



UvA-DARE (Digital Academic Repository)

Parts of speech and dependent clauses: A typological study

van Lier, E.H.

Publication date
2009

[Link to publication](#)

Citation for published version (APA):

van Lier, E. H. (2009). *Parts of speech and dependent clauses: A typological study*. LOT.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

PARTS OF SPEECH IN THE LANGUAGES OF THE SAMPLE

5

5.1 Introduction

In this chapter I present the classification of parts of speech (PoS) classes attested in the sample languages, following the definitions developed in Chapter 2. In particular, for every language in the sample I determine the set of PoS classes it displays, and define which propositional functions each PoS class can express without extra structural coding. These data are presented in section 5.2. Subsequently, in section 5.3, I evaluate the attested PoS classes and PoS systems in terms of the predictions of the implicational map model of parts of speech, as presented in Chapter 2.

The remainder of the chapter addresses a variety of issues regarding lexical classification in the sample languages. These issues have already been touched upon in the theoretical discussion on PoS typology in Chapter 2, but their practical implications are now further illustrated with actual language data. First, in section 5.4, I discuss the generality problem and the subclass problem. I show how the generality problem is circumvented by disregarding behavioural potential as a criterion for PoS classes. In addition, I discuss the problem of fuzzy boundaries between PoS classes. I first illustrate cases of lexeme classes with variable distributional patterns. Second, I consider different types of ‘restricted’ lexeme classes, i.e. small, closed classes and classes consisting of derived lexemes. Section 5.5 discusses the application of the typology of non-verbal predication to the sample languages. Section 5.6 focuses on the identification of flexible PoS classes. Finally, section 5.7 is a brief conclusion.

5.2 Overview of the language data

Table 5.1 below shows the PoS classes that are attested in the languages of the sample. The languages are listed in the leftmost column; they are ordered from maximally flexible, via less flexible, to fully rigid. Within the group of languages with rigid PoS systems, those with four large, open lexeme classes precede the ones with (an increasing amount of) restricted and/or lacking PoS classes. Every PoS class appears in a separate row. The terms for different types of PoS are abbreviated as follows: *C* for contentives, *non-V* for non-verbs, *Pred* for predicatives, *Nom* for nominals, *V* for verbs, *N* for nouns, *Adj* for adjectives and *mAdv* for manner adverbs⁷⁰. These abbreviations appear in every column corresponding to a function that can be expressed without extra structural coding by the PoS class in question⁷¹. The names of the four propositional functions appear in the top row of the Table: *Pred Head* for head of a predicate phrase, *Ref Head* for head of a referential phrase, *Ref Mod* for modifier in a referential phrase, and *Pred Mod* for modifier in a predicate phrase. The two rightmost columns are used to add information about the size and status of a PoS class: When it is a small, closed class, an *S* appears in the penultimate column (with the heading *Small*). When a PoS class consists of derived members only, a *D* appear in the rightmost column (with the heading *Derived*). All other, unmarked cases involve large, open PoS classes.

Table 5.1: PoS classes in the sample languages

Language	Pred Head	Ref Head	Ref Mod	Pred Mod	Small	Derived
Tagalog	C	C	C	C		
Kharia	C	C	C	C		
Kambera	C	C	C	C	S	D
	V			mAdv		
Samoan	C	C	C	C	S	D
	V			mAdv		

⁷⁰ Some of the PoS classes predicted in Chapter 2 do not appear in Table 5.1. This is discussed further in section 5.2.

⁷¹ If a non-verbal PoS class has an additional predicative use (with a zero-1 expression strategy, see Chapter 2, section 2.3.1), then the abbreviation of the particular PoS class appears also in the column for Head of a Predicate Phrase).

Language	Pred Head	Ref Head	Ref Mod	Pred Mod	Small	Derived
Guarani	C V	C	C	C		
Santali	C V	C	C	C		
Warao	V	non-V N	non-V	non-V		D
Turkish	V	non-V	non-V Mod	non-V Mod		D
Kayardild	Pred	non-V	non-V Adj	Pred non-V mAdv	S	
Paiwan	Pred N Adj	N	Adj	Pred	S	
I. Quechua	V	Nom	Nom	mAdv	S	
Ma'di	V	Nom	Nom Mod	Mod mAdv mAdv	S S	D
Gooniyandi	V	Nom	Nom	mAdv		
Hungarian	V	Nom	Nom	mAdv mAdv		D
Japanese	V	N Nom	Nom	mAdv		
Hmong Njua	V Mod	N	Mod	Mod		

Language	Pred Head	Ref Head	Ref Mod	Pred Mod	Small	Derived
Lango	V	N	Mod	Mod mAdv		
Ket	V	N	Mod Adj Adj	Mod mAdv	S S	D
Koasati	V	N	Mod Adj	Mod mAdv mAdv	S S	D D
Itelmen	V	N	Mod Adj	Mod mAdv	S	D D
Thai	V Adj Mod mAdv	N	Adj Mod	Mod mAdv	S	D
Basque	V	N	Adj Mod	Mod mAdv	S	D
Abun	V Adj Mod	N	Adj Mod	mAdv Mod	S	
Bambara	V	N	Adj	mAdv	S	

Language	Pred Head	Ref Head	Ref Mod	Pred Mod	Small	Derived
Georgian	V	N	Adj	mAdv		
Bukiyip	V	N	Adj	mAdv mAdv	S	
Abkhaz	V	N	Adj	mAdv mAdv	S	D
Polish	V	N	Adj	mAdv mAdv	S	D
Burushaski	V	N	Adj	mAdv mAdv	S	D
Lavukaleve	V	N	Adj	mAdv	S	
Alamblak	V	N	Adj	mAdv	S	
Pipil	V	N	Adj			
Wambon	V	N	Adj			

Language	Pred Head	Ref Head	Ref Mod	Pred Mod	Small	Derived
Dhaasanac	V Adj	N	Adj Adj		S	
Berbice D.C.	V Adj	N	Adj			
Babungo	V	N	Adj Adj	mAdv	S S	D
Nama	V	N	Adj Adj	mAdv	S	D D
Hdi	V	N	Adj	mAdv mAdv	S S	D
Mandarin C.	V	N	Adj	mAdv	S S	
Tamil	V	N	Adj	mAdv	S S	
Kisi	V	N	Adj Adj		S	D
Nung	V Adj	N	Adj		S	
Garo	V	N		mAdv mAdv	S	D

Language	Pred Head	Ref Head	Ref Mod	Pred Mod	Small	Derived
Krongo	V	N		mAdv	S	
Hixkaryana	V	N		mAdv	S	
Slave	V	N				
Nivkh	V	N				
W.Greenlandic	V	N				
Nunggubuyu	V	N				
Tuscarora	V	N			(S)	

5.3 Results in relation to the implicational map of parts of speech

5.3.1 Introduction

In Chapter 2, three hierarchical relations were proposed as constraints on cross-linguistic variation in PoS classification: A primary relation *predication* \subset *reference*, a secondary relation *head* \subset *modifier*, and a ranking between these two: ((Predication/Reference) \subset (Head/Modifier)). These dominance relations were combined into an implicational map model of parts of speech. On the basis of this model, a number of possible PoS systems were predicted. In addition, it was hypothesized that, as long as no constraints are violated, in individual languages the predicted PoS systems can be mixed or supplemented with additional classes.

In this section, I evaluate the data from the sample languages, as presented in Table 5.1 above, in terms of the predictions of the implicational PoS map. I start out with a presentation of the languages that exhibit one of the predicted PoS systems in its ‘pure’ form. Second, I turn to languages with mixed or ‘intermediate’ PoS systems. Third, I discuss cases of PoS systems that were predicted in Chapter 2 but are not attested in the sample. There are also some potential examples of the reverse situation: PoS systems that

were excluded in Chapter 2 on the basis of constraint violation(s), but that are nonetheless attested in the sample.

5.3.2 Predicted and attested ‘pure’ systems

In this section I present the languages that exhibit one of the PoS systems predicted in Chapter 2, without any mixing or additional classes. I start with flexible systems, followed by rigid systems, in the same order as in Table 5.1.

Tagalog and Kharia both have contentives as their only PoS class⁷². This type of system is represented in (1) (cf. (14) in Chapter 2):

(1)

	Head	Modifier
Predication	Contentives	
Reference		

Hmong Njua has verbs, nouns, and a class of modifiers that can also be used predicatively⁷³. This system is represented in (2) (cf. (16) in Chapter 2):

(2)

	Head	Modifier
Predication	Verb	Modifiers
Reference	Noun	

Gooniyandi and Hungarian have a PoS system with verbs, nominals, and manner adverbs, as in (3) (cf. (21) in Chapter 2). Note that Hungarian has both simple and derived manner adverbs.

(3)

	Head	Modifier
Predication	Verb	
Reference	Nominal	

⁷² In section 5.6 the analysis of flexible PoS systems is discussed in full detail.

⁷³ In Table 5.1 this is indicated by the abbreviation *Mod* in the column for *Pred Head*.

Georgian has four large, open classes, each of them specialized for a single function. This is represented in (4) (cf. (17) in Chapter 2):

(4)

	Head	Modifier
Predication	Verb	Manner adverb
Reference	Noun	Adjective

The PoS system represented in (5) (cf. (18) in Chapter 2) has verbs, nouns, and adjectives, but no manner adverbs. It is attested in Pipil, Wambon, Dhaasanac, and Berbice Dutch Creole. Note that these languages differ in terms of the ability to use adjectives in predicative function with a zero-1 strategy (see Chapter 2, section 2.3.1 and the discussion in section 5.4 below): In Pipil and Wambon adjectives cannot be used as verbs, while in Berbice Dutch they can. Dhaasanac has both types: a large class of non-predicative adjectives, and a small class of predicative ones. In Table 5.1, predicative adjectives are recognisable by the fact that the abbreviation *Adj* appears in the column for modifier in a referential phrase *and* in the column for head of a predicate phrase. In contrast, classes of non-predicative adjectives are indicated with *Adj* in the column for modifier in a referential phrase only (cf. notes 71 and 73).

(5)

	Head	Modifier
Predication	Verb	–
Reference	Noun	Adjective

Finally, there are languages with the PoS system represented in (6) (cf. (19) in Chapter 2), with verbs and nouns, but no lexeme classes for the two modifier functions. This system is attested in Slave, Nivkh, West Greenlandic, Nunggubuyu, and Tuscarora.

(6)

	Head	Modifier
Predication	Verb	–
Reference	Noun	–

The PoS system of Tuscarora may alternatively be analyzed as intermediate between the system in (6) above and the system represented in (7) below (cf. (20) in Chapter 2).

(7)

	Head	Modifier
Predication	Verb	–
Reference	–	–

The reason for this is that Tuscarora (like other Iroquoian languages) often uses verbal constructions where other languages would use nouns. Mithun describes this phenomenon as follows:

“A striking feature of natural speech in Iroquoian languages [...] is the relative rarity of nouns. It can be attributed in part to noun incorporation, as in ‘he suit-case-carried’. [...] More important are the functions that verbs can serve. [...] They can be used as descriptive labels for entities (objects, animals, people) and even proper names [...]” (Mithun 2000: 412)

An example of such a descriptively used verb is given in (8):

Tuscarora (Mithun-Williams 1976: 30)

- (8) *Ra-hren-abs*
MASC-CUT-PROGR
‘he cuts’ → ‘surgeon’

Apparently however, this “*relative rarity of nouns*” in Tuscarora does not extend to the extreme situation in which only verbs remain. I have tried to capture this in Table 5.1 by adding a bracketed *S* to Tuscarora’s noun class.

