

## Contents

<b>0</b>	<b>Introduction</b>	<b>1</b>
0.1	Aims and general strategy . . . . .	1
0.2	Guaraní . . . . .	2
0.3	Theoretical background . . . . .	3
<b>1</b>	<b>The verbal person system of Guaraní</b>	<b>5</b>
1.1	The person prefixes . . . . .	5
1.2	Reference hierarchy . . . . .	7
1.3	Conjugational classes . . . . .	8
<b>2</b>	<b>An integrational analysis of the person system of Guaraní</b>	<b>9</b>
2.1	Traditional conceptions based on Latin grammar . . . . .	9
2.2	Beyond Latin grammar . . . . .	11
2.3	Transitive verbs . . . . .	12
2.4	Reflexive, reciprocal, desiderative, and command . . . . .	14
<b>3</b>	<b>Structural and functional systems, and the system connection</b>	<b>15</b>
3.1	The structural system . . . . .	15
3.2	The functional system . . . . .	16
3.3	The system connection . . . . .	17
3.4	From the description of structures to the description of units . . . . .	22
<b>4</b>	<b>Inflectional units contained in forms</b>	<b>23</b>
4.1	Inflectional units . . . . .	23
4.2	An example . . . . .	23
4.3	Being contained . . . . .	24
<b>5</b>	<b>The functional effects of inflectional units</b>	<b>26</b>
5.1	Containing sets and inflectional-unit-categories . . . . .	26
5.2	Marking pairs of inflectional units and their description . . . . .	29
5.3	Marking effects of inflectional units — specificity of inflectional units and language types . . . . .	30
	<b>References</b>	<b>34</b>

# **Inflectional Units and their Effects: The Case of Verbal Prefixes in Guaraní**

Sebastian Drude

Freie Universität Berlin (Germany)

## **0 Introduction**

### *0.1 Aims and general strategy*

With the present essay I pursue a threefold aim as will be explained in the following paragraphs. Since I cannot expect my readers to be familiar with the language studied, Guaraní, more information about this language will be given in the next subsection.

During the last years I tried to develop an adequate description (especially a lexicographic description) of the Guaraní prefixes that are used with verbal stems in predicative expressions to ‘cross-reference’ the semantic agent and/or patient of the activity or state being expressed by the predicate. In doing so I eventually achieved a new categorial structuring of the person system of this language. The (practical) main purpose of the present essay is to present this new structuring to readers who are studying this language or other languages with similar conditions.

In the theoretical field I developed definitions of a couple of concepts that may be useful for describing inflectional affixes and auxiliaries in arbitrary languages. Presentation of these concepts is the second major aim of the present paper.

Since these concepts as well as the new analysis presuppose the Integrational Theory of Language that is being developed in the general framework of Integrational Linguistics (IL), I will first give a brief introduction to those parts of this theory that are relevant to understand the subsequent sections of my paper and go beyond the general introduction to basic concepts of Integrational Linguistics given in the first chapter of this book (p.5ff). Furthermore, this overview is justified by the fact that H. Lieb’s (1992) essay on the concept of paradigm has been published in German. This paper is of significant importance in more than just the area of interest of the present essay.

After giving a brief general description of Guaraní, the last subsections of this introduction establish a basis for the subsequent sections. These will be of minor interest to readers familiar with IL. In the next sections I describe the relevant facts of Guaraní (sec. 1) and try to outline an integrational account of these facts

(secs 2 and 3). They should mainly be of interest to linguists who investigate Tupí–Guaraní and similar languages but also show the concepts of sec. 0.3 ‘at work’.

Finally, the last two sections (secs 4 and 5) extend the theory in order to facilitate the description of inflectional affixes and auxiliary words within the framework of IL. In this third part (with a higher level of technicality), data from the previous sections is used by way of example but no new facts nor new accounts of facts are introduced. Therefore, readers with an interest in Guaraní only may safely omit reading this part.

## 0.2 *Guaraní*

With ‘Guaraní’ I refer to modern Paraguayan Guaraní which is one of the largest indigenous languages of South America. It is the only one that is spoken by the vast majority of a non-indigenous society (only 5–10 % of the population of Paraguay are native speakers of Spanish).

Together with some minoritarian Guaranian languages spoken by native groups (Mby’á, Kaiwa, Chiriguano, etc.) it forms one of eight subgroups in the large Tupí–Guaraní family of languages (Rodrigues 8586). All other subgroups have been and still are located in the Amazonian basin and almost everywhere along the Brazilian coast. For this and for other historical reasons, Tupí–Guaraní languages were among the first to be studied and used as *linguae francae* by the Europeans and are nowadays the best studied language family of the lowland part of South America, mainly due to the work of Aryon Rodrigues dall’Igna and his followers.

Guaraní, like other Tupí languages, is most often considered to belong to the agglutinative type of languages. Normal ‘units of speech’ (which can be word forms or more complex entities, so-called ‘macrosegments’ in the terminology of Gregores & Suares 1967) usually consist of a stem surrounded by a few prefixes and many suffixes. The prefixes are more paradigmatic and mostly serve as person markers. The verbal prefixes are the subject matter of the present paper. The suffixes are more freely combinable and are used as markers for tense (even with substantives), mode and other pragmatically determined features. There is no obligatory order of the main parts of a clause, but subject - predicate - object is a very frequent type, maybe due to the influence of Spanish. Subordinating particles such as postpositions are found at the end of the subordinated constituent.

In this paper I will use the excellent grammar by Emma Gregores and Jorge A. Suares (Gregores & Suares 1967) as a basis and even use their examples. However, I adapted them to the modern Paraguayan orthography although this continues to differentiate between allophones of certain phonemes, which is avoided in the phonemical orthography of Gregores / Suares (Gregores & Suares). For the convenience of the reader I separate different morphemes by a hyphen as in *a-guata* instead of the usual *aguata*. The pronunciation rules are as follows (all letters not

listed here are pronounced in the same way as the corresponding IPA symbols):

⟨Y⟩ represents a high unrounded central vowel, varying between [i̠ ~ ī̠ ~ ɨ̠]. All six vowels /i, i̠, u, e, a, o/ and many consonants have nasal allophones that occur inside a ‘nasal span’ (see below). The apostrophe ⟨'⟩ represents the glottal stop [ʔ], the pronunciation of ⟨H⟩ varies between [h], [x] and [ç].

⟨M⟩ and ⟨MB⟩ represent two allophones [m] and [ᵐb], respectively, of a single phoneme /m/, and the same holds analogously for the phonemes /n/ and /ɲ/: ⟨N⟩ stands for [n] and ⟨ND⟩ for [ᵐd], both allophones of /n/; ⟨Ñ̃⟩ represents [ɲ] and ⟨J⟩ the phones [ᵈɲ] and [ɲ], all allophones of /ɲ/. The first allophones of each pair are used in nasal spans, the other one(s) in oral surroundings. /ŋ/ has two allophones [ŋ] (nasal) and [ᵑg] (oral), too, but it is represented by ⟨NG⟩ everywhere.

⟨V⟩ varies between [v] and [ʋ] (both occur also in nasalized variants), ⟨R⟩ represents a single flap [ɾ] (also nasalized), and ⟨G⟩ is pronounced as (nasal / oral) [ɣ] or [ʉ]. Finally, ⟨CH⟩ stands for the alveolar fricative [ʃ], not for the corresponding affricate (as in Spanish).

