs in relation to which the AL constructor will have to make decisions. We shall offer recommendations on some of these, but it should be borne in mind that sometimes the language constructor will be dealing with implicationally related properties and, therefore, decisions made for other grammatical subsystems may limit or even predetermine the choices within the subsystem of RelPs. Further, what may be best for one type of language, e.g., a pan-Romance language, may not be at all appropriate for another type, e.g., a pan-Altaic or *a*

priori language. Our recommendations to some extent should be seen as recommendations for the most common and successful type of AL, an *a posteriori* language using largely elements of several major Western European languages.

Also, there is a range of fundamental considerations which the AL constructor must take into account whatever the specific linguistic structure he is involved in designing may be. We addressed some of these in our paper on the form and function of reflexive pronouns (Libert & Moskovsky 2002), particularly the inherent conflict between clarity and simplicity (of form). Another closely related issue is the issue of learnability, which (in view of the fact that an AL would almost always be learnt as a second language) should be of central concern for the AL constructor. Simple forms would be easier to learn and use, but being underspecified they would also be a potential source for ambiguity. Providing an appropriate level of specification on the structure in question can eliminate even the remotest possibility for ambiguity, but this would inevitably lead to problems in learning and using these forms. Thus one crucial question is about the (un)acceptable level of ambiguity in an AL: a question which (as we suggested in our paper on reflexives) may be extremely hard (or impossible) to answer on a principled basis, and the view we take in this regard is that the language constructor will have to deal with this issue on a case-to-case basis.

Of course, the first decision that an AL constructor will have to make is whether for an AL it is necessary to have RelPs (and, more generally, relative clauses) at all. RelPs, and relative clauses, exist in many, but arguably not all of the world's languages. Some languages, such as Turkish, have, at least in native constructions, nothing precisely analogous to the relative clauses of e.g., English. Rather, participial clauses are used, as shown in (9). In fact, neither relative clauses strictly speaking nor such equivalent constructions are necessary to fully express the meanings which they convey; for example the sentence *The man whom I saw is a teacher* is truth-

conditionally identical to *I saw a man. He is a teacher* (assuming that *He* is corefential with *a man*). Nevertheless, the fact that all natural languages, and to our knowledge all ALs, have at least equivalents of relative clauses is unlikely to be coincidental, and it seems that relative clauses (and their equivalents) provide languages with flexibility and expressiveness which they may not have without them. We, therefore, believe that an AL should have relative clauses in the strict sense, and RelPs.

If an AL is to have relative clauses and RelPs, this would inevitably predetermine certain aspects of the function that RelPs will perform: minimally, the RelP will have to provide reference to an ant NP in the main clause. It could also function as a link between the relative clause it is a part of and the main clause. We believe that while only the anaphoric function is essential, RelPs in Als could easily combine the two, as is the case in so many natural languages, and as these two functions are not contradictory or mutually exclusive, having the RelP perform both will be desirable in terms of economy. Of course different design solutions involving different function of the RelP are possible. In the rest of this section, we touch upon some of these possibilities.

As regards the form of RelPs in ALs, the language constructor will have to make decisions in relation to the following: Will the RelP in this AL be a simplex featureless form or will it be a complex form specified for a range of features (such as phi-features), morphological case, etc.? Will the RelP be identical with, or similar to, other elements in that AL (e.g., interrogatives, demonstrative, conjunctions)? Will there be more than one type of RelP (as, e.g., is the case in English, where we have wh-relatives and *that*)? Will zero RelP forms be allowed in this AL? Will the AL under construction use relative adjectives and, if yes, will they have same form as the RelP?

In addition to these questions about the form, the AL designer will also have to make some syntactic decisions: e.g., whether to use

pronoun retention to mark the relativized position in the relative clause, which positions in the relative clause can be relativized, what position in the relative clause will be occupied by the RelP, etc.

The question of feature specification and complexity of the form of the RelP is of great importance as it bears direct relevance to what we said earlier about the conflict between clarity and simplicity, and the related learnability issue. English RelPs are, e.g., specified for the feature [HUMAN]: the RelP *who* is [+HUMAN], while the RelP *which* is [-HUMAN]. The presence of this feature might be instrumental in removing potential ambiguity in some sentences, e.g.:

(65) I read a book by a former teacher of mine [which/who] was awful.