Many other languages, besides Tuscarora, also display a PoS system which combines features of more than one of the systems predicted in Chapter 2. To these ‘intermediate’ systems I turn in the next section.

5.3.3 Predicted and attested ‘intermediate’ systems

First, there are two languages in the sample that combine a class of contentives with a large, open class of verbs: Guaraní and Santali. Thus, the PoS system of these languages can be analyzed as a combination of the systems represented in (1) and (7) above. Other languages have contentives in combination with a class of *derived* verbs. This holds for Kambera and Samoan. In addition, these languages both display another extra class, namely a small, closed class of manner adverbs. This is an unexpected pattern to the extent it involves a head-modifier distinction (derived verbs versus adverbs) without a full-blown predication–reference distinction (see constraint (13) of Chapter 2). However, it seems that adverbs in Kambera and Samoan have many characteristics of function words, as opposed to content words (see also section 5.4.3.2).

Second, the system in (9) (cf. (15) in Chapter 2) is attested in Warao, but in combination with a class of derived nouns.

(9)

	Head	Modifier
Predication	Verb	
Reference	Non-verb	

The PoS system of Warao may thus be represented as in Figure 5.1 below. This case is discussed in more detail in section 5.6.3.1.

	Head	Modifier
Predication	Verbs	Non-verbs
Reference	Derived nouns	

Figure 5.1: The PoS system of Warao

There are two more languages with intermediate PoS systems involving the system in (9). First, Ma’di combines a PoS system of verbs and non-verbs with a system of verbs, nominals, and manner adverbs. The latter is represented in (10) (cf. (21) in Chapter 2):

(10)

	Head	Modifier
Predication	Verb	Manner adverb
Reference	Nominal	

Turkish, as discussed in Chapter 2 (see Figure 2.5), also has a PoS system with verbs and non-verbs (see (9) above), but combines it with a class of derived modifiers, i.e. lexemes with the distributional possibilities of non-verbs minus the function of head of a referential phrase.

Imbabura Quechua has an intermediate PoS system that involves a combination of the system in (10) above, and the one in (11) below (cf. (22) in Chapter 2), since it has verbs, nominals, and a closed class of manner adverbs.

(11)

	Head	Modifier
Predication	Verb	–
Reference	Nominal	

Japanese combines the system in (10) above with the system in (12), which has verbs, nouns, and manner adverbs (cf. (30) in Chapter 2)⁷⁴.

(12)

	Head	Modifier
Predication	Verb	Manner adverb
Reference	Noun	–

Another type of intermediate PoS system involves a system with verbs, nouns, and flexible modifiers, in combination with rigid classes of adjectives and/or manner adverbs. This is attested quite often, namely in Lango, Ket, Koasati, Itelmen, Thai, Abun, and Basque. There is variation in the size and status of the flexible and rigid PoS classes that can express the modifier functions in these PoS systems. Lango, for instance, has a large class of

⁷⁴ Alternatively, the Japanese system can be described as the system in (10) supplemented with a class of rigid nouns. The PoS system of Japanese is discussed in more detail in section 5.4.2.1.

flexible modifiers and a large class of rigid manner adverbs, but no rigid adjectives. Thus, its PoS system can be analyzed as a combination of the system in (12) above and the one in (13) below:

(13)

	Head	Modifier
Predication	Verb	Modifier
Reference	Noun	

Ket also has a large class of modifiers, but in combination with small classes of rigid adjectives and manner adverbs (cf. (4) above). In addition, Ket has derived adjectives.

Another possibility is that the rigid adjective and/or adverb class(es) are large, while the flexible modifier class is relatively small. In Abun, for instance, the classes of adjectives and manner adverbs are larger than the modifier class. Thai and Basque both have full-blown adjective classes, combined with (very) restricted classes of modifiers and manner adverbs. Finally, it may be the case that all three types of PoS classes, i.e. modifiers, adjectives and adverbs, are small and/or derived. This is attested in Koasati and Itelmen.

A number of other languages have a fully differentiated PoS system with four rigid PoS classes, in which either the adjective class or the manner adverb class or both are small and/or derived. The following languages have a small and/or a derived class of manner adverbs: Abkhaz, Polish, Burushaski, Lavukaleve, and Alamlak. This type of PoS system is represented in (14). It may be regarded as a system intermediate between the pure systems presented in (4) and (5) above.

(14)

	Head	Modifier
Predication	Verb	S/D Manner adverb
Reference	Noun	Adjective

Bukiyip has two types of manner adverbs. One type takes verbal inflection, but cannot appear independently as the head of a predicate phrase. Conrad and Wogiga (1991) do not give information about the size of this class,

nor about its open or closed status. An example of this type of adverb is given in (15) below. The second class is described as a small, closed class of uninflected manner adverbs, consisting of some 13 items (Conrad & Wogiga 1991: 41).

Bukiyip (Conrad & Wogiga 1991: 41)

- (15) *Awou* *w-a-gamu*
 3PL.FEM 3PL.FEM.SUBJ-REAL-well
w-a-dúkemech
 3PL.FEM.SUBJ-REAL-understand
 ‘The women understand well.’

There are also languages in which both adjectives and manner adverbs constitute small and/or derived classes, as is represented in (16) below. This is the case in Babungo, Nama, Hdi, Mandarin Chinese, and Tamil.

(16)

	Head	Modifier
Predication	Verb	S/D Manner adverb
Reference	Noun	S/D Adjective

Finally, there are languages in which either the adjective class is small/derived and the manner adverb class is altogether missing, or in which the adjective class is missing and the adverb class is small/derived. The former situation is represented in (17). It is attested in Kisi and Nung. Kisi has both a small and a derived class of adjectives, neither of which can be used predicatively. Nung has only a small class of adjectives, which can also be used as verbs.

(17)

	Head	Modifier
Predication	Verb	–
Reference	Noun	Adjective

Finally, the reverse situation – a restricted class of manner adverbs and no adjectives – is represented in (18). This type of system is attested in Garo, Krongo, and Hixkaryana.

(18)

	Head	Modifier
Predication	Verb	S/D manner adverb
Reference	Noun	–

5.3.4. Predicted but not attested systems

(and attested but not predicted systems)

When considering all the predicted PoS systems in Chapter 2, we find that the following ones are not attested in any of the sample languages:

- (i) PoS systems with flexible predicatives: see (19) and (20) below (and cf. (23) and (24) in Chapter 2);
- (ii) PoS system with flexible heads: see (21) below (and cf. (25) in Chapter 2);
- (iii) PoS systems with a class of flexible lexemes that can be used in all functions except one of the modifier functions (*Flex A* and *Flex B*): see (22) and (23) below (and cf. (26) and (27) in Chapter 2);
- (iv) PoS systems that involve either one or both of the ‘cross-wise’ flexible classes *Flex C* and *Flex D*: see (24) and (25) below (and cf. (28) and (29) in Chapter 2).

(19)

	Head	Modifier
Predication	Predicative	
Reference	Nominal	

(20)

	Head	Modifier
Predication	Predicative	
Reference	–	–

(21)

	Head	Modifier
Predication	Head	–
Reference		–

(22)

	Head	Modifier
Predication		–
Reference	Flex A	

(23)

	Head	Modifier
Predication		
Reference	Flex B	–

(24)

	Head	Modifier
Predication	Verb	Flex C
Reference	Flex C	–

(25)

	Head	Modifier
Predication	Flex D	Flex C
Reference	Flex C	Flex D

I start out with the systems involving predicatives. First, the fact that the system in (20) is not attested does not appear to be very surprising, since this system has no lexical strategy to express any function within the referential domain. Even though constraint (11) in Chapter 2 determines that the predicative domain ranks higher than the reference domain, this constraint seems to have more bearing on functional specialization (as opposed to flexibility) of PoS classes, rather than predicting the presence versus

absence of a lexical strategy in the first place. Apparently, as the discussion of Tuscarora in the previous section also suggests, languages may make relatively extensive use of verbal as opposed to nominal constructions, but a complete lack of lexical strategies to express any function in the referential domain seems to be a marginal option, if it occurs at all.

In contrast, the system in (19) does have a lexical strategy for the expression of referential functions: it has predicatives *and* nominals. In fact, even though this system is not attested in its ‘pure’ form, there are two languages in the sample with PoS systems that resemble it to some extent. One of these languages is Paiwan. It seems that all lexemes in this language that usually occur as predicative heads (i.e. ‘verbs’) can also be used as predicate modifiers (‘manner adverbs’), and vice versa. There are only two items that possibly qualify as lexical adverbs, because they cannot be used predicatively (Egli 1990: 158, 313). Since only one of these items appears to be a manner adverb, I do not make reference to a (small) class of manner adverbs in Paiwan (see Table 5.1). Even though Paiwan thus has a class of predicatives, its PoS system differs from the system in (19): Paiwan has rigid nouns and a restricted class of rigid adjectives, whereas the system in (19) has nominals. Thus, the PoS system of Paiwan can be analyzed as intermediate between the systems represented in (26)* and (27)* below (cf. the systems in (15)* and (16)* of Appendix 11). As the asterixes are meant to indicate, these systems are both excluded by the implicational PoS map, in particular on the basis of constraint (11a) in Chapter 2, which states that the availability of a rigid class for heads of referential phrases (nouns) implies the availability of a rigid class for heads of predicate phrases (verbs).

(26)*

	Head	Modifier
Predication	Predicative	
Reference	Noun	Adjective

(27)*

	Head	Modifier
Predication	Predicative	
Reference	Noun	–

There is, however, an alternative analysis, which hinges on the fact that it is not easy to distinguish in Paiwan between a predicative head modified by a simple lexical manner expressions and a serial verb construction, since the language has very little verbal morphology. Therefore, it is possible that Paiwan predicatives are in fact rigid verbs (cf. Egli 1990: 312). If this is the right analysis, then constraint (11a) of Chapter 2 is no longer violated.