A major systematic feature in Guaraní is nasality. It is a ‘long component’ that stretches from the ‘core’ (a nasal consonant or a stressed nasal vowel) mostly to the left over several syllables. In each of the affected nasal syllables any vowel or sonorant is represented by its ‘nasal’ allophone. As Gregores & Soares (1967:65ff) put it, such a sequence of nasal syllables is either a nasal span, caused by a nasal accent (that has the nasalized syllables as its domain), or it is caused non-phonemically by one of the nasal consonants /n, m, ŋ, ŋʷ/ (no matter whether these occur as an oral or nasal allophone). The domain of the long component is limited to the left by a (different) accent group (that may be nasal or oral), or by a word boundary. Usually the stressed nasal vowel that is the core of the nasal span is its last vowel, but if the nasal span is the last accent group of a word, nasality may spread to unstressed syllables to the right up to the end of the word.

In writing, a nasal accent is represented by a tilde ⟨˜⟩ above the core vowel of the span. The stressed vowel of an oral accent group will be marked by an acute ⟨´⟩ if it is not the last vowel of the word. In the latter case it is not indicated at all. That means that all words without a graphical accent have their main stress on the last syllable.

### 0.3 Theoretical background

It is a characteristic feature of Integrational Linguistics that it is based substantially on linguistic tradition by attempting to reconstruct traditional conceptions in a formally consistent way. One central concept of traditional grammar, widely ignored in many current linguistic approaches, is the concept of *paradigm*. Under an IL perspective, (syntactic) paradigms rely on word forms and on categories (sets) of such forms (see S3).

The functional categories such as Singular, First Person or Nominative that appear in paradigms (viz., as elements of the second component of each element of any paradigm) are given by the SUO (as explained in S2).

In order to account for the overall internal structure of paradigms we need *bases for paradigms* as introduced in Lieb (1992b) that consist of six components. These may be very briefly characterized as follows:

1. The *starting set* is a basic constituent categorie (typically, Noun Form (-,S) or Verb Form (-,S)) that is the source of two proper parts of the SUO, viz. the functional and structural classification systems.
2. The *functional classification system* yealds the functional categories mentioned above as its endpoints.
3. The *structural classification system* determines formal (or structural) categories, that is, sets of forms that share a property of their inner composition, such as containing a certain affix or auxiliary word - or a form of a stemlexeme that belongs to a certain morphological categorie (see M2 and M3).
4. The *system connection* ensures that each form of a paradigm is assigned to the right functional categories.

The two systems and their connection take care of the right composition of each pair that is an element of a paradigm. The last two components of bases for paradigms have to ensure that only formally and semantically related forms are actually put together in a single paradigm.

5. For the formal aspect we need a *basis for compatibility* that ensures that complex forms (sequences of more than one member) can be reduced to their main parts and that these main parts can be compared with each other with respect to occurrence of – possibly different – forms of the same stem.
6. For the semantic aspect we need an *external relation* between lexical meanings and forms. Only forms with the same meaning will actually appear in a given paradigm.

I will now discuss the system connection (no. 4, above) in greater detail.

#### *The system connection*

The formal representation of the interaction between the structural and the functional systems is provided by the *system connection* ('Systemverbindung' in German), a relation between sets of structural categories and sets of functional categories. It allows (i) determination of each functional category by referring to form categories and (ii) to answer the question whether a given form (build in a

certain manner) combines with a given set of functional categories (Lieb 1992, p. 10).

Note that the system connection relates *sets* of formal to sets of functional categories. This is because a single formal property often doesn't allow one to determine whether the form belongs to a certain functional category, whereas a *combination* of properties is sufficient. Take, for instance, the case of three German composite forms (I was unable to find an easy example in English, where not much of inflection is found; Categorical names which are not needed in the subsequent discussion are abbreviated):

- (1) a. *wird sehen* ('will see') : 3.Ps, Sg, Active, Ind, Future 1,
- b. *wird gesehen* ('is seen') : 3.Ps, Sg, Passive, Ind, Present,
- c. *hat gesehen* ('has seen') : 3.Ps, Sg, Active, Ind, Perfect.

Now consider the following two structural categories of German idiolect systems (names of structural categories are enclosed in square brackets to differentiate them easily from functional ones):

- (2) a. [*werden*<sub>V</sub>] = the set of all verb forms that contain a form of the auxiliary *werden*,
- b. [Participle] = the set of all verb forms whose main part is the participle perfect (the main part is that part of a form in which a form of the stem occurs).

The second form *wird gesehen* belongs to the functional category *Passive*, cf. (1 b). This fact can not be accounted for by referring to the membership of *wird gesehen* to [*werden*<sub>V</sub>] alone: the active form *wird sehen* (future) also belongs to this class. Neither is it sufficient to point at the belonging of *wird gesehen* to the class [Participle], as it is the same in the case of the active perfect form *hat gesehen*. Thus, a set containing at least *both* structural categories is needed to be able to assign the functional category *Passive*. (In fact, not even these two categories are sufficient since *wird gesehen haben* ('will have seen') belongs to both and still is an active form.)

Even with these scarce explications it seems reasonable to state that the system connection plays a key role in construing paradigms and should figure, in one form or another, in the description of any language.

## 1 The verbal person system of Guaraní

### 1.1 The person prefixes

The following presentation of the facts in Guaraní grounds on the grammar of E. Gregores and J. Suares (Gregores & Suares 1967). They set up three major

classes of prefixes for the person system proper, that is, prefixes that ‘cross-reference’ or express the grammatical subject or object of the predicate — as *-s* is used for ‘cross-referencing’ the third person singular in English. I here adhere to this term, also used by Cheryl Jenssen (1990:passim), who treats the same subsystem of Tupí–Guaraní languages in the context of reconstructive studies. (Sometimes, ‘indicate’ or ‘mark’ are used for the same relation.)

In doing so we are confronted with a terminological problem: In describing the characteristics of these prefixes, we are forced to use either terms for grammatical relations such as “subject” or “object”, or terms for semantic roles such as “agent” and “patient”. Some authors even use both, for instance in order to differentiate between the subject of transitive and intransitive verbs. Since I have not been able to offer any better terms I will use “subject” and “object”, too. Consequently, in this article these terms don’t (or only indirectly) refer to the grammatical relations but to the categories of verb forms that stand in agreement to the corresponding external constituents. It should be clear, though, that the prefixes themselves do *not* stand in the subject or object relation to the verb stem as may be seen in

- (3) a. *che* *che-* *mandu’a*                      b. *ha’e* *o-* *guata*  
       *I*        *1Sg-* *remember*                      *he*    *3-* *go*  
               ‘I remember.’    ‘He goes.’

where the free form *che* – a pronoun meaning “I” – is the (external) subject of *che-mandu’a*, and *ha’e* the subject of *o-guata*. The subject constituent is most often omitted by ellipsis.

The first of the three major classes is called *personal reference* by Gregores & Soares (1967:130f) and contains six prefixes that belong to the following categories (as they put it):

- (4) a. speaker (first person singular, 1Sg) — *che-*  
       b. addressee (second person singular, 2Sg) — *nde-/ne-*  
       c. neither speaker nor addressee (third person, 3) — *i-* (with many allomorphs: *i-*, *h-*, *ij-*, *iñ-*, *hi-*)  
       d. speakers and others including addressee (first person plural inclusive, 1Pl<sub>i</sub>) — *ñande-/ñane-*  
       e. speaker and others excluding addressee (first person plural exclusive, 1Pl<sub>e</sub>) — *ore-*  
       f. addressee and others (second person plural, 2Pl) — *pende-/pene-*

The prefixes of this set (which I will call “Set One”) are used in the forms of certain intransitive verbs, few in number and often called “chendal verbs” in studies on Guaraní, such as *mandu’a* ‘(to) remember’. They are identical with the set of

possessive noun-prefixes and are a subset of the prefixes of Set Three, below (with the exception of *i-* that does not occur in Set Three). In our analysis, each prefix exists only once but has different *effects* when used with forms of stem lexemes that belong to different classes.