Clearly the feature specification on the RelP here helps eliminate potential ambiguity, which an unspecified form cannot do. An AL constructor might thus need to decide whether the risk of ambiguity in sentences like (65) warrants including that feature in the specification of the RelP of the AL they are designing. The English RelP is not specified for phi-features (person, number, gender) but RelPs in many languages are, and such specification can also play a role in minimizing the level of potential ambiguity. For instance, in a Bulgarian sentence analogous to (65):

(66) Pročetoh kniga ot edin moj predishen uchitel
I-read book from one my former teacher
[kojato /kojto beshe uzhasna /uzhasen
which-F which-MASC was awful-F awful-MASC

the form of the RelP will be feminine if the intended reference is the book or masculine if the intended reference is the teacher. Thus the same question holds here as well: is such level of specification essential for the RelPs in an AL? The position that we take in relation

to this is that it is not. Our reasons are as follows: 1) ambiguity of the type illustrated in (65)–(66) would arise relatively infrequently, and, even though in these examples it clearly exists on a sentence-level analysis, it is very likely that in the vast majority of cases the larger context (i.e., the preceding or following discourse) will reduce or completely remove any existing sentence-level ambiguity; and 2) the potential of feature specification on the RelP to disambiguate is limited — e.g., in the two sentences above, the [HUMAN] or gender feature would have been of no value if the potential referents for RelP in the main clause were both [+HUMAN] or masculine. We would, therefore, recommend that the form of the RelP in an AL be a completely unspecified simplex form. We suggest that it start with [k], as do many RelPs of Indo-European languages, not only in Romance languages, but also in Slavic languages and Persian. Further, RelPs in some of the most successful existing ALs, such as Esperanto, have [k] as their initial segment, so people who have already learned an artificial language may well already associate this segment with relative words, even if they do not speak one of the natural languages just mentioned. We believe that the question of case marking for the RelP (a different, though related issue, bearing on the form of RelPs) may hang on decisions made for different grammatical domains, in particular the nominal system: if the nominal system of that AL is generally designed to have case inflections, then the same should probably hold for the pronouns (including RelPs) of that language.

We do not see it as a problem if the same forms are used for relatives and interrogatives: questions and relative clauses can generally be distinguished by intonation (or by punctuation in writing). There may be some small potential for confusion between indirect questions and relative clauses, but ambiguous cases will not arise very often, and it seems better to allow for a slight chance of ambiguity than to force to learner to memorize two sets of pronouns rather than one. Given that some relatives and interrogatives have the same

form in so many languages, this may appear to language learners as a natural feature of an AL. However, we do not advise that the same forms be used for RelPs and demonstrative pronouns, as happens in German, or for RelPs and subordinate conjunctions or complementizers (as happens in English and many other languages); also we do not recommend that an AL have two different types of RelPs (i.e., one wh-form and one conjunction-like form).

We would recommend against allowing RelPs to be dropped as *in The man I saw yesterday*. This is a feature which is not only rare among the world's languages, but also can unnecessarily put serious additional demands on the processing capacity of the AL user.

If an AL is to have relative adjectives, we would suggest that they have the same form as RelPs. The immediate context may often or always make it clear whether the relative word is functioning as a pronoun or an adjective.

As we indicated in the previous sections, not all natural languages relativize all possible syntactic positions in the relative clause. The language constructor must, therefore, decide whether the AL under construction will allow the full range of relativizable positions (as English does), and if not, which will be the permissible relativizable positions. We believe that, as in terms of learnability ALs are very much like second languages, such findings as there are in the field of second language acquisition would be relevant to how ALs are acquired. Second language acquisition literature (e.g., Gass 1979) indicates that more marked relative clauses (i.e., relative clauses lower on the Accessibility Hierarchy) do not necessarily pose serious learnability problems, and that the system of relative clauses in the L2 can be taught in a very efficient way (see the end of section 2). We can think of no plausible reasons why the full range of possible syntactic positions inside the relative clause should not be allowed.

In some natural languages, pronoun retention[38] is used to mark the relativized position inside the relative clause. Pronoun retention involves constructions like (67):

(67) the man$_i$ [that I saw him$_i$]

While we recognize that a combination of a relative conjunction and a personal pronoun in the relativized position is one possible design, we would nevertheless contend that having a single RelP performing the dual function of anaphor and conjunction is more economical and, therefore, the preferable option.

Finally a brief word about the position of the RelP inside the relative clause. Again this is a design decision which at least in part hangs on other design decisions within the relative clause domain and/or elsewhere. The anaphoric function of the RelP entails that it must be able to corefer with the ant NP. Morphological agreement between RelP and ant NP can be used to encode such coreference. Alternatively, if the RelP is designed as a simplex featureless form (which we recommend), then structural adjacency between RelP (or the phrase containing the RelP) and ant NP will be required to encode the coreference between them. Further, if the RelP is designed to function not only as an anaphor, but also as a conjunction, then obviously it will have to be positioned in clause initial or final position (depending on the basic word order of that language).

References

Amador. M. 1936. *Fundaments of Panamane*. Pueblo Nuevo, Panama: Imprenta Barcelona.