The second language in the sample with predicatives is Kayardild. According to Evans, in this language

“[V]erbals primarily denote actions and processes, but may also provide adverbial type information about the manner in which these are carried out. [...] A few lexemes only permit the modifier function.” (Evans 1995: 86)

In fact, Evans lists only four lexemes that are restricted to the function of modifier in a predicate phrase (i.e. rigid manner adverbs), and about these he remarks that “*it is possible that a bigger corpus would see even these used as main verbs*” (Evans 1995: 303-304). Thus, it seems that Kayardild indeed has predicatives. However, Kayardild’s PoS system also differs in some respects from the system in (19) above: First, Kayardild has a class of non-verbs, rather than nominals. In addition, there is a class of rigid manner adverbs, which are analyzed as a subclass of non-verbs, since, unlike predicatives, they take nominal inflection (Evans 1995: 227-229). There are thus two lexical strategies available to express the function of modifier in a predicate phrase: a ‘verbal’ and ‘nominal’ one. Moreover, Kayardild has a reasonably large class of rigid adjectives (about 100 items). This is unexpected, since it involves a specialized class for modification in the referential phrase, without a specialized class for the corresponding head function (i.e. without rigid nouns, cf. constraint (12a) of Chapter 2). In Figure 5.2 the PoS system of Kayardild is schematically represented:

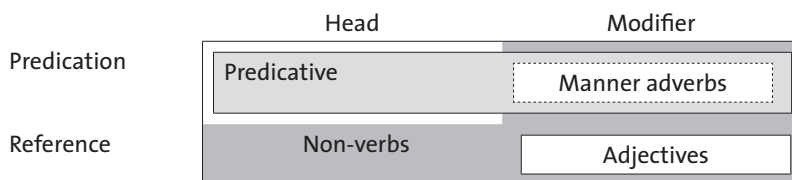


Figure 5.2: The PoS system of Kayardild

I now turn to the predicted but unattested PoS system in (21) above. Even though this system does not occur in its pure form, there are some languages in the sample with a PoS system consisting of rigid nouns and verbs in combination with a restricted group of lexemes that can apparently be used as both. The relevant languages are Slave, Nivkh, West Greenlandic, and possibly Nunggubuyu. The question is, however, whether these apparent ‘heads’ are truly flexible lexemes, or whether they rather involve zero-derivation/conversion or polysemy. In section 5.6, where I discuss lexical flexibility in the sample languages, I will argue that the non-compositional semantic interpretation of the relevant lexemes in verbal and nominal function suggests that they do not represent cases of true flexibility.

The non-attestation of flexible heads may also be related to the fact that the propositional function(s) of modification appear not to be relevant in all languages. Rather, in some languages ‘modifiers’ take the form of either separate predications or of appositional referential expressions. In such cases the head-modifier parameter becomes irrelevant, which means that a hypothetical class of ‘heads’ would in fact be a class of ‘contentives’, in the sense of a lexeme class whose members can be used in all (relevant) functions⁷⁵.

The issue of irrelevance of propositional function(s) probably plays a similar role in explaining why the systems in (22) and (23) above are not attested. It is quite possible that some of the languages classified as having a single class of contentives, in fact have a PoS class with maximally flexible lexemes, which can be used in the complete set of *relevant* functions; a set that excludes one of the two modifier functions. In Tagalog, for instance, it seems that the slot for modification in a predicate phrase is irrelevant (see Hengeveld & Van Lier 2009).

Finally, the systems in (24) and (25) above seem intuitively rather improbable, even though they cannot be excluded on the basis of the constraints as formulated in Chapter 2. In particular, these systems involve one or two flexible lexeme class(es), which combine the opposite values of both the predication-reference and the head-modifier parameter. In other words, the relevant flexible classes combine the head function of one domain with the modifier function of the other domain. The markedness

⁷⁵ Note that the possibility of irrelevant propositional function(s) in specific languages suggests that the functional space as defined in Chapter 2 is not universal. This is in line with the assumptions of the framework of Functional Discourse Grammar, which explicitly distinguishes between language-specific pragmatic and semantic functions, which belong to the grammar of a language, and universal conceptualisation, which belongs to general cognition (Hengeveld & Mackenzie 2008).

of this situation, i.e. the neutralization of maximally divergent functional distinctions, may explain the non-occurrence of the systems (24) and (25) in the sample languages.

Of course, in these and all other cases discussed above, it is possible that the non-attestation of a predicted PoS system is purely coincidental, and that the relevant system will turn up in an investigation of a larger sample. In addition, as I have suggested several times above, both the non-attestation of predicted systems and the attestation of excluded systems may be due to particular difficulties in the analyses of the available language data. Finally, it may very well be that the formulation of the restrictions on PoS systems in Chapter 2 is not yet optimal, and/or that the model should take into account more functional parameters.

5.3.5 Summary

To summarize, in this section I have related the PoS systems attested in the languages of the sample to the PoS systems predicted in Chapter 2. It was shown that most predicted systems are actually attested, either in their 'pure' form or in the form of an 'intermediate' system. I have suggested possible explanations for the cases of predicted but unattested PoS systems. Apart from the unexpected attestation of rigid adjective and noun classes in the two languages with predicatives (Paiwan and Kayardild), none of the systems that were excluded on the basis of the constraints in Chapter 2 (see Appendix II) have been found in the sample languages. In general, these results suggest that the implicational map model of PoS developed in Chapter 2 has a reasonable typological adequacy.

5.4 The generality and the subclass problems in practice

5.4.1 The generality problem

As explained in Chapter 2, the Hengeveldian method for identifying PoS classes crucially depends on whether or not extra structural coding (specific function-indicating morpho-syntax) is required to use a member of a particular lexeme class in a certain propositional function. In contrast, behavioural potential (the expression of morpho-syntactic categories belonging to a particular function) does not play a role in the definition of PoS classes.

In the present study, the practical application of this method has in some cases resulted in analyses of PoS systems that differ from the ones

proposed in the relevant descriptive sources. In all these cases, the difference in analysis is of the same nature: I propose a flexible PoS class where the reference grammar distinguishes two (or more) rigid word classes. The reverse situation – in which a reference grammar claims flexibility, while I propose multiple rigid PoS classes – does not occur. Obviously, this ‘lumping’ effect is due to the fact that structural coding is a rather restricted defining criterion. Most reference grammars identify PoS classes using the full set of language-specific distributional data, as advocated by Croft (2001). However, as discussed in Chapter 2, the Croftian method raises the generality problem: PoS classes identified in this way are not cross-linguistically comparable. In what follows, I will give examples to show how the Hengeveldian method gets around the generality problem. At the same time, these examples illustrate how the results obtained with the Hengeveldian method may differ from the results of a complete language-specific distributional analysis.

First, consider Kambera⁷⁶. According to Klamer (1998: 91–95), Kambera has verbs, nouns, and a class of multifunctional lexemes that can be used as both. Nouns and verbs are distinguished by Klamer on the basis of the following criteria:

- (i) Nouns can occur with articles and with the marker *bai/bi*, which expresses appreciation or derogation of (properties of) the noun. Noun phrases can be modified by emphatic or demonstrative pronouns. Nouns also have specific quantifying properties. They may occur, for instance, with a bare, un-derived numeral and may have a classifier.
- (ii) Verbs can be derived and they can be modified by adverbs (adverbs do not modify nouns).

These criteria involve language-specific aspects of behavioural potential, rather than structural coding. Therefore, following the Hengeveldian approach, these facts are not regarded as evidence for a lexical distinction between nouns and verb.

Moreover, there is explicit evidence in favour of a flexibility analysis. First, Klamer (1998: 96) acknowledges the presence of “*a considerable number*

⁷⁶ Recall that Kambera is one of the two languages added to the balanced sample (see Chapter 4), and will therefore not be included in the analyses of Chapter 7.

of multifunctional lexical items” in Kambera. In addition, ‘nouns’ and ‘verbs’ have identical distributional possibilities in terms of the ability to be used predicatively and referentially. This is illustrated in examples (28) and (29) below. In example (28a), an action-denoting lexeme is used as the head of a predicate phrase and takes the inflection belonging to this function, while in (28b) the same lexeme is used as the head of a referential phrase and combines with an article. In (29) two object-denoting lexemes are used as predicative heads⁷⁷.

Kambera (Klamer 1998: 105, 107)

(28) a. *Jāka manganga-ma-na-yna-i una, ...*
 if steal-EMPH-3SG.GEN-3SG.DAT-ITER EMPH.3SG
 ‘If he does steal again, ...’

b. *Na ma-kaloru-nya na manganga...*
 ART REL.SUBJ-arrange-3SG.DAT ART steal
 ‘Who is engaged in theft, ...’

(29) *Nina nda tustel-a-ya, senter-ya*
 DEIC-3SG NEG camera-MOD-3SG.ACC torch-3SG.ACC
 ‘This is not a camera, it’s a torch.’

Apart from expressing the two head functions, ‘verbs’ and ‘nouns’ in Kambera can also be used, without structural coding, as modifiers in predicate phrases and referential phrases. The examples in (30) show an action-denoting lexeme (a) and an object-denoting lexeme (b) functioning as modifiers in a referential phrase. Examples (31a) and (31b) show an action-denoting and an object-denoting lexeme, respectively, in the function of modifier in a predicate phrase.

Kambera (Klamer 1998: 108, 109)

(30) a. *iyang wāu* b. *meu rumba*
 fish smell cat grass
 ‘smelly fish’ ‘wild cat’

⁷⁷ These examples show that the semantic interpretation of the lexemes in either function is fully compositional: an action-denoting lexeme in nominal function denotes ‘the act of X-ing’, while an object-denoting lexeme in verbal function denotes ‘be X’. I return to this issue in section 5.6.

- (31) a. *dedi meti-ma-a-nanya na ina-na*
be.born die-EMPH-MOD-3SG.CONT ART mother-3SG.GEN
 ‘His mother died in labour’ (lit: ‘died while giving birth’)
- b. *jangga eti*
 be.tall **liver**
 ‘be arrogant’ (lit.: ‘be tall liver-wise’)

In short, these examples show that there is no difference between nouns and verbs in Kambera, at least not in terms of the definitions developed in Chapter 2.

Ma’di represents another case in which the Hengeveldian method results in a different classification of a language’s PoS system than the one proposed in the reference grammar. Blackings and Fabb (2003) distinguish between nouns and adjectives in Ma’di, on the basis of the following distributional facts:

- (i) Most adjectives are inflected for number (except for colour terms), while most nouns are not;
- (ii) Adjectives always combine with an article when they are used as the head of a referential phrase, whereas nouns can also appear without one.

However, adjectives and nouns do not differ in terms of their possibility to be used, without structural coding, as the head and the modifier of a referential phrase. This is illustrated in (32) and (33). In the first example a property-denoting lexeme is used as the head of a referential phrase; in the second example an individual-denoting lexeme is used as a modifier in a referential phrase. Therefore, I have classified Ma’di as a language with flexible nominals.

Ma’di (Blackings & Fabb 2003: 106, 304)

- (32) *āli ri pi ē-tsá ádʒmī*
 short(PL) DEL PL.PRON (3)-VE-arrive yesterday
 ‘The short ones arrived yesterday.’

- (33) *Mādī ādrúpi ri ɔdā rō ilé-ni*
 person **brother**(INDEF) DEF act REFL that-like
kū
 NEG(NON-PST)
 ‘A person who is a brother won’t behave like that.’

Note in passing that a determiner is also obligatory in Ma’di when both a modifier and a head are present in the referential phase. The ungrammatical example in (34) shows this:

Ma’di (Blackings & Fabb 2003: 302)

- (34) **bási ingwē ē-ɬfá rá*
 bus white (3)-VE-arrive AFF
 ‘The/a white bus has certainly arrived.’

The case of Ma’di may be contrasted with a case like Hdi. Frajzyngier and Shay (2002: 71) use the following two criteria to define the class of adjectives in Hdi:

- (i) Adjectives can modify a noun without any intervening marker (whereas nouns cannot);
- (ii) Adjectives cannot be used as arguments (whereas nouns can).

Example (35a) and the ungrammatical (35b) show that these two criteria qualify the lexeme *xɔ̀úzá* ‘big-bellied’ as an adjective:

Hdi (Frajzyngier & Shay 2002: 72)

- (35) a. *ndá ngh-íyù tá gù xɔ̀úzá*
 ASSOC see-1SG OBJ goat big-bellied
 ‘I saw a big-bellied goat.’
- b. **ndá ngh-íyù tá xɔ̀úzá*
 ASSOC see-1SG OBJ big-bellied
 ‘I saw a big-bellied one.’