The same six ‘categories’ are described for the second set, called “subject” by Gregores & Soares (1967:130):

- (5) 1Sg: *a-*, 2Sg: *re-*, 3: *o-*, 1Pl<sub>i</sub>: *ja-/ña-*, 1Pl<sub>e</sub>: *ro-*, 2Pl: *pe-*

The prefixes of this class (Set Two, for short) occur with the majority of stems of transitive and intransitive verbs: *a-guata*: ‘I walk’ (*guata*: ‘(to) walk’).

The third class, “object”, “has seven specific categories, four of which express also person of subject” (Gregores & Soares, 1967:131):

- (6) 1s: *che-*, 2s with subject=3: *nde-/ne-*, 2s with subj=1: *ro-*,  
1pi: *ñande-/ñane-*, 1pe: *ore-*, 2p with subj=3: *pende-/pene-*,  
2p with subj=1: *po-*.

*Examples*: *che-juhu*: ‘he (she, it, they, you) meet(s) me’ (*juhu*: ‘(to) meet’); *pende-juhu*: ‘he (she, it, they) meet(s) you(pl)’; *po-juhu*: ‘I (we excl.) meet you(pl)’.

Quite obviously, five of these prefixes also figure in the first set (put differently: all prefixes but *i-* of Set One reappear in the “object” set). The remaining two, *ro-* and *po-*, both involve the speaker as subject and the addressee as object. (In the case of *ro-*, addressee is the only object, in the case of *po-*, together with others. Number of subject is irrelevant.) This is no coincidence: In many Tupí–Guaraní languages we find a phenomenon called “reference hierarchy”.

## 1.2 Reference hierarchy

### *Prefixes of transitive verbs*

In an instructive article called “Hierarquia Referencial em Línguas Tupí”, R. Monserrat and M. Facó Soares (1970) show that in many languages of this language stem, especially in the Tupí–Guaraní family, transitive verbs may take a prefix that cross-references the subject or cross-references the object, but not two such prefixes at a time (the ‘mixed prefixes’ *ro-* and *po-* and their equivalences will be discussed below).

Thus, there has to be a criterion for selecting one of the twelve possible prefixes. This criterion is called the *reference hierarchy*. In the reference hierarchy, the speaker takes precedence over the addressee, and the addressee over others. Spelled out: If the speaker is involved, alone or with others, in the course of events expressed by the verb, a first person prefix will go with the verb stem. If the speaker is subject, *a-*, *ja-*, or *ro-* will be chosen, if he is object, either *che-*, *ñande-*, or *ore-* will occur. (From here on, morphemes with a nasal and an oral form will be represented by the oral allomorph, as in *ja-* (*ña-/ja-*) or *ñande-* (*ñan(d)e-*).



A second person prefix from either set may only occur if the speaker is *not* involved but the addressee is. If the addressee is subject, *re-* or *pe-* is chosen, if he is object, *nde-* or *pende-*.

If neither speaker nor addressee are involved, only the third person prefix *o-* is chosen to go with the verb stem. For this reason, the third person prefix *i-* of Set One is not included among the prefixes for transitive verbs. In the context of reference hierarchy, this is stated by some authors (but not by Gregores & Soares 1967) to be an additional rule of reference hierarchy, determining that in case of equal valence of subject and object, the subject prefix will be employed. In fact, in case of reflexivity or reciprocity only Set Two prefixes are found. (See below, subsec. 1.3.)

#### *Mixed prefixes*

Facó Soares and Monserrat argue (1970:164f) that a non overlapping prefix system specifying only either subject or object in accordance to reference hierarchy had been present in the protolanguages which are still being reconstructed, that is, in the early proto-Tupí and probably also in later proto-Tupí-Guaraní. (In turn, A. Rodrigues and his follower Ch. Jensen (1990) reconstruct mixed prefixes for proto-Tupí-Guaraní.)

If we encounter exceptions from the stated above rules (as is the case in virtually all Tupí-Guaraní languages), according to Facó Soares and Monserrat it should be analyzed as a dissolution rather than an incomplete development, of reference hierarchy — a dissolution that started with a competition between first and second person when the former was subject and the latter object. For this case, they postulate a stage in which both prefixes were present and assume that the new (combined) prefixes have undergone functional and sound changes that lead to the additional forms *ro-* and *po-* of the third set ((6), “object” in the grouping of Gregores & Soares 1967).

### **1.3 Conjugational classes**

The reference hierarchy is obviously relevant only in case of transitive verbs. Put differently, a certain set of stems occurs with prefixes of both Set Two and Set Three (a morphological criterion), and this set coincides with the set of stems of transitive verbs (a criterion based on syntactic facts).

Intransitive verbs (more exactly, their stems) are divided in two classes: most take the prefixes of Set Two (*a-*, *re-*, etc.), but an important class of so-called quality verbs (cf. Gregores & Soares 1967:137) take the prefixes of the first class only (*che-*, *nde-*, etc.). The resulting predicate constituents show a certain similarity with nominal constructions: *che-mandu'a nderehe* (*mandu'a*: ‘(to) remember’/‘memory’, with postpositional group *nde-rehe*) could be translated as ‘my memory of you’ or as ‘I remember you’. Still, the difference between nominal and

verbal constructions exists as becomes apparent when different temporal suffixes are added, when the predicate is negated by the circumfix *n(d)a-...-i*, or when the verbal desiderative prefix *ta-* (several allomorphs) stands before the verb, which is impossible with the noun. This leads us to the remaining four prefixes that will be discussed in the present article.

### *Other prefixes*

As shown above, all prefixes introduced so far are mutually exclusive. Instead of one of these, the prefix for command, *e-*, may go with the verb stem: *e-guata*: ‘walk!'; *e-juhu*: ‘meet (it/him)!’.

A prefix with a similar but somewhat ‘softer’ meaning, called “desiderative” by Gregores & Soares (1967:132), is *ta-* (it has many allomorphs due to a kind of vowel harmony with the following prefix). Differently from all other prefixes mentioned up to now, *ta-* may stand *additionally* in front of any of the person prefixes (except *e-*), as it may be seen in the following examples:

- (7)      a.    *t-*        *o-*        *guata*  
              (*desid*)    *he*        *go*  
                  ‘may he go!’
- b.    *ta-*        *che-*        *mandu'a*  
              (*desid*)    *I*         *remember*  
                  ‘I hope I remember’

The prefix for reflexive and passive, *ñe-/je-*, enters *between* a prefix of Set Two and a transitive verb stem: *a-ñe-hendu*: (*hendu*: ‘(to) hear’) ‘I hear myself’/‘I am heard’; *ña-ñe-hendu*: ‘we hear ourselves’ (I hear myself and you hear yourself) / ‘we are heard’ (somebody/something hears you and me).