Beerman. E. 1907. *Die internationale Hilfsprache Novilatin*. Leipzig: Theodor Weicher.

Bianchi, V. 2000. The Raising Analysis of Relative Clauses: A Reply to Borsley. *Linguistic Inquiry* 31, 123-140.

Bollack, L. 1900. *Abridged Grammar of the Blue Language*. ("English Version by Professor Tischer"). Paris: Editions of the Blue Language.

Borsley, R. 1997. Relative Clauses and the Theory of Phrase Structure. *Linguistic Inquiry* 28, 629-647.

Cárdenas, C. 1923. *Hom-Idyomo* (2nd edition). Leipzig: Fischer & Wittig.

Collins, P. & Holo, C. 2000. *English Grammar: An Introduction*. London: MacMillan.

Comrie, B. 1989. *Language Universals and Linguistic Typology: Syntax and Morphology* (2nd edition). Oxford: Blackwell Publisher.

Couturat, L. & L. Leau. 1903, 1907/1979. *Histoire de la Langue Universelle and les Nouvelles Langues Internationale*. Hildesheim: Georg Olms Verlag.

Curme, G. 1935/1962. *A Grammar of the English Language*. Essex, CT: Verbatim.

Dominicus, R. 1982. *SPL*. Wisconsin: Dominicus Publishing House.

Donisthorpe, W. 1913. *Uropa*. Guildford: W. Stent & Sons.

Downing, B. 1978. Some Universals of Relative Clause Structure. In J. Greenberg (ed.), *Universals of Human Language* 4, 375-418. Stanford: Stanford University Press.

Dutton, R. 1951. *Teach Yourself Dutton Speedwords*. London: The English Universities Press.

Elam, C. 1932. *The Case for an A Priori Language*. Cincinnati, OH: The Open Sesame Press.

Foulk, R. 1937. *Americai Spek*. New York: J. J. Little & Ives.

Gass, S. 1979. Language Transfer and Universal Grammatical Relations. *Language Learning* 29, 327-344.

Gode, A. & H. Blair 1951. *Interlingua*. New York: Storm Publishers.

Hogben. L. 1943. *Interglossa*. Harmondsworth: Penguin Books.

Kayne, R. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Keenan, E. 1985. Relative Clauses. In T. Shopen (ed.), *Language Typology and Syntactic Description* 2, 141-170. Cambridge: Cambridge University Press.

Keenan, E. & B. Comrie. 1977. Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry* 8, 63-99.

Keenan, E. & B. Comrie. 1979. Noun Phrase Accessibility Revisited. *Language* 55, 649-664.

[38] In other words, the use of resumptive pronouns.

Kellerman, I. 1910. *A Complete Grammar of Esperanto*. New York: D. C. Heath.
Lewis, C. 1879. *A Latin Dictionary*. Oxford: Oxford University Press.
Lefebvre, C. 1979. Réanalyse de Que/qui, Inversion Stylistique et Mouvement de WH en Français. *Montreal Working Papers in Linguistics* 13, 73-90.
Libert, A. & C. Moskovsky. 2002. On the Form and Function of Reflexives in Artificial Languages. *Journal of Universal Language* 3, 97-132.
Martellotta, V. 1909. *Latinulus*. Bari: F. Casini & Figlio.
Morneau, R. 1994. Designing an Artificial Language: Syntax. Available at URL <http://www.eskimo.com/~ram/syntax.html>.
O'Connor, C. 1917. *American: The New Pan-American Language*. Buffalo: Hausauer-Jones.
Oehl, M. 2001. The Dialect of Bolton. Available at URL <http://www.haus arbeiten.de/rd/ faecher/ani_9.html>.
Okamoto, F. 1962. *Universal Auxiliary Language Babm*. Ms., Tokyo.
Quirk, R., S. Greenbaum, G. Leech, & J. Svartvik. 1985. *A Grammar of Contemporary English*. London: Longman.
Radford, A. 1988. *Transformational Grammar: A First Course*. Cambridge: Cambridge University Press.
Rosenfelder, M. no date. The Language Construction Kit. At URL <http://www.zompist.com/kit.html>.
Russell, B. 1966. *Suma* (3rd edition). Plainview, NY.
Talmey, M. 1925. *Arulo*. New York: Ilo Press.
Wald, M. 1909. *pan-kel* (4th edition). Ms., Gross-Beeren.
Weilgart, J. 1979. *aUI*. Decorah, IA: Cosmic Communication Co.
Weisbart, J. 1912. *Europal*. Berlin: Michaelis u. Neumann.
Wennergren, B. 2003. *PMEG: Plena Manlibro de Esperanta Gramatiko* (Versio 12.1). Available at URL <http://www.bertilow.com/pmeg/index.php>. (We refer to this work as the PMEG.)