Thus, the distinction between nouns and adjectives, as proposed by Frajzyngier and Shay (2002) is based on functional possibilities and

structural coding only. Therefore, their PoS classification is the same as the one I arrive at using the Hengeveldian method.

In this section I have illustrated how the criterion of structural coding allows for the establishment of cross-linguistically comparable PoS classes and PoS systems. It should be clear that the down-side of this method is that it may result in lumping together groups of lexemes that would be regarded as separate classes when the full set of language-specific distributional facts would be taken into account. This brings us to the related subclass problem, which is discussed in the next subsection. Further on, in section 5.6, I return in more detail to the issue of defining flexible PoS classes.

5.4.2 The sub-class problem

In this section I discuss a number of issues related to the sub-class problem. First, as pointed out in Chapter 2, PoS classes in particular languages may have fuzzy boundaries in terms of their distributional behaviour. Such PoS classes display a combination of properties associated with two other groups of lexemes and/or they display varying behaviour in a single propositional function. These phenomena are illustrated in section 5.4.2.1. A second problem concerns the boundary between major, open PoS classes and various types of more restricted classes. In section 5.4.2.2 I consider two types of such restricted classes: (i) small, closed lexeme classes, and (ii) classes of derived lexemes.

5.4.2.1 Fuzzy boundaries

In Chapter 2, I discussed the distinction between nouns and adjectives in Japanese. This case was adduced by Croft (2001) to illustrate that the Hengeveldian approach to PoS definition cannot handle the phenomenon of fuzzy boundaries between language-specific lexeme classes. Since Japanese is one of the languages of my sample, I consider here once more the relevant data, which are presented in Figure 5.3. The different groups of lexemes identified in Croft's analysis are indicated here with numerical codes (cf. Figure 2.13 in Chapter 2).

PoS class	Function/Coding		
	Head of Pred Phrase	Head of Ref Phrase	Mod in Ref Phrase
1	<i>da</i> (COP)	∅	<i>no</i> (GEN)
1/2	<i>da</i> (COP)	∅	<i>no</i> (GEN)/ <i>na</i> (LK)
2	<i>da</i> (COP)	∅	<i>na</i> (LK)
2/3 a	<i>da</i> (COP)/∅ + Tense	∅	<i>na</i> (LK)/∅ + Tense
2/3 b	∅ + Tense	∅	<i>na</i> (LK)/∅ + Tense
3	∅ + Tense	X	∅ + Tense

Figure 5.3: *The noun-adjective distinction in Japanese (2)*

Following the Hengeveldian approach, the data in Figure 5.3 result in the distinction of three PoS classes in Japanese:

- (i) A class of nouns (class number 1): lexemes that are used without structural coding as the head of a referential phrase. These lexemes need a copula when used in predicative function, and a genitive marker *no* when used as modifier in a referential phrase;
- (ii) A class of nominals (class number 2): lexemes that are used without structural coding as the head of a referential phrase, and with the linking element *na* (rather than *no*) when used as a modifier in a referential phrase. In predicative function these lexemes need a copula.
- (iii) A class of verbs (class number 3): lexemes that are used without structural coding as the head of a predicate phrase, and that cannot be used as the head of a referential phrase. When used as a modifier in a referential phrase, these lexemes remain verbal (i.e. are relativized): they inflect for tense and take neither the genitive nor the linking element, as do nouns and nominals, respectively.

The remaining classes, i.e. the ones with the codes 1/2, 2/3, 2/3a and 2/3b in Figure 5.2 above, are not taken into account as separate classes. This is reflected in Table 5.2, which repeats the relevant data for Japanese from Table 5.1.

<i>Language</i>	<i>Pred Head</i>	<i>Ref Head</i>	<i>Ref Mod</i>	<i>Pred Mod</i>
Japanese	V	N Nom	Nom	

Table 5.2: *Verbs, nouns and nominals in Japanese*

In other languages of the sample I have come across comparable cases of fuzziness. For instance, in Mandarin Chinese the existence of a separate class of adjectives is still debated. In terms of the Hengeveldian approach the following facts are relevant: There is a large group of property-denoting lexemes (i.e. semantic ‘adjectives’) that can function as modifiers in referential phrases only when they are structurally coded with the marker *de*, whereas they can be the heads of predicate phrases without any structural coding. The modifying construction is illustrated in (36) below. This construction is exactly the same as the one used for action-denoting lexemes in modifier position, i.e. relative clauses, as illustrated in example (37). For this reason, semantic ‘adjectives’ are analyzed as (stative) verbs.

Mandarin Chinese (Li & Thompson 1981: 32, 581)

(36) [*Hóng=hong de*] *būa*
 red=red LK flower
 ‘flowers that are really red’

(37) [*Jintiān yíng de*] *qián*
 today win LK money
 ‘the money that we won today’

However, according to Paul (2005), there is also a relatively small group of so-called “*non-predicative adjectives*”. When these property-denoting lexemes modify a noun, they appear either with or without the linking element *de*, as is shown in (38) below. In addition, these lexical items cannot function as the head of a predicate phrase without structural coding; they need the copula *shi* in combination with *de*, as is illustrated in (39).

Mandarin Chinese (Paul 2005: 760, 759)

(38) *ta mai-le yi-ge fang (de) panzi*
 3SG buy-PFV 1-CL square LK plate
 ‘He bought a square plate.’

(39) *Zhei-ge panzi shi fang de*
 this-CL plate COP square LK
 ‘This plate is square.’

On the basis of these distributional facts, I have analyzed Mandarin Chinese as having (i) a class of verbs: lexemes that can only be used predicatively, and (ii) a small class of adjectives: lexemes that can (at least sometimes) function without structural coding as modifiers in referential phrases, and that cannot be the heads of predicate phrases without a copula. The fact that the adjective class had ‘fuzzy edges’, in the sense that its members show variable behaviour in the modifier position, is not taken into account. This analysis is represented in Table 5.3, which repeats the data for Mandarin Chinese from Table 5.1:

<i>Language</i>	<i>Pred Head</i>	<i>Ref Head</i>	<i>Ref Mod</i>	<i>Pred Mod</i>	<i>Small</i>
Mandarin Ch.	Verb		Adj		S

Table 5.3: Verbs and Adjectives in Mandarin Chinese

In sum, this section further illustrated the practical repercussions of the restrictive structural coding criterion of the Hengeveldian method. In particular, fuzzy boundaries between language-specific lexeme groups and class-internal behavioural variability are ignored, in order to maintain the cross-linguistic comparability of PoS classes.

5.4.3.2 Restricted classes

Small, closed classes

Lexemes and lexeme classes in languages can be divided (broadly) into content words and function words. Generally, content word classes are open, which means that new members can easily be added to them. Also, content word classes are large; they comprise hundreds or thousands of items. Function word classes, on the other hand, tend to be closed and small.

It can be difficult, however, to determine the boundary between these two types of word classes in individual languages. Especially adjectives and/or (manner) adverbs (or flexible modifier classes, for that matter) are in some languages more like function words than like content words, in that they constitute relatively restricted, closed classes (Haspelmath 2001: 16539-42). The available descriptive sources do not always provide clear indications of the size and status of adjective, adverb, and/or modifier classes. And even if such information is available, definitions remain far from absolute.

I already briefly touched upon this point in Chapter 2, in connection with the adjective class in Alamblak. A comparable case is attested in Abun. According to Berry and Berry:

“The class of adjectives in Abun hovers on the borderline of being an open or a closed word class (...). In the corpus a total of 38 adjectives has been distinguished, but there may well be others, thus it is included as an open class.” (Berry & Berry 1999: 36)

The number of manner adverbs in Abun is somewhat more limited; Berry and Berry (1999: 25) identify 25 items. They state that:

“The subclass of manner adverbs is not as ‘open’ as a regular open class because many adverbial meanings are expressed in other ways, such as associative prepositional phrases or the use of adjectives without any special marking.” (Berry & Berry 1999: 37)

Even though manner adverbs are thus not very numerous, Berry and Berry include them as an open word class in Abun. In addition to adjectives and adverbs, there is a group of about 12 lexical items in Abun that can function as both (see also the second quotation above). This group is analyzed as a small (and probably closed) class of flexible modifiers.

In some other languages even smaller classes are distinguished. Hdi, for instance, has 8 lexical adjectives (Frajzyngier & Shay 2002: 71). Classes of manner adverbs may likewise consist of (very) few members. For example, Tamil is described as having only 3 manner adverbs: *mella* ‘softly’, *molla* ‘slowly’, and *jalti* ‘quickly’ (Asher 1985: 115).

In many other cases, the available descriptive sources do not give specific numbers or lists of items that belong to classes of adjectives, adverbs, or modifiers. Rather, such classes are characterized for instance as ‘restricted’ compared to other classes like nouns and verbs, which comprise much larger amounts of items. Clearly, the bottom line of this section is that at least some of the decisions to mark a PoS class as ‘small’ (S) in Table 5.1 involved a certain degree of arbitrariness.

Derived classes

PoS classes that are marked as ‘derived’ (D) in Table 5.1 involve items that are formed by means of productive derivational processes only⁷⁸. Furthermore, derived PoS classes are mentioned separately only in cases where there is either no categorically equivalent class of simple items, or where the categorically equivalent simple items constitute a small, closed class.

For example, if a language has a large, open class of simple nouns and a productive process to derive nouns, then the latter is not mentioned in the data. In contrast, if a language has no class of simple adjectives, but does have a productive process to derive adjectives, then a derived class of adjectives is included in the data. Also, when a language has but a small, closed class of adjectives, then the possibility to productively derive adjectives is also mentioned separately. In the present section I give some examples of derived PoS classes from the sample languages.

Consider once more Kambera. We already saw that, as far as un-derived items are concerned, Kambera is analyzed as having contentives and adverbs. However, the language also has two productive verbalization processes: (i) *pa-* prefixation, which derives verbs with causative, permissive, factitive, resultative, intensive, infinitive, or reciprocal interpretation; and (ii) *-ng* suffixation, which forms applicative verbs. Some examples of these two processes are given in (40) and (41):

Kambera (Klamer 1998: 179, 199)

- (40) *ànga* ‘foolish/useless’ – *pa-ànga* ‘confuse, cheat on someone’
ana ‘child’ – *pa-ana* ‘have children’
dua ‘two’ – *pa-dua* ‘divide’
ndia ‘no’ (emphatic negator) – *pa-ndia* ‘deny’

- (41) *bungga* ‘open X’ – *bunggabu-ng* ‘open X for Y’
riki ‘laugh’ – *riki-ng* ‘to laugh at/about X’
angu ‘friend’ – *angu-ng* ‘have X as a friend’
nàmu ‘towards the speaker’ – *nàmu-ng* ‘move towards the speaker’

⁷⁸ The notion of ‘productivity’ is of course rather slippery. I have, as much as possible, relied upon unambiguous statements in reference grammars in deciding whether or not a certain derivational process is productive, i.e. applicable to all members of the input class(es).