Finally, the prefix for reciprocity, *jo-*, occurs between a ‘pluralic’ prefix of Set Two (including third person: *ja-*, *ro-*, *pe-*, or *o-*) and a transitive verb stem: *ja-jo-juhu*: (*juhu*: ‘(to) meet’) ‘You and I, we meet us (one another)’; *o-jo-juhu*: ‘they meet’. *jo-* is mutually exclusive with *je-*.

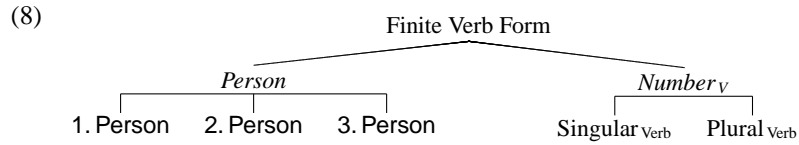
We see: all prefixes (except *ta-*) involve subject, and many specify object as well. A unified theoretical description of this system, taking the reference hierarchy into account, will be given in the next two sections.

## **2 An integrational analysis of the person system of Guaraní**

### *2.1 Traditional conceptions based on Latin grammar*

All descriptions of Guaraní that I am familiar with use traditional terminology when describing the person system. In traditional grammar two classifications on (finite) verb forms are assumed: one (called “person”) that yields the three classes 1. Person, 2. Person, and 3. Person, and another one (“number”) that has as its classes (i.e., its elements) Singular<sub>V</sub> (“Singular for verb forms”) and Plural<sub>V</sub> (in some languages further elements, like Dual). From the point of view of IL, the two

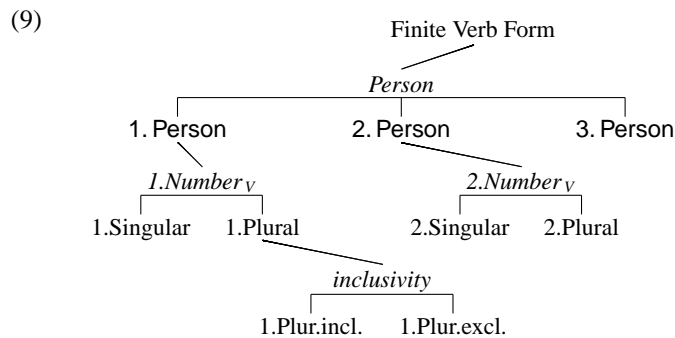
classifications are cross-classifications on the set of finite verb forms (or a subset thereof) and may be represented by the following graph:



Under an IL view, these two classifications (or similar, possibly with unmarked categories) are elements of the syntactic unit ordering in most European languages.

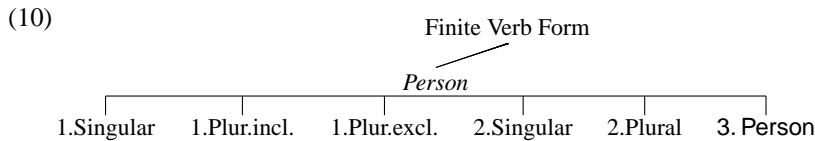
In Guaraní we have two facts that don't fit into this conception: (i) the third person has only one single form for singular and plural (and this doesn't seem to be due to syncretism), and (ii) the so-called first person plural has two forms: one including and one excluding the addressee.

If we tried to account for these facts using the traditional classifications, we would have to create a subclassification for inclusivity — but a subclassification of which set? A solution like the following may be envisaged:



However, such a system shows many inadequacies: (i) We would need two analogous classifications for number; (ii) we would gain many classifications, all containing at least one class with forms with only one prefix of each, Set One and Two; (iii) the reference hierarchy couldn't be explained referring to endpoints of the system, and (iv) in general, the resulting system would be 'deep' (using three levels of sub-classifications) – an effect of this being usually that we need more classifications.

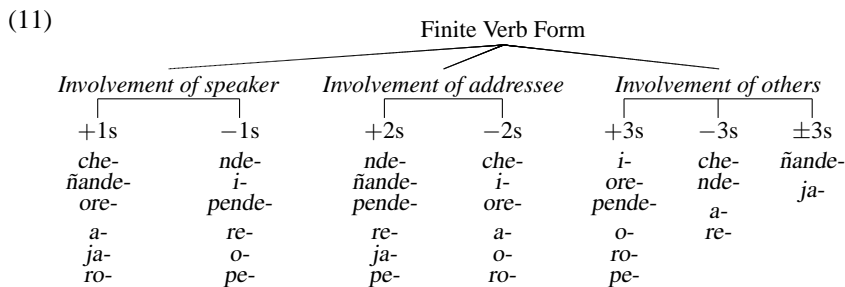
An important advance was made by E. Gregors and J. Soares when they referred to speaker, addressee, and others (1967:131ff). (The same is done in Rodrigues' 1990 treatment of the personal system of Tupinamb, and the present work is inspired by that article.) But their further terminology falls back to the traditional system, and no attempt for any internal grouping was tried. This is not convincing either:



Now it is true that many concepts and classifications that were based on traditional (Latin) grammar and that worked reasonably well for European languages had to be changed when languages from overseas were considered (think of the traditional categories of case or of the temporal / aspectual / modal systems). In view of the facts in Guaraní I think it is time to reconsider the analysis of the Guaranian person system as well in order to avoid inadequate analyses as in (9) and (10) or similar ones that are at least suggested by the terminology.

### 2.2 Beyond Latin grammar

My own proposal assumes *three* cross-classifications on the set of finite verb forms, one for ‘participation’ of speaker, of addressee, and of others, respectively, allowing for neutral categories<sup>1</sup>. This proposal which arises from considering forms of intransitive verbs is shown in the following diagram:



Below the names of the classes I list the prefixes for which each intransitive verb form with this prefix belongs to the class. (For reasons of readability, the superscripts “L” are omitted in the names of the prefixes in (11). However, from here on the ontological status of all linguistic entities will be indicated in accordance with the IL conventions. In particular, “L” indicates lexemes, “P” paradigms, “W” lexical words, and “1” unit sequences, for details see A6 of this book.) This means that the classifications include intransitive verbs of both conjugational classes: normal intransitive verbs (with prefixes of Set Two (5)), and so-called quality verbs

<sup>1</sup> Neutral categories have shown to be useful by research within the IL framework; consider English forms like *good*<sup>1</sup> that are neutral – or unmarked – with respect to number, rather than members of both singular and plural. Hence, such forms belong to the category *Unmarked<sub>N-Num</sub> (-,S)* — the set of noun forms of S which are unmarked for noun form number, where S is an English idiolect system.

(with prefixes of Set One (4)). As we will see in the next subsection, the same classifications (with a minor extension) even apply to the subject-part of transitive verbs.

This subsystem of functional classifications in the syntactic unit ordering of Guaraní idiolect systems corresponds to the following incompletely characterized part of the paradigm of the intransitive verb *mba'apo*<sup>W</sup> '(to) work':

- (12) *a-mba'apo*<sup>1</sup> 'I work' : +1s (Speaker involved as subject),  
 -2s (addressee *not* involved as subject),  
 -3s (others not involved as subject<sup>2</sup>)
- re-mba'apo*<sup>1</sup> 'you(sg) work' : -1s, +2s, -3s  
*o-mba'apo*<sup>1</sup> 'he (she, it, they) work(s)' : -1s, -2s, +3s  
*ña-mba'apo*<sup>1</sup> 'we (including you) work' : +1s, +2s, ±3s<sup>3</sup>  
*ro-mba'apo*<sup>1</sup> 'we (excluding you) work' : +1s, -2s, +3s  
*pe-mba'apo*<sup>1</sup> 'you(pl) work' : -1s, +2s, +3s

### 2.3 Transitive verbs

The three classifications for subject work fine with forms of intransitive verbs. In the case of transitive verbs there are twelve possible forms (not taking into account reflexive, reciprocal, desiderative, or command forms), six with the prefixes of Set Two (5) and six with the prefixes of Set Three ((6), *ro*<sup>-L</sup> appears twice).