In languages with rigid PoS systems, it is often the case that a class of derived items exists alongside a small, closed class of underived items with the same categorial value. Usually, the relevant lexeme classes are adjectives and/or manner adverbs. Nama, for instance, has a small class of simple adjectives and two types of productively derived adjectives, formed with the suffixes *-xà* ('attributive') and *-ò* ('privative'), respectively (Hagman 1979: 60). However, classes of derived lexemes also occur without the presence of any categorially equivalent class of simple items. Basque, for example, has derived manner adverbs only. These are productively formed from adjectival bases with the suffix *-ki*, as in *sendo* 'strong', *sendo-ki* 'strongly' (Hualde & Ortiz de Urbina 2003: 193)⁷⁹.

A factor that may complicate the identification of classes of productively derived lexemes is the influence of diachronic processes. First, it may be the case that a certain proportion of the derived forms becomes lexicalized and that their base forms no longer occur independently. Consider for example Babungo: The large majority of the adjectives in this language is derived from verbs. However, there are also a few adjectives for which no corresponding verbal form exists. Historically, these forms probably did exist, but synchronically it seems appropriate to say that Babungo has a small class of simple adjectives, next to a class of derived adjectives.

The case of manner adverbs in Babungo is similar, except that the process of lexicalization is apparently more advanced: Many adverbs take the form of reduplications. Only a few of them, however, can still be identified as derivations from verbs. For instance, the manner adverb *bwāŋŋɔ̄-bwāŋŋɔ̄* 'gently' is derived from the verb *bwāŋŋɔ̄* 'to be well, to be soft'. In contrast, for the lexeme *wèè-wèè* 'slowly, gently' no verbal base form is available (Schaub 1985: 246). In accordance with cases of the second type, Babungo is synchronically analyzed as having a restricted class of underived manner adverbs. Furthermore, since the derivational process is no longer productive, no class of derived manner adverbs is included for Babungo in Table 5.1.

A related potential problem concerns the distinction between lexical and clausal derivation. In Tamil, for example, it may seem as if there is a class of derived adjectives, formed from verbal and nominal bases by means

⁷⁹ In one case, namely Hungarian, I have coded both a class of simple and a class of derived manner adverbs. In fact, however, the status of these classes is not entirely clear. According to Kenesei et al. (1998: 222), simple adverbs are 'numerous'. However, according to Rounds (2001: 180) most manner adverbs are derived, which seems to suggest that the class of simple items is in fact relatively restricted.

of the suffixes *-a* and *-aana/-ulla*, respectively. However, these derivations are in fact all dependent clauses: The suffix *-a* operates on the phrase level, and the argument(s) of the dependent predicate retain the same expression as main clause arguments. This is shown in example (42):

Tamil (Asher 1985: 28)

- (42) [*Akkaa taṅkaccikki caata poṭṭ-a*] *karaṅṭi*
 elder.sister younger.sister-DAT rice put:PST-PTC spoon
 ‘The spoon with which elder sister gave rice to younger sister.’

Derivations from nominal bases with *-aana/-ulla* are no different: These suffixes are in fact dependent, participial forms of the verbs *aaku* ‘to become’ and *ullu* ‘(existential) be’. This can be seen in example (43):

Tamil (Lehmann 2005: 13)

- (43) [*aZak-aan-a*] *kaṅṅati*
 beauty-become.PST-PTC mirror
 ‘beautiful mirror’ (lit. ‘mirror that has become beautiful’)

In sum, what may look like a derived adjective in Tamil is in fact a relative clause construction.

Before rounding off this section, it is worth mentioning that the notion of derived PoS classes clashes in some sense with the criterion of structural coding, since word-class changing derivation can be regarded as an instance of structural coding in itself. Therefore, it may seem strange to say that the members of a class of derived lexemes can be used without structural coding in a particular function. On the other hand, derivation is also a means by which new group of lexical stems is created, which is especially suited for the expression of a specific function. This group of stems may in turn form the input of further processes of morphological or syntactic derivation and structural coding. Moreover, in the context of the present study, the crucial question is whether derived PoS classes may serve as functional models for DC constructions, and there seems to be no principled reason why they may not.

5.5 Non-verbal predication

As explained in Chapter 2, the function of head of a predicate phrase has an exceptional status within Hengeveld’s PoS theory. This is because this function is the *only* possible function of the PoS class defined as ‘verbs’,

while it can be an *additional* function of PoS classes other than verbs. In Chapter 2 I discussed the possible strategies that languages may employ to express non-verbal predication. These strategies were divided into those with a copula and those without one. The strategies without a copula were further subdivided into zero-1 and zero-2 strategies. A zero-1 strategy means that the non-verbal predicate shows the same behavioural potential as a verbal predicate in the language under analysis. A zero-2 strategy means that the non-verbal predicate does not express any verbal categories and is simply juxtaposed to its argument.

The use of a copula strategy counts as structural coding. Therefore, a non-verbal PoS class that employs this strategy is analyzed as not including the function of head of a predicate phrase in its range of functional possibilities. The same holds for non-verbal PoS classes that use a zero-2 strategy: they are also analyzed as non-predicative, since they cannot express the same categories as verbal predicates. In addition, zero-2 strategies often occur in alternation with a copula strategy. In such cases, the copula does not appear when there is no functional need to express verbal features like TAM distinctions and/or Person. Typically this involves ‘default’ cases, such as present tense and third person singular. The Hungarian examples (44) (in present tense) and (45) (in past tense) illustrate the alternation of a zero-2 strategy and a copula strategy.

Hungarian (Kenesei et al. 1998: 62, 78)

(44) *A lány diák*
 the girl student
 ‘The girl is tall/a student.’

(45) *Péter diák volt.*
 Peter student COP(PST)
 ‘Peter was a student.’

In short, PoS classes that employ a copula or a zero-2 strategy when used as the head of a predicate phrase are regarded as non-predicative. In contrast, PoS classes that use the zero-1 strategy are analyzed as including the function of head of a predicate phrase in their range of distributional possibilities.

However, several difficulties may arise when classifying the strategies used for non-verbal predication in individual languages. First, it may be the case that only *part* of the behavioural potential of verbal predicates

can be carried over to non-verbal predicates. This happens for instance in Abkhaz and Hdi. Non-verbal predicates in Abkhaz can take verbal suffixes that express present tense, past tense, and imperative mood. However, other verbal suffixes can be added to dynamic stems only. A copula intervenes when one of the latter suffixes must be expressed on a non-dynamic, non-verbal predicate. In Hdi nouns do not need a copula when they are used in predicative function and tense can be expressed on them. However, unlike verbal predicates, nouns cannot be marked for aspect. The reason for this discrepancy is that tense is marked by means of independent morphemes, while aspect marking involves different stem forms of verbs, and these are not available for nouns. In cases like Abkhaz and Hdi I have opted for the strongest possible definition of the zero-1 strategy: a non-verbal predicate is analyzed as employing a zero-1 strategy only if the full set of verbal behavioural potential can be expressed on the non-verbal predicate.

Second, it can be problematic to apply this criterion for the zero-1 strategy to isolating languages. In these languages, which lack inflection by definition, it is not possible to decide in favour of a zero-1 strategy (instead of zero-2) on the basis of the expression of verbal inflection on the non-verbal predicate. In such cases, there are two possible ways to distinguish between a zero-1 and a zero-2 strategy:

- (i) The possibility to add a copula;
- (ii) The possibility to combine the non-verbal predicate with freestanding particles that express verbal categories such as TAM.

For instance, Paiwan does not have a copula, and non-present tense can be expressed with non-verbal predicates by means of the freestanding particles *na* (for past) and *urbi* (for future), as is shown in (46) below. This is interpreted as evidence that there is no difference between verbal and non-verbal predication in terms of behavioural potential, i.e. that a zero-1 strategy is used in Paiwan.

Paiwan (Egli 1990: 61)

- (46) *Na/urbi sivitai ti kama*
 PST/FUT soldier FOC father
 ‘Father was/will be a soldier.’

In other isolating languages, such as Hmong Njua, Abun, Thai, and Nung, nominal predicates do take a copula. Under certain circumstances this copula can be omitted, which results in a zero-2 strategy. In contrast, adjectival predicates never take a copula. In addition, there appear to be no principled restrictions on the usage of grammatical particles with adjectival predicates. In (47) I give an example from Thai, which shows the expression of imperfective aspect with an adjectival predicate by means of the particle *yùu*. On the basis of these data, Thai is analyzed as having predicative adjectives. Similar evidence is available for the other isolating languages mentioned above.

Thai (Iwasaki & Ingkaphirom 2005: 155)

- (47) *aahǎan kamlaj rɔɔn yùu*
 food ADV hot IPFV
 ‘The food is still hot.’

It may be worthwhile at this point to stress the difference between (i) a situation in which adjectives can be used predicatively with a zero-1 strategy, and (ii) a situation in which there is no difference between adjectives and verbs. While the latter case involves bi-directional flexibility, in the former case there is only a unidirectional identity relation: all adjectives can be used as verbs, but not all verbs can be used as adjectives without structural coding⁸⁰. Rather, when a verb is used as a modifier in a referential phrase, it must be relativized, as in example (48), again from Thai:

Thai (Iwasaki & Ingkaphirom 2005: 243)

- (48) *Khon [thii duuleɛ] ni pen pen acaan lǎ*
 person REL take.care PRT COP COP teacher Q
 ‘Is the person who takes care (of the students) a teacher?’

Notably, the use of a verb in the function of modifier in a referential phrase does not always require an overt relativizer in Thai. Subject relative clauses that ascribe a general property to their head can also appear without the relativizer, as shown in (49). Nonetheless, the availability of the relativizer as a structural coding mechanism, even though it is not always needed, is interpreted as evidence that there is no complete overlap between verbs and adjectives in Thai.

⁸⁰ This amounts to what Stassen (1997) terms “*verby adjectives*”.

Thai (Iwasaki & Ingkaphirom 2005: 250)

- (49) *dèk* [*rian kèŋ*]
child study well
'a child who studies well.'

I will return to the issue of bi-directionality as a criterion for flexible PoS classes in the course of the following section, which is concerned more generally with the identification of lexical flexibility in the languages of the sample.

5.6 Identifying flexibility

5.6.1 Introduction

In Chapter 2, I discussed three criteria for lexical flexibility, as proposed by Evans and Osada (2005): (i) the criterion of *semantic compositionality*, (ii) the criterion of *exhaustiveness*, and (iii) the criterion of *equivalent combinatorics*. In this section, I present evidence for flexible PoS classes in the languages of the sample. As far as the data allow it, I organize the discussion along the lines of Evans and Osada's criteria. In section 5.6.2 I apply these criteria to languages that appear to display the most radical type of flexibility: a class of contentives that can be used in all propositional functions. In section 5.6.3 I focus on languages that display less flexible lexeme classes, i.e. classes that can express the functions of head and modifier in referential phrases (*nominals*), modifier in predicate and referential phrases (*modifiers*), or all three functions (*non-verbs*).

5.6.2 Languages with contentives

In some languages contentives are the only available lexeme class. This is the case in Tagalog and Kharia. Other languages combine a large class of contentives with one or more rigid PoS classes, which can but need not be small or derived. This is attested in Samoan, Kambara, Santali, and Guarani.

5.6.2.1 Semantic compositionality

I argued in Chapter 2 (section 2.5.2.2) that flexible languages are characterized by a mismatch between lexical categorization and syntactic categorization. Derivational processes pertaining to the former type of categorization typically involve semantically (and phonologically) unpredictable interpretations. Syntactic categorization of flexible lexemes, in contrast, yields compositional semantic meanings. The presumably flexible

languages discussed in Chapter 2 were shown to exhibit the latter type of categorization, and can thus be analyzed as behaving in accordance with Evans and Osada’s compositionality criterion. The relevant languages are Samoan, Tagalog, and Kharia. At the same time, however, these languages were shown to also display (zero-marked) processes of lexical derivation, with unpredictable semantic outcomes. In this section, I will further discuss this situation. In addition, I will show that similar situations are attested in Kambera and Santali.

For Kambera I have given examples in (28) and (29) above, showing that the insertion of presumably flexible lexemes into verbal versus nominal syntactic functions yields compositional semantic interpretations. However, there are also other types of processes in this language that are accompanied by more idiosyncratic meaning shifts. As illustrated in (50) below, the types of shifts are comparable to the ones discussed in Chapter 2. In particular, they involve meaning such as: ‘the typical cognate object of the action denoted by the root’, ‘a location associated with the action denoted by the root’, ‘the result of an action’, etcetera. As in Chapter 2, I argue that the examples in (50) are instances of lexical zero-derivation, whereas examples (28) and (29) illustrate ‘true’ flexibility.

Kambera (Klamer 1998: 110)

(50) Form	Nominal meaning	Verbal meaning
<i>hukung</i>	‘law/penalty’	‘punish’
<i>hindi</i>	‘attic’	‘dry/smoke’
<i>bāndil</i>	‘rifle’	‘shoot’
<i>dadu</i>	‘dice’	‘play dice’
<i>lindi</i>	‘bridge’	‘cross a bridge’
<i>ramuk</i>	‘pool/swamp’	‘be in a pool/swamp’
<i>tanda</i>	‘sign’	‘know’
<i>ludu</i>	‘song’	‘emit sound/sing’

Like the other flexible languages discussed in Chapter 2, Kambera also has overtly marked lexical derivations. In fact, these forms are only formally derived; the process itself is no longer productive. As would be predicted, formally derived forms can be used in nominal as well as in verbal syntactic environments. This is illustrated in (51a-b):

Kambera (Klamer 1998: 113)

- (51) a. *Ku-manàhal-nya na n-jala-nggu*
1SG.NOM-regret-3SG.DAT ART mistake-1SG.GEN
'I regret my sins.'
- b. *Eba! N-jala-mbu-nggunja-i-ka nú karwài*
EXCL be.wrong-also-1SG.CA-ITER-PFV DEIC just.now
'Darn! I was mistaken here too.'

On the other hand, *Kambera* also has derivational processes that produce verbal output forms, i.e. forms that are no longer flexible, but rather categorized in terms of phrase structure. I have illustrated these verbalizing processes in (40) and (41) above.

Recall from Chapter 2 that the flexible language *Kharia* has the possibility to insert complex phrases consisting of multiple roots or stems into verbal as well as nominal syntactic functions (see example (51) in Chapter 2). Interestingly, complex phrases in *Kambera* are flexible as well. Consider for instance example (52), in which the translational equivalent of an English NP functions as the head of a predicate phrase. The semantic interpretation is compositional: 'be X'.

Kambera (Klamer 1998: 107)

- (52) [*Tau mayila*]-*mbu-kai nyimi ná*
person poor-also-2PL.A you(PL) DEIC
'Moreover, you are also poor people.'

Similar examples can be found in *Samoan*, as is shown in (53): The complex phrase *atu mata tasi* 'one-eyed bonito' appears as an object argument with the possessive form *la'u* in (53a), and as a predicate with the future tense particle *ole'ā* in (53b). Again, the semantic interpretation of the phrase in these two functional environments is entirely predictable.

Samoan (Mosel 2004: 286, 287)

- (53) a. *fia maua la'u [atu mata tasi]*
want find my bonito eye one
'I want to find me a one-eyed bonito.'

- b. *Ia*, ‘*ole’ā* [*atu mata tasi*]
 well FUT bonito eye one
 ‘Was that going to be the one-eyed bonito?’

Consider now Santali, which resembles the other flexible languages in that it combines very regular meaning shifts with more idiosyncratic ones. The former type of shift, which characterizes ‘true’ flexibility, is illustrated in (54) and (55) below. First, examples (54a-b) show the compositional interpretation of action-denoting roots in the nominal functions of direct object (zero-marked for accusative), and dative-marked locative adjunct, respectively:

Santali (Neukom 2001: 17; Rau, forthcoming)

- (54) a. *Uni-then ədi rubet’-ijn jəm-akat’-ma*
 that-DAT much scold-1SG.SUBJ eat-PFV:ACT-IND
 ‘I got scolded badly by him.’ (lit: ‘I ate much scolding from him’)
- b. *Gapa-dɔ am-geg si-ok’-then*
 tomorrow-TOP you-FOC plough-MID-DAT
dangra-dɔ laga-əgu-kin-me
 bullock-TOP drive-bring-3DAT.OBJ-2SG.SUBJ
 ‘Tomorrow you shall drive the bullocks to where I am ploughing.’

The examples in (55) illustrate the compositional interpretation of semantically nominal lexemes in predicative function. In (55a) an individual-denoting lexeme is used as the head of a predicative phrase. Examples (55b-d) show that the same possibility exists for proper names, onomatopoeic forms, and complex NPs, respectively⁸¹.

Santali (Rau, forthcoming; Neukom 2001: 15)

- (55) a. *adɔ-e raj-en-a*
 then-3SG.SUBJ king-PST:M-IND
 ‘So he became king.’

⁸¹ See Neukom (2001: 15) for comparable examples with pronouns, numerals, adverbs, demonstratives, and quantifiers. Similar examples are provided by Peterson (2006: 62ff) for Kharia.

- b. *hɔpɔn-tɛt'-dɔ* *ɔnuə-a-e-a*
 son-3POSS-TOP Anua-APPL-3SG.OBJ-IND
 'The name of their son was Anua.' (lit.: Their son was Anua-ed.)
- c. *Bar pɛ dhao āā-y-en-a*
 two three time-3SG.SUBJ ONOM-y-PST:M-IND
 'It (the buffalo) groaned two or three times'

In contrast, there are also instances of more idiosyncratic meaning shifts. Some examples are provided in (56). These cases are interpreted as involving zero-marked derivation, rather than flexibility.

Santali (Neukom: 2001: 20)

(56)	Form	Nominal meaning	Verbal meaning
	<i>bicər</i>	'judgement'	'consider'
	<i>dhər-wək'</i>	'branch'	'send a branch' (to inform someone of the date of an event)
	<i>jɔm</i>	'food'	'eat'
	<i>bəhu</i>	'bride'	'take a wife for somebody'

Interestingly, Santali has the same overt lexical derivational process of nasal-vowel-infixation that was discussed for Kharia in Chapter 2 (see the examples in (52) of that chapter). However, in Santali the process is no longer productive. Some examples appear in (57) (note in passing that in Santali the process has an allomorph, namely *-t-vowel*):

Santali (Neukom 2001: 59)

(57)	Base	Lexical derivation
	<i>osar</i> 'to be broad'	– <i>o-no-sar</i> 'breadth'
	<i>sarec</i> 'to remain'	– <i>sa-ta-rec</i> 'remainder'
	<i>ɛhɔp</i> 'to begin'	– <i>ɛ-tɔ-hɔp</i> 'beginning'

As in Kharia, these derived forms remain syntactically flexible, even though their noun-like semantics suggest otherwise. This flexibility is illustrated in (58), where *ɛtɔhɔp* 'beginning' appears as the head of a predicate phrase:

Santali (Neukom 2001: 60)

- (58) *orak'-in* *etəhəp'-akat'-a*
house-1SG.SUBJ beginning-PFV:ACT-IND
'I have begun (to build) the house.'

Thus, the process exemplified in (56) is interpreted as the zero-marked counterpart of overt lexical derivation, as illustrated in (57). If this is the right analysis, then it is predicted that both the 'nominal' and the 'verbal' instantiations of the examples in (58) can be used nominal as well as verbal syntactic environments. Unfortunately, however, the available data are not conclusive on this point⁸².

Finally, as mentioned in section 5.3.4 above, there are some languages in the sample that seem to combine rigid classes of verbs and nouns with a class of flexible *heads* that can be used as both. The languages in which this is attested are Slave, Nivkh, West Greenlandic, and possibly Nunggubuyu. However, it appears that in all cases the meaning shifts displayed by the relevant lexemes are unpredictable, which disqualifies them as instances of flexibility. In (59) I give some examples of presumably flexible items in Nivkh.

Nivkh (Matissen & Drossard 1998: 62)

(59) Form	Nominal meaning	Verbal meaning
<i>k'u</i>	'arrow'	'shoot'
<i>ŋa</i>	'beast'	'pursue'
<i>kelma</i>	'step'	'step'
<i>k'ər</i>	'hunger'	'starve'
<i>mos</i>	'porridge'	'grind'
<i>əlv</i>	'roof'	'thatch'

Similarly, Sadock (2003: 79) notes for West Greenlandic that "*there is usually an incompletely predictable meaning relation between homophonous noun/verb pairs of stems.*" Some examples are provided in (60):

⁸² For Guaraní, there are unfortunately no clear data about the semantic relationship between nominal and verbal uses of flexible items.

West Greenlandic (Sadock 2003: 4)

(60) Form	Nominal meaning	Verbal meaning
<i>iga/igavooq</i>	'cooking pot'	'cook'
<i>imeg/imerpog</i>	'water'	'drink'
<i>kalleq/kallerpoq</i>	'thunder'	'be thundering'
<i>kuuk/kuuppoq</i>	'river'	'flow'
<i>niu/niuvooq</i>	'leg'	'get out of a vehicle'
<i>sianeq/sianerpoq</i>	'bell'	'ring up, telephone'
<i>siku/sikuvo</i>	'ice'	'be frozen over'

Finally, also in Slave there appear to be different types of unpredictable semantic shifts associated with the nominal and the verbal use of the same root. Some examples appear in (61)⁸³:

Slave (Rice 1989: 161)

(61) Form	Nominal meaning	Verbal meaning
<i>t'éh</i>	'charcoal'	'cook'
<i>tɛ</i>	'ice'	'freeze'
<i>seeh</i>	'saliva'	'spit'
<i>dzéh</i>	'gum'	'be sticky'
<i>tthih</i>	'axe'	'chop with an axe'

On the basis of the unpredictable semantics of the meaning shifts illustrated in (59)-(61), I argue that these are not to be analyzed as instances of flexibility. In addition, there is a crucial difference between the data in (61)-(63) from Nivkh, West Greenlandic, and Slave on the one hand, and the lexical zero-derivations discussed for the flexible languages on the other hand (see examples (50) and (56)). This difference is due to the separation in flexible languages between lexical and syntactic categorization. In flexible languages, such as Samoan, the two meanings pertaining to a particular form reflect lexical categorization only. Syntactically, i.e. in terms of phrase-structural possibilities, both zero-derived forms are flexible, i.e. can be inserted in verbal as well as in a nominal syntactic frames, with compositional semantic interpretations. In contrast, in rigid languages, such as Nivkh, lexical and syntactic categorization go hand in hand. Therefore, the lexical processes

⁸³ The case of Nunggubuyu is less clear. As I have no data for this language of the type discussed for Nivkh, West Greenlandic and Slave in (59)-(61), I postpone the discussion of Nunggubuyu to the next subsection.

which zero-derive the two unpredictably related meanings of a single root form *also* determine the phrase-structural possibilities of their output. In other words, any lexically categorized item is also syntactically categorized: it is either a noun or a verb, and can no longer be used in another function than its defining function, at least not without further measures (see also Don & Van Lier, forthcoming).