Following the description of Gregores & Soares (1967), we need three analogous classifications for involvement of speaker / addressee / others *as object*. Only some forms belong to subject and object classes simultaneously. Can this view be maintained?

Taking into account the reference hierarchy as it was presented above (subsection 1.2), in my opinion it is inappropriate to state that all other but four prefixes (*nde*<sup>-L</sup>, *ro*<sup>-L</sup>, *pende*<sup>-L</sup>, and *po*<sup>-L</sup>) express only one category, subject *or* object. On the contrary, a form as *a-juhu*<sup>1</sup> (traditionally glossed as 'I meet') indicates explicitly that the addressee is *not* involved as object: if he was, *ro*<sup>-L</sup> or *po*<sup>-L</sup> would have been chosen as prefixes. So, *a-juhu*<sup>1</sup> should better be glossed as 'I meet him (her, it, them)'. This can be said analogously for almost all forms.

There seem to be some cases in which the addressee (in the case of our example *a-juhu*<sup>1</sup>) is introduced as semantic patient outside the predicate constituent by a postpositional group such as *nde - ve* or *pende - ve*, often called "free object". This might be seen as an objection to my proposal. However, for semantic reasons not to be explained here, I came to the conclusion that these constituents should be analyzed as modifiers, not complements, of the predicate constituent.

<sup>2</sup> The following abbreviated categorial names are formed analogously

<sup>3</sup> ±3s: 'neutral with respect to others being involved as subject'

Thus, every form of a transitive verb may be analyzed in terms of subject *and* object involvement, and I arrive at the following partial and incompletely characterized paradigm of the sample verb *juhu*<sup>L</sup> ‘(to) meet’ (*partly* because other forms would have to be included in the paradigm, *incomplete* with respect to the categorizations given for each form). (In the glosses, “he” and “him” stand generally for third person and could be replaced by “he, she, it, they” / “him, her, it, them”, respectively. Abbreviations: sg: singular; pl: plural, i: inclusive, e: exclusive. — For two forms, *che-juhu*<sup>L</sup> and *ore-juhu*<sup>L</sup>, we had to assume a new neutral category:  $\pm 2s$ , neutral with respect of involvement of addressee as subject.)

- (13) a. *a-juhu*<sup>L</sup> ‘I meet him’ +1s, -2s, -3s, -1o, -2o, +3o  
 b. *re-juhu*<sup>L</sup> ‘you(sg) meet him’ -1s, +2s, -3s, -1o, -2o, +3o  
 c. *o-juhu*<sup>L</sup> ‘he meets him’ -1s, -2s, +3s, -1o, -2o, +3o  
 d. *ja-juhu*<sup>L</sup> ‘we(i) meet him’ +1s, +2s,  $\pm 3s$ , -1o, -2o, +3o  
 e. *ro-juhu*<sup>L</sup> ‘we(e) meet him’ +1s, -2s, +3s, -1o, -2o, +3o  
 f. *pe-juhu*<sup>L</sup> ‘you(pl) meet him’ -1s, +2s, +3s, -1o, -2o, +3o  
 g. *che-juhu*<sup>L</sup> ‘you/he meet(s) me’ -1s,  $\pm 2s$ ,  $\pm 3s$ , +1o, -2o, -3o  
 h. *nde-juhu*<sup>L</sup> ‘he meets you(sg)’ -1s, -2s, +3s, -1o, +2o, -3o  
 i. *ro-juhu*<sup>L</sup> ‘I/we(e) meet you(sg)’ +1s, -2s,  $\pm 3s$ , -1o, +2o, -3o  
 j. *ñande-juhu*<sup>L</sup> ‘he meets us(i)’ -1s, -2s, +3s, +1o, +2o,  $\pm 3o$   
 k. *ore-juhu*<sup>L</sup> ‘you/he meet(s) us(e)’ -1s,  $\pm 2s$ ,  $\pm 3s$ , +1o, -2o, +3o  
 l. *pende-juhu*<sup>L</sup> ‘he meets you(pl)’ -1s, -2s, +3s, -1o, +2o, +3o  
 m. *po-juhu*<sup>L</sup> ‘I/we(e) meet you(pl)’ +1s, -2s,  $\pm 3s$ , -1o, +2o, +3o

*Note:* Forms that start with *ro*-<sup>L</sup> appear twice in this paradigm. This is indeed seen as a case of syncretism. If we had the ambition to obtain a single unique description for these forms, we would have to assume an additional neutral category for addressee being involved as object, too. We would achieve the following paradigm:

- (14) As (13), but substituting (13 e) and (13 i) by:  
*ro-juhu*<sup>L</sup> ‘I/we(e) meet you(sg)/him’ +1s, -2s,  $\pm 3s$ , -1o,  $\pm 2o$ ,  $\pm 3o$

This possibility will not be pursued any further here. Instead, the unique description for the subject part of intransitive and transitive verbs will be maintained.

I consider this proposal to be more adequate than traditional accounts for two main reasons: (i) with intransitive and transitive verbs, it accounts for the absence of the distinction of number within the third person category (in fact, no number classification is assumed at all) as well as presence of an inclusive / exclusive distinction with the ‘first person plural’, and (ii) it avoids double descriptions of the prefixes of Set One and Set Three (they are, in fact, the same prefixes with

an expanded functionality in the case of transitive verb forms), especially with respect to ‘mixed forms’. In this way we take reference hierarchy into account and thus the fact that every finite form of a transitive verb includes information about subject *and* object. Admittedly, the information about subject is, as an empirical fact, not necessarily very specific — at least in the case of forms with *che*-<sup>L</sup> and *ore*-<sup>L</sup>; it is not.

I will show the corresponding functional classifications in (16) in the next section after having treated the remaining prefixes.

#### 2.4 *Reflexive, reciprocal, desiderative, and command*

A form of one of the thirteen prefixes treated in the last subsections has to be present with the verb stem in any predicate constituent in Guaraní clauses (there is, to my knowledge, no infinitive form at all in Guaraní).

As said above in subsection 1.3 there is one more prefix that has this status and could therefore be added to the other ones: the prefix for command (or imperative mode), *e*-<sup>L</sup>. Only second person singular can be subject in case of a command form.

The other mode-indicating prefix (*ta*-<sup>L</sup>, desiderative) has a totally different status: one of its numerous allomorphs stands additionally in front of one of the thirteen person prefixes (*ta*-<sup>L</sup> is not possible with *e*-<sup>L</sup>). It “indicates wish, permission, exhortation, command” (Gregores & Soares, p. 132).

It should be clear that we have to assume a functional classification of mode with at least three categories: command, desiderative, and neutral mode (indicative). More categories might have to be added due to suffixes.