In sum, truly flexible languages can, in accordance with Evans and Osada's compositionality criterion, use simple roots, (zero-)derived stems, and complex phrases in more than one syntactic function without extra structural coding and with predictable semantic interpretations. The possibility of semantic compositionality is interpreted as a direct consequence of the fact that in flexible languages lexical and syntactic categorization do not coincide as they typically do in rigid languages.

In the next sub-section I address Evans and Osada's second criterion, which is concerned with the quantitative measurement of lexical flexibility.

5.6.2.2 Exhaustiveness

According to the criterion of exhaustiveness, 'true' flexibility requires that all members of a specific PoS class be flexible (see Chapter 2, section 2.5.3). Thus, if a language is claimed to have a class of simple contentives, then all content items belonging to that class should be usable in all relevant propositional functions.

I already indicated in Chapter 2 that my descriptive sources for Tagalog, Kharia, and Samoan are quite unambiguous about the fact that flexibility in these languages concerns all or almost all content words (depending on whether the class of contentives is combined with some (restricted) additional rigid PoS class). First, Himmelmann claims for Tagalog that

“(...)all Tagalog content words (both roots and derived words) are categorically indistinct, i.e. they may *all* occur in essentially the same basic syntactic positions.” (Himmelmann, 2007: 249; emphasis added, EvL)

For Kharia, Peterson writes the following:

“*Almost all* underived lexical morphemes in Kharia may be considered precatatorial in the sense that they may appear in referential, attributive and predicative function, and [...] there is no reason to assume that

this is any different for derived stems.” (Peterson 2006: 83; emphasis added, EvL)

Finally, for Samoan it is claimed that:

“The categorisation of full words into nouns and verbs is not given a-priori in the lexicon. [...] Derivations are more restricted in their usage; some derivations can occur only in verbal functions.” (Mosel & Hovdhaugen 1992: 73)

For the other languages in the sample that I have analyzed as displaying a class of contentives the data are somewhat less straightforward. Regarding Kambara, I explained earlier (see section 5.4.1) that Klamer (1998) distinguishes between nouns and verbs in this languages on the basis of differences in behavioural potential. However, if structural coding is the only criterion, then all lexemes in Kambara that are not adverbs or derived verbs apparently belong to a single class of contentives.

Regarding Guaraní and Santali, I have already mentioned that both combine a class of contentives with a class of rigid, un-derived verbs. (Nordhoff 2004: 56, 59 and p.c.; Neukom 2001: 13, 17). According to Neukom, about one third of all content lexemes in Santali are verbs; the rest are contentives. For Guaraní, no such quantitative information is available.

Finally, there are some indications of the proportion of roots that serve as the basis of both nominal and verbal zero-derivations in Slave, Nivkh, West Greenlandic and Nunggubuyu, as discussed above. Even though these cases are not regarded as cases of true flexibility, their degree of pervasiveness seems interesting to report. For Nivkh, Matissen and Drossard (1998: 61) remark that “*most root types are fixed with respect to lexical categories*”. This suggests that the class of un-categorized roots is at least restricted as compared to the classes of rigid nouns and verbs.

For West Greenlandic, Sadock points out that:

“The neat division of stems into two distinct part-of-speech classes is somewhat complicated by the fact that there are quite a few stems that occur both as a noun and a verb. Bergsland (1955) estimated that about 200 out of 1500 roots are noun-verb homophones, the rest being roughly equally divided between the two classes.” (Sadock 2003: 4-5)

possibilities. A lexeme like *susun* ‘dance’ in (62) needs to be relativized only if it is to receive the interpretation of ‘the one who is dancing’. As Evans and Osada remark themselves, all verbs can indeed be used directly in what they call “*clausal argument positions*”, and this yields a compositional semantic interpretation, in which the relevant lexeme still denotes an action rather than an object or an individual. Example (63) illustrates this:

Mundari (Evans & Osada 2005: 377)

- (63) *dub=ko laga-ja-n-a*
 sit=3PL be.tired.of-INCEP-INTR-IND
 ‘They are tired of sitting.’

However, in some cases other semantic interpretations are found. An example would be the root *jom*, which means ‘to eat’ in predicative function, and ‘food’ (rather than ‘eating’) in referential function. In fact, this pattern is very reminiscent of the situation in other flexible languages. In particular, I would argue that *jom* is a case of conversion: lexical, zero-marked derivation with an unpredictable semantic outcome. This is in contrast to the case in (63), which illustrates flexibility. In short, by adhering to the criterion of semantic compositionality as proposed in the previous subsection, we may also solve this potentially problematic aspect of the criterion of equivalent combinatorics.

A second aspect of equivalent combinatorics involves the claim that truly flexible items should be usable *to an equal extent* in multiple functions. As already pointed out in Chapter 2 (section 2.5.4), I disagree with Evans and Osada that each flexible lexeme should appear equally frequently in all propositional functions. In some of my descriptive sources, this issue of usage frequency is explicitly addressed. For instance Peterson, in his comment on Evans and Osada’s (2005) article, argues for Kharia that:

“[I]n fact, the only real restriction on this type of flexibility is what might be termed semantic compatibility. For example, the loan *tebal/tebul* ‘table’ can function as the complement of a predicate with the meaning ‘table’, in the middle voice with the meaning ‘become a table’ or in the active voice with the meaning ‘turn (something/someone) into a table’. Needless to say, objects seldom turn into tables, hence the predicative use of this lexeme is virtually never found in actual conversations.” (Peterson 2005: 396)

Crucially however, such predicative usage is not excluded for grammatical reasons. Peterson argues that a similar situation is likely to hold for Mundari, since speakers of this language accept constructions like the one in (64):

Mundari (Peterson 2005: 400)

- (64) *Sibonga am-e tebal-ked-me-a.*
God 2SG-3SG table-ACT.PST-2SG-IND
'God turned me into a table.'

In the same vein, Mosel and Hovdhaugen say the following about Samoan:

"Although certain full words seem to be used more as a verb or more as a noun phrase nucleus for semantic reasons, there are no lexical or grammatical constraints on why a particular word cannot be used in one or the other function." (Mosel & Hovdhaugen 1992: 73)

Regarding Tagalog, Himmelmann makes clear claims about the frequency with which action-denoting roots are used in nominal function (he terms this "*unaffixed uses*", i.e. without voice marking):

"It should be clearly understood that in general unaffixed uses of action roots are not in any way exceptional. Instead, they are reasonably common both in terms of types as well as in terms of tokens. [...] Although there are differences with regard to how frequent and natural it is for a given action root to occur without affixes, it is clearly the case that Tagalog action roots quite generally allow for unaffixed uses." (Himmelmann 2007: 284)

In sum, the available data strongly suggest that the languages I have analyzed as flexible indeed satisfy the criterion of equivalent combinatorics, as far as the *possibility* to use lexemes in multiple functions is concerned. As regards usage *frequency*, there are indications of asymmetry, but this only to be expected on the basis of semantic-pragmatic markedness.

5.6.3 Non-verbs, nominals, and modifiers

In this sub-section I focus on languages that are characterized by a certain degree of lexical flexibility regarding the following three functions: (i) head of a referential phrase, (ii) modifier in a referential phrase, and (iii)

modifier in a predicate phrase. As we have seen in section 5.2, there are some languages, in which these three functions are expressed by a single PoS class termed *non-verbs*. This is the case in Warao, Turkish, Kayardild, and (if small classes are included) Ma'di. Other languages have a class of *nominals* that can fulfil the two functions of head and modifier in a referential phrase. This is attested in Imbabura Quechua, Ma'di, and Hungarian. Finally, a number of languages have flexible modifiers, i.e. lexemes that can function as *modifiers* in both referential and predicate phrases. Below I discuss data from languages with non-verbs, nominals, and modifiers in turn.

5.6.3.1 Non-verbs

In Warao, according to Romero-Figeroa (1997: 49) “*there is no clear-cut distinction between nouns and adjectives; Warao nouns may function as attributives modifying other nouns.*” Romero-Figeroa illustrates this claim with examples involving lexemes that denote abstract properties, such as *yak* ‘beauty’ in (65):

Warao (Romero-Figeroa 1997: 50)

- (65) *Hiaka yak-era auka saba tainisa-n-a-e*
 garment **beauty**-AUGM for she buy-SG-PUNCT-PST
 ‘She bought a beautiful dress for her daughter.’

Note however, that it is not clear in how far object- and individual-denoting lexemes can also function without structural coding as modifiers in referential phrases. Moreover, there is no explicit evidence for bi-directional flexibility. That is to say, there are no examples in which items like *yak* are used without structural coding as heads of referential phrases (even though this possibility is expected on the basis of the ‘nominal’ translation of the lexeme).

In addition to the fuzzy noun-adjective distinction, the lexical flexibility in Warao presumably also includes the manner function. Example (66) below shows that the item *yak* ‘beauty’ can be used without structural coding as a modifier in a predicate phrase:

Warao (Romero-Figeroa 1997: 70)

- (66) *Tai hi saba yak-era ana tan-a-e*
 it you for **goodness**-AUGM NEG happen-SG-PUNCT-PST
 ‘It happened not so well for you.’

On the basis of these data, I have classified Warao as having a class of non-verbs, but admittedly the evidence is rather patchy.

Turkish has a class of non-verbs as well (Göksel & Kerslake 2005: 49-50). In Chapter 2 this was illustrated with the lexeme *güzel* ‘beauty’. Note that, like *yak* in Warao, Turkish *güzel* denotes a property concept. Therefore, it may be questioned in how far object-denoting lexemes have the same distributional possibilities. According to Göksel and Kerslake however, indeed many lexemes display noun-adjective flexibility, even though each lexeme has a ‘primary’ function in which it occurs most frequently for semantic reasons.

Finally, Kayardild has a large class of lexemes that can be used flexibly as heads and modifiers in referential phrases. Some examples are given in (67):

Kayardild (Evans 1995: 85, 238)

- (67) *maku*: ‘woman/female’
balumbanda: ‘westerner/western’
wurkura: ‘boy/male’
jambanda: ‘hollow log/hollow’

Such items are also productively used as what Evans (1995) calls ‘secondary predicates’. I regard secondary predicates as participant-oriented manner expressions (see Himmelmann & Schultze-Berndt 2005). This means that in Kayardild the functions of nouns, adjectives, and manner adverbs can be expressed by a single class of lexemes: non-verbs. Notably, there is also a small class of rigid manner adverbs in Kayardild (see figure 5.1 above). Evans interprets these as a subclass of nominals, because they take nominal morphology, as is illustrated in (68) below. However, he explicitly mentions the difference in distributional freedom between these ‘manner nominals’ on the one hand and non-verbs on the other: “*The distinguishing feature of this minor subclass is that it is restricted to one function, whereas the other nominal classes are versatile.*” (Evans 1995: 227; emphasis added, EvL).

Kayardild (Evans 1995: 259)

- (68) *Ngaɽwarri bukaɽwa-th*
thirsty.NOM die-ACT
‘(He) died thirsty.’