The two remaining prefixes only occur in forms of transitive verbs and with prefixes of Set Two (5). *je*-<sup>L</sup> (with the phonetical alternative *ñe*-) “indicates that the subject is undergoer, or actor and undergoer simultaneously” (Gregores & Soares 1967:132). The resulting forms are sometimes called reflexive and sometimes passive, depending on context and translation, I suppose. It can be concluded that the grammatical subject (cross-referenced by the prefixes of Set Two) of a form with *je*-<sup>L</sup> semantically is patient — the agent being either the same person(s) (reflexive) or anybody/anything else (passive).

The other prefix, *jo*-<sup>L</sup> (*ño*-), has an even more restricted distribution: it occurs only after the ‘pluralic’ prefixes — in our analysis, only in forms that are marked for involvement of others (+3S: *o*-<sup>L</sup>, *ro*-<sup>L</sup>, *pe*-<sup>L</sup>) or forms that are neutral with respect to such involvement (±3S: *ja*-<sup>L</sup>). Forms with *jo*-<sup>L</sup> express that subject and object are reciprocally the same.

Forms with these two prefixes specify the object by themselves and are therefore incompatible with the object categories as presented above. This has to be taken into account when conceiving the structural and functional systems and their connection, a step that will be tried in the next section.

### 3 Structural and functional systems, and the system connection

#### 3.1 The structural system

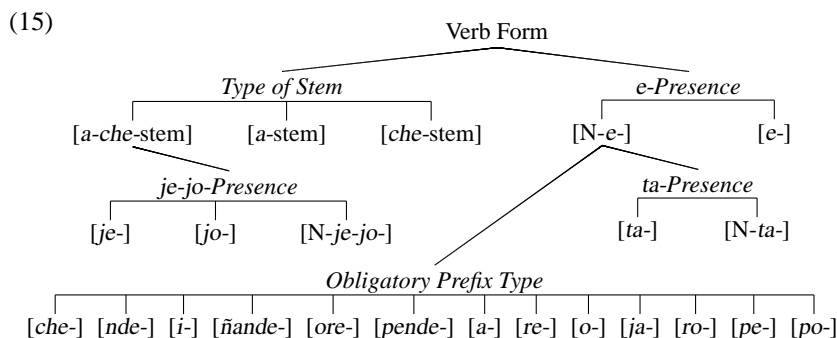
Observing the structural properties considered so far (i.e., the properties referring to their inner composition) of the verb forms in Guaraní that might serve as criteria for structural classifications (elements of the structural part of the classification system on Guaraní verb forms), we see that they fall into three types.

The first concerns the stem: as shown in subsection 1.3 the forms of some stems (stems of so-called quality verbs) occur *only* with forms of prefixes of the first set (4). I will call these stems CHE-STEMS because labels for these categories are traditionally derived from the first person singular verb forms. (Names of categories of lexemes are given in capitals – as are categories of words – in order to mark their difference to categories of forms; cf. p.7.) Other stems (of what traditionally are called regular intransitive verbs) only occur with prefixes of the second set (5) and with  $e^{-L}$ . In the following I name them A-STEMS. Finally, a third class of stems (of transitive verbs) may occur with any of these prefixes (except  $i^{-L}$  that is restricted to intransitive verbs) and also with  $po^{-L}$ . I will call this last class of stems A-CHE-STEMS. This classification of stems can be used in an obvious way to define classes of verb forms that ‘contain’ a form of a stem of each respective class: [che-stem], [a-stem], [a-che-stem].

The second type of criteria is given by the question which of the ‘obligatory’ fourteen prefixes is present in a given form. This yields fourteen corresponding classes. The  $e^{-L}$  containing class has a somewhat special status due to non-co-occurrence with  $ta^{-L}$  (see subsection 1.3).

Finally, the third type accounts for the presence or absence of the optional prefixes.  $ta^{-L}$  may be present with any stem or other prefix except  $e^{-L}$ , thus it seems useful to single out all forms not containing  $e^{-L}$  first.  $je^{-L}$  and  $jo^{-L}$  are mutually exclusive and can both occur only with A-CHE-STEMS.

I come to the (incomplete) structural system of verb forms of Guaraní idiolect systems given in the following diagram: As before, square brackets indicate struc-





tural categories and reference to idiolect systems is omitted. Any category [x] (or [N-x]) may informally be defined as: “the set of all *verb* forms that (don’t) ‘contain’ a form of a stem of class x or a form of the affix  $x^L$ ”. Reference to verb forms is omitted in the abbreviated names. The notion of containing in the above scheme has to be made more precise. This will eventually be done in the next section.

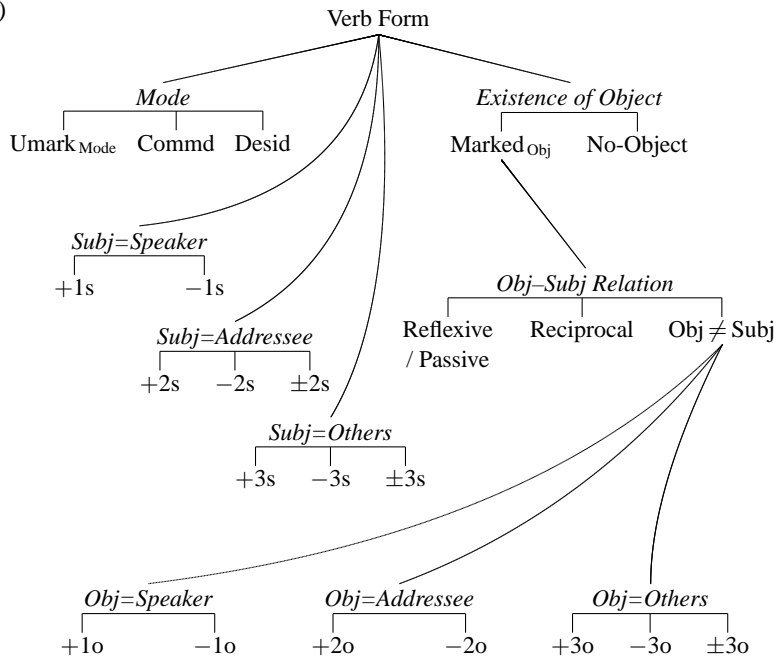
### 3.2 *The functional system*

The functional system for verb forms is more complex. The three classifications for involvement of speaker / addressee / others that were shown above in (11) should be renamed to involvement of speaker / addressee / others *as subject*, abbreviated: “Subj=Speaker” etc. They are cross-classifications on the set of all verb forms — and so is the classification of mode.

But not all verb forms are marked for the existence of an object, and thus the three analogous cross-classifications for involvement of speaker / addressee / others as object (“Speaker=Obj” etc.) should be based on the subset of verb forms that *are* marked for existence of an object (as opposed to the forms that explicitly can’t have an object [No-Object]) and are neither reflexive / passive nor reciprocal, that is, the forms for that an object exists and is not related to the subject [Obj ≠ Subj].

This yields the (incomplete, since no suffixes are considered) functional system for verb forms. Like the structural system it is a subset of the SUO in Guaraní idiolect systems:

(16)



### 3.3 The system connection

The combination between the two systems (15) and (16) is provided, as explicated above (subsection 0.3), by the system connection. It is the third and last complex formal entity of Guaraní idiolect systems that is to be presented in this article. It ensures that a given form is assigned to the right functional categories (to which it is related by the paradigm), taking into account the inner composition of forms via their membership to structural categories only.

Below I give the relevant elements of the system connection for verb forms of Guaraní idiolect systems. Each element is, as explicated above, a pair such that the first component is a set of structural categories (endpoints or other) and the second component a set of functional categories (endpoints of the SUO). Any form that belongs to the intersection of the first component (that is: any form that is an element of all elements) also belongs to the intersection of the second. In the names of the functional sets space has been arranged for easier recognition of relevant elements.

Sometimes, one formal property is sufficient to determine the belonging of a given form to one or more functional categories. For instance, occurrence of the *ta*<sup>L</sup>-prefix indicates desiderative mode without any further conditions. This is the case of all prefixes relevant in the following elements of the system connec-

tion (this is characteristic for ‘agglutinating languages’ as Guaraní, compare the remarks in the last section):

- (17)  $\langle \{ [a\text{-stem}] \}, \{ \text{No-Object} \} \rangle$   
 $\langle \{ [che\text{-stem}] \}, \{ \text{No-Object} \} \rangle$   
 $\langle \{ [je\text{-}] \}, \{ \text{Reflexive / Passive} \} \rangle$   
 $\langle \{ [jo\text{-}] \}, \{ \text{Reciprocal} \} \rangle$   
 $\langle \{ [ta\text{-}] \}, \{ \text{Desiderative} \} \rangle$   
 $\langle \{ [N\text{-}ta\text{-}] \}, \{ \text{Unmarked}_{\text{Mode}} \} \rangle$

Some examples may help to understand the functioning of the system connection. Consider the form *re-guata*<sup>L</sup>. As its stem never co-occurs with the prefixes *che*<sup>L</sup>, *nde*<sup>L</sup> etc., the form belongs to the structural category [a-stem]. Therefore, the first pair in (17) can be applied. Any form that belongs to the intersection of the first component also belongs to the intersection of the second. Thus, due to the analysis given by the system connection, the form *re-guata*<sup>L</sup> is an element of **No-Object**, since it belongs to the intersection of the first component (trivially so, as the first component only has one element, the set [a-stem]). Of course, this statement seems to be true but redundant since it corresponds to the observation that all verbs with A-STEMS are intransitive. However, it is pure coincidence that in Guaraní, a property of words (intransitive verbs) is reflected by a property of all their forms (belonging to **No-Object**). If a better analysis that avoids this doubling could be shown, I would be happy to replace this point.

As another example consider the form *ta-ja-jo-juhu*<sup>L</sup> (‘may we meet one another’). As it belongs to [jo-] and [ta-], we can conclude that the form belongs to the functional categories **Reciprocal** and **Desiderative**, referring to the fourth and fifth elements in (17).

Note that the structural category [N-ta-] is a subset of [N-e-], so the last pair states efficiently that a form in that neither *ta*<sup>L</sup> (nor *e*<sup>L</sup>) is present is in neutral mode (or unmarked for mode, i.e. ‘indicative’).

We continue with those forms that contain a prefix of Set Two (except forms that contain *ro*<sup>L</sup>, which will be treated separately below). Their subject is well defined no matter whether they are transitive or intransitive. (Note that, according to this analysis, forms with *e*<sup>L</sup> behave as forms with *re*<sup>L</sup>, but are additionally marked for Command mode.)

- (18)  $\langle \{ [a\text{-}] \}, \{ +1s, -2s, -3s \} \rangle$   
 $\langle \{ [re\text{-}] \}, \{ -1s, +2s, -3s \} \rangle$   
 $\langle \{ [e\text{-}] \}, \{ -1s, +2s, -3s, \text{Command} \} \rangle$   
 $\langle \{ [o\text{-}] \}, \{ -1s, -2s, +3s \} \rangle$   
 $\langle \{ [ja\text{-}] \}, \{ +1s, +2s, \pm 3s \} \rangle$   
 $\langle \{ [pe\text{-}] \}, \{ -1s, +2s, +3s \} \rangle$

The functioning of the system connection becomes clearer when we observe these

somewhat more complex cases. Consider again the form *re-guata*<sup>1</sup> that belongs to [re-]. The second element in (18) says that *re-guata*<sup>1</sup> belongs to all three functional categories in the set (that is its second component)

$$(19) \{-1s, +2s, -3s\}$$

This is to say (in traditional terms) that *re-guata*<sup>1</sup> is a Second Person Singular form. The traditional categories can be defined as *intersections* of sets like (19) (that is, in this example, the set of all forms that are elements in every element of set (19)).

It should be seen as an important result of the account given here that with the suggested interpretation of the traditional categories, it is possible to integrate the results of other, non-IL work into this conception. I will from here on make use of the traditional terms, conceiving them as names of intersection sets in the way described above.

The fifth pair in (18) is relevant for our other sample form *ta-ja-jo-juhu*<sup>1</sup>. We saw that it belongs to **Desiderative** (due to the occurrence of *ta*<sup>-L</sup>) and to **Reciprocal** (due to the occurrence of *jo*<sup>-L</sup>). Now we see that the form also belongs to the three categories +1s, +2s and ±3s (traditionally, the form is ‘first person plural inclusive’) since it belongs to the category [ja-]. No more elements of the system connection (in the part presented here and possibly generally) concern this form, thus we can conclude that

$$(20) \langle ta-ja-jo-juhu^1, \{+1s, +2s, \pm 3s, \text{Desiderative, Reciprocal}\} \rangle$$

is an element of the paradigm of the verb *juhu*<sup>W</sup> ‘(to) meet’. This exemplifies how the different elements of the system connection work together in the determination of paradigms, especially, in the determination of their elements’ second components, i.e. the categorizations of the forms that are the first components of the elements of the paradigms.

Up to now sets of functional categories have been related to sets with only one structural category by the system connection. However, in the majority of cases only a combination of formal categories suffices to allocate forms to certain formal categories. For instance, if a form belongs to [che-] and to [che-stem] it can be assigned to the formal categories +1s, -2s and -3s, but [che-] alone isn’t sufficient to specify anything. (This may come as a surprise since *che*<sup>-L</sup> ‘indicates first person singular’, but as shown above, we have to distinguish between ‘first person singular *subject*’ and the corresponding ‘first person singular *object*’.)

Consequently, the forms of quality verbs are affected by the following pairs:

$$(21) \begin{array}{l} \langle \{ [che\text{-stem}], [che\text{-}] \}, \{ +1s, -2s, -3s \} \rangle \\ \langle \{ [che\text{-stem}], [nde\text{-}] \}, \{ -1s, +2s, -3s \} \rangle \\ \langle \{ [che\text{-stem}], [\tilde{n}ande\text{-}] \}, \{ +1s, +2s, \pm 3s \} \rangle \\ \langle \{ [che\text{-stem}], [ore\text{-}] \}, \{ +1s, -2s, +3s \} \rangle \\ \langle \{ [che\text{-stem}], [pende\text{-}] \}, \{ -1s, +2s, +3s \} \rangle \\ \langle \{ [i\text{-}] \}, \{ -1s, -2s, +3s \} \rangle \end{array}$$

For instance, *che-mandu*'a<sup>1</sup> belongs to the categories +1s, -2s and -3s (traditionally: it is a 'first person singular [subject]' form) because the stem does never co-occur with *a*<sup>-L</sup>, *re*<sup>-L</sup> etc., and because it contains the prefix *che*<sup>-L</sup>. Since *i*<sup>-L</sup> can appear only with forms of CHE-STEMS, it is alone by itself sufficient for assigning a form such as *i-mandu*'a<sup>1</sup> to the related functional categories of traditional 'third person'.