5.6.3.2 Nominals

Languages with non-verbs and nominals have in common that they lack a distinction between nouns and adjectives. As explained in section 2.5.4 of Chapter 2, this type of flexibility can be difficult to ascertain. In particular, it is often problematic to establish whether property-denoting lexemes can be referentially independent, or whether they rather function as modifiers of understood heads.

Hengeveld (1992: 63) argues for Imbabura Quechua that property-denoting lexemes can indeed be referentially independent. Example (69) shows the use of the property-denoting word *hatun* ‘big’ as the sole element of a referential phrase.

Imbabura Quechua (Schachter & Shopen 2007: 17)

- (69) *Rikaška hatun-ta*
see.PST-1SG big(PL)-ACC
‘I saw the big one(s).’

In contrast, Beck (2002: 146-149; see also Cerrón-Palomino 1987) claims that Imbabura Quechua does have a lexical noun-adjective distinction. The evidence Beck advances for this claim is that modifying lexemes with nominal semantics, such as *rumi* ‘stone’ in *rumi jan* ‘stone road’, can occur only once in a referential phrase, whereas modifying lexemes with ‘adjectival’ semantics are recursive. According to Beck, this makes constructions with ‘nominal’ modifiers look very much like compounds. Support for this analysis is found in the fact that when such ‘compounds’ function themselves as modifiers in a referential phrase, the only possible reading is the one where the combination of the two nouns in the ‘compound’ modifies the head. Thus, in *hara çakra rumi* ‘stone from a cornfield’ the combination *hara çakra* ‘corn field’ modifies the head *rumi* ‘stone’. On the other hand, as Beck acknowledges, there is no phonological evidence for a ‘compound’ analysis. On a more general note, Beck observes that, on the basis of isolated sentences in a descriptive grammar, it is quite hard to decide upon the issue of noun-adjective flexibility in any language. This is obviously true, and in the case of Imbabura Quechua it poses a serious problem. Nonetheless, for lack of conclusive evidence to the contrary, I maintain that Imbabura Quechua has a class of flexible nominals.

The second language with nominals is Hungarian. It was already briefly mentioned in Chapter 2, where I pointed out that Hungarian property-denoting lexemes can be referentially independent (Moravcsik 2001).

Examples (70a-b) and (71a-b) show that object-denoting and property-denoting lexemes can both function as heads and modifiers in referential phrases. Moreover, these usages can be distinguished from appositional constructions, which are also open to both semantic types of lexemes. This is illustrated in (72a-b)

Hungarian (Moravcsik 2001: 339)

(70) a. *A ház-ak-at látom.*
 the house-PL-ACC I.see
 ‘I see the houses.’

b. *A nagy-ok-at látom.*
 The big-PL-ACC I.see
 ‘I see the big ones.’

(71) a. *Ezr a bestia nő-t utálok.*
 this:ACC the beast woman-ACC I.hate
 ‘I hate this beast of a woman.’

b. *A kék ház-ak-at látom.*
 the blue house-PL-ACC I.see
 ‘I see the blue houses.’

(72) a. *A ház-at, a szülőhely-em-et, látom*
 The house-ACC the birth.place-SG1-ACC I.see
 ‘I see the house, my birthplace.’

b. *A ház-at, a kék-et, látom*
 The house-ACC the blue-ACC I.see
 ‘I see the house, the blue one.’

The third language in the sample with a class of nominals is Gooniyandi. McGregor states that in this language:

“[N]early all words which can realise the Entity role can also realise the role of Qualifier: in other words, words which can be referential can also be qualifying, and may indicate qualities or properties of things.” (McGregor 1990: 142).

It should be noted that the above quotation makes reference to unidirectional flexibility only: the use of entity-denoting lexemes (semantic ‘nouns’) in modifying function. With respect to the use of semantic ‘adjectives’ in referential function, things are slightly more complicated. As McGregor notes, for a number of property-denoting lexemes such usage has not yet been attested. However, according to the author, this is due to pragmatic markedness and “*limitations of the data*”, since many semantically similar (i.e. other property-denoting) lexemes do occur as heads of referential phrases. Therefore, I conclude with McGregor that there is insufficient evidence to justify a noun-adjective distinction in Gooniyandi.

Noun-adjective flexibility in Ma’di has already been addressed in section 5.4.1. It seems that in this language property-denoting lexemes can indeed be referentially independent. Note that, in addition to nominals, Ma’di has a small, closed class of what appear to be flexible non-verbs. These are property-denoting lexemes, which can be used as modifiers in both referential and predicative phrases (and also as heads of referential phrases if one assumes that there is no lexical noun-adjective distinction). Some examples are given in (73)⁸⁴:

Ma’di (Blackings & Fabb 2003: 125)

(73)	Form	Adjectival meaning	Adverbial meaning
	<i>pélére</i>	clean	properly
	<i>l̩s̩̩</i>	good	well
	<i>r̩r̩̩</i>	quick	quickly
	<i>tʃé tʃé</i>	slow	slowly

At this point, it may be worthwhile to consider, for the sake of contrast, some examples of languages that may seem to display noun-adjective flexibility, while in fact the absolutive use of adjectives requires the presence of an understood head in the discourse context. In these languages, evidence for such an understood head comes from the fact that it triggers agreement on the absolutive modifier. In Chapter 2, I already gave an example of this phenomenon from Spanish. Abkhaz is a similar case. Chirikba (2003: 29) remarks that in this language it may be difficult to distinguish adjectives from nouns, when the latter are used as the sole element of a referential

⁸⁴ It is not clear whether these lexemes would remain property-denoting (i.e. become abstract nouns such as ‘cleanliness’) or would rather be interpreted as an object/individual with the property denoted by the base form (‘clean one’) when used as the head of a referential phrase.

phrase and combine with determiners and number markers. According to Hewitt (1979: 226) however, the morphological form of an absolutive adjective is determined by the unexpressed noun to which it functions as an attribute. Also in Basque, absolutive use of adjectives is “*exceedingly common*”. However, the animacy marking on the adjective is dependent on the animacy status of its understood head. For instance *beltza* ‘the black one’ takes animate morphology when the understood head is a horse, but inanimate morphology when it is a skirt (Hualde & Ortiz de Urbina 2003: 149-150).

5.6.3.2 Modifiers

I will now discuss what can be regarded as the most marginal form of lexical flexibility, involving the two modifier functions. As already mentioned, there are some languages in which flexible modifiers constitute quite large, open classes, while other languages rather have small or derived modifier classes⁸⁵.

As a first example, consider Lango, a language in which modification in predicate and referential phrases is expressed in the same way: with the linking particle *à* and a member of the class of flexible modifiers (Noonan 1992: 181). Notably, this class of modifiers exists alongside a class of rigid manner adverbs, which take the particle *nî*, as illustrated in (74) below. Noonan (1992: 181) explicitly mentions the difference between flexible modifiers and rigid manner adverbs: “*Unlike type 1 [i.e. modifiers, EvL], these forms [i.e. manner adverbs, EvL] can never have another grammatical function*”.

Lango (Noonan 1992: 181)

- (74) *ɲámó* *cèm nî mɔ̀ðk- mɔ̀ðk*
 3SG.SUBJ.chew.HAB food PRT noisily
 ‘He chews (his) food noisily.’

⁸⁵ The derived class of modifiers in Turkish had been mentioned several times already. It is interesting to note that, although *-CA* derivations can be used as modifiers in both predicative and referential phrases, they are always event-oriented (also when used attributively). Attributive use is only possible with nominal bases (notional nouns or nominalizations) that describe a result, or can at least take a resultative interpretation. The attributively used *-CA* form expresses the manner in which the action has been performed, and stands in opposition with a non-derived attribute, which simply predicates a property of the resultative noun, without specifying the orientation of this property:

<i>Akıllı-ca</i>	<i>birplan</i>	<i>Akıllı</i>	<i>birplan</i>
smart-MOD	a.plan	smart	a.plan
‘A smartly made plan.’		‘a smart plan.’ (Schroeder 2004: 201)	

Also in Ket, most modifying lexemes belong to the same flexible class. However, there are a few lexemes that can be used as adjectives only, and some others which are used exclusively as manner adverbs (Werner 1997: 119-120; Vaida 2004: 40).

In Itelmen, modifying lexemes can be divided into a morphologically regular class and a morphologically irregular class (Georg & Volodin (1999: 106). The members of the former group have different forms when they are used as modifiers in referential versus predicate phrases: In the former function they take the suffix *-lab*, and in the latter the suffix *-q*. This is illustrated in (75a) and (75b), respectively:

Itelmen (Georg & Volodin 1999: 111, 112)

- (75) a. *K-k'ol-knen* *newen as-lab* *nejŋ-anke*.
 INF.III>COME<INF.III DEM high-ADJ mountain-DAT
 'He came to a high mountain.'
- b. *A t'salaj iseŋ-q* *saŋte-s-kinen*
 but fox quiet-ADV follow-PRS-3.SG:3.SG.I
 'But the fox was following him quietly.'

However, Itelmen also has a small, closed group of lexemes that are not morphologically recognizable as adjective or adverb: They have the same form in both functions. This is illustrated in (76a-b) with the lexeme *miça* 'beautiful':

Itelmen (Georg & Volodin 1999: 119)

- (76) a. *Kninə-kit miça mimsx t-ənnksə-çen*.
 2.SG-CAUS beautiful woman 1.SG-let.go-3.SG.PAT
 'Because of you I have let go of a beautiful woman.'
- b. *E, çasit miça sun-s-ç*
 INTERJ now beautifully live-PRS-2.SG
 'Oh well, now he is doing well.'

Finally, Thai also has a small class of flexible modifiers, existing alongside rigid adjectives and rigid (derived) manner adverbs. An example of a flexible modifier is *dii* 'good, well', which can be used as an modifier in a referential phrase, as in *kbon dii* 'good person', and as a modifier in a predicate phrase,

as in *tham dii* ‘do well’. (Iwasaki & Ingkaphirom 2005: 92). A similar system attested in Abun has already been discussed in section 5.2.2.

5.6.4 Summary

In this section I presented data in favour of lexical flexibility in the relevant languages of the sample. First, I discussed languages with contentives in terms of Evans and Osada’s three criteria for flexibility. In line with the discussion in Chapter 2 (section 2.5), I have proposed that the crucial characteristic of a ‘truly’ flexible language lies in the separation of lexical and syntactic categorization. This allows for the use of basically all content material in all syntactic functions, with semantically compositional interpretations. The only restriction on this flexibility concerns differences in frequency of use, which can be explained in terms of relative semantic-pragmatic markedness rather than absolute grammatical constraints.

Second, I discussed types of lexical flexibility with respect to the functions of head and modifier in a referential phrase, and modifier in a predicate phrase. In general, it seems that there has been less debate in the literature – both theoretical and typological – regarding these types of lexical flexibility. As a result, the kind of argumentation that should support the identification of non-verbs, nominals, and modifiers in actual languages is less developed, and sources provide less data that are relevant to the issue. This also has repercussions for the thoroughness of the discussion of these issues in the above section.

5.7 Conclusion, outlook

In this chapter I have presented the data on PoS classification in the languages of the sample, and evaluated them in light of the predictions and theoretical issues of Chapter 2.

In the next chapter I turn to the classification of the dependent clause constructions in the sample languages, within the functional-typological framework developed in Chapter 3.