Now elements of the system connection of Guaraní idiolect systems that take account of forms of transitive verbs will be observed. In the case of forms with prefixes of Set Two that don't contain neither *je*<sup>-L</sup> nor *jo*<sup>-L</sup>, according to our analysis (compare subsections 2.3 and 2.4) *all* forms contain information about the object as well. The same holds true for prefixes of Set Three. As the subject-classes of the forms with prefixes of Set Two have been treated already (see (18)), we need not repeat this here. Consider

$$(22) \begin{array}{l} \langle \{ [a\text{-che-stem}], [che\text{-}] \}, \{ -1s, \pm 2s, \pm 3s, +1o, -2o, -3o \} \rangle \\ \langle \{ [a\text{-che-stem}], [nde\text{-}] \}, \{ -1s, -2s, +3s, -1o, +2o, -3o \} \rangle \\ \langle \{ [a\text{-che-stem}], [\tilde{n}ande\text{-}] \}, \{ -1s, -2s, +3s, +1o, +2o, \pm 3o \} \rangle \\ \langle \{ [a\text{-che-stem}], [ore\text{-}] \}, \{ -1s, \pm 2s, \pm 3s, +1o, -2o, +3o \} \rangle \\ \langle \{ [a\text{-che-stem}], [pende\text{-}] \}, \{ -1s, -2s, +3s, -1o, +2o, +3o \} \rangle \\ \langle \{ [N\text{-je-jo-}], [a\text{-}] \}, \{ -1o, -2o, +3o \} \rangle \\ \langle \{ [N\text{-je-jo-}], [re\text{-}] \}, \{ -1o, -2o, +3o \} \rangle \\ \langle \{ [N\text{-je-jo-}], [e\text{-}] \}, \{ -1o, -2o, +3o \} \rangle \\ \langle \{ [N\text{-je-jo-}], [o\text{-}] \}, \{ -1o, -2o, +3o \} \rangle \\ \langle \{ [N\text{-je-jo-}], [ja\text{-}] \}, \{ -1o, -2o, +3o \} \rangle \\ \langle \{ [N\text{-je-jo-}], [pe\text{-}] \}, \{ -1o, -2o, +3o \} \rangle \end{array}$$

Please note that all the classes [*je*-], [*jo*-] and [N-*je-jo*-] are proper subsets of [*a-che-stem*]. Since prefixes of Set Three (not considering *ro*<sup>-L</sup>) don't co-occur with *ja*<sup>-L</sup> or *jo*<sup>-L</sup>, [N-*je*<sup>-L</sup>*jo*-] could have been chosen instead of [*a-che-stem*] as well. I'm not shure about what type of criteria to use in these cases to make a decision.

Take again as an example two forms of the verb *juhu*<sup>W</sup> '(to) meet': *nde-juhu*<sup>1</sup> 'he meets you(sg)' and *re-juhu*<sup>1</sup> 'you(sg) meet him'. All forms of this verb belong to [*a-che-stem*]. In addition to this, the first form, *nde-juhu*<sup>1</sup>, belongs to [*nde*-]. Thus the form belongs to all structural categories that are elements of the first component of the second pair given in (22) — and consequently we may conclude

that the form belongs to all six categories in the second component. Traditionally speaking, *nde-juhu*<sup>1</sup> is a ‘third person subject’ form (belongs to  $-1s$ ,  $-2s$ , and  $+3s$ ) and ‘second person singular object’ form (is an element of  $-1o$ , of  $+2o$ , and of  $-3o$ ).

For a complete characterization of the second sample form, *re-juhu*<sup>1</sup>, we have to take in account more than one element of the system connection. The form belongs to  $[N-je-jo-]$  and to  $[re-]$ , consequently it belongs to any category in the second component of the seventh pair in (22), that is to say, to  $-1o$ , to  $-2o$ , and to  $+3o$  (traditionally, to ‘third person object’). But its belonging to  $[re-]$  suffices to assign the form to the intersection of (19) — just as the form *re-guata*<sup>1</sup> discussed above. This is due to the second pair in (18) that treats forms of transitive verbs as well as forms of intransitive verbs. Therefore, we get as an element of the paradigm *juhu*<sup>P</sup>:

$$(23) \langle re-juhu^1, \{ -1s, +2s, -3s, -1o, -2o, +3o \} \rangle.$$

Finally, we have to account for forms with the ‘mixed prefixes’ *ro*<sup>-L</sup> and *po*<sup>-L</sup>, cf. (13 e, i, m). It turns out that their behaviour is quite different since *ro*<sup>-L</sup> appears as a normal prefix with intransitive verbs, cross-referencing ‘first person plural exclusive’, and as a ‘mixed prefix’ proper, specifying that ‘second person singular’ is object and that speaker participates as subject. This is a case of syncretism. (By the way, syncretism in paradigms can be predicted if in the system connection a first component of elements appears more than once.)

The behaviour of *po*<sup>-L</sup> is in a way contrary to that of *ro*<sup>-L</sup>: it appears only in forms of transitive verbs and is *specific* (in a sense that will be clarified in the next section).

$$(24) \langle \{ [po-] \}, \{ +1s, -2s, \pm 3s, -1o, +2o, +3o \} \rangle \\ \langle \{ [ro-] \}, \{ +1s, -2s \} \rangle \\ \langle \{ [a-stem], [ro-] \}, \{ +3s \} \rangle \\ \langle \{ [je-], [ro-] \}, \{ +3s \} \rangle \\ \langle \{ [jo-], [ro-] \}, \{ +3s \} \rangle \\ \langle \{ [N-je-jo-], [ro-] \}, \{ +3s, -1o, -2o, +3o \} \rangle \\ \langle \{ [N-je-jo-], [ro-] \}, \{ \pm 3s, -1o, +2o, -3o \} \rangle$$

Thus, the first pair in (24) describes for instance that *po-juhu*<sup>1</sup> ‘I/we(excl) meet you(pl)’ belongs to all six categories in the second component since *po-juhu*<sup>1</sup> belongs to  $[po-]$ .

Considering the second and fourth pair in (24) and the fourth pair in (17), we see that the form *ro-jo-juhu*<sup>1</sup> ‘we(excl) meet one another’ appears in the paradigm of the verb *juhu*<sup>W</sup> as follows:

$$(25) \langle ro-jo-juhu^1, \{ +1s, -2s, +3s, \text{Reciprocal} \} \rangle,$$

whereas the relevant element in the paradigm *guata*<sup>P</sup> for *ro-guata*<sup>1</sup> is:

(26)  $\langle ro-guata^1, \{ +1s, -2s, +3s, No-Obj \} \rangle$ .

(To arrive at (26), compare the first pair in (17) and the second and fifth pairs in (24).)

Finally, the syncretistic form  $ro-juhu^1$  appears twice in the paradigm  $juhu^P$ . In both cases the second pair of (24) is applied, the first time together with the sixth, the second time with the seventh:

(27)  $\langle ro-juhu^1, \{ +1s, -2s, +3s, -1o, -2o, +3o \} \rangle$ .  
 $\langle ro-juhu^1, \{ +1s, -2s, \pm 3s, -1o, +2o, -3o \} \rangle$

### 3.4 From the description of structures to the description of units

The first sections of this article presented a description of a small but essential part of the syntax of Guaraní, using the framework of Integrational Linguistics. The central question was: how is it that verbal paradigms are construed in this language, as far as the person system is concerned? We saw that it is a small set of prefixes that play a key role in the construction of paradigms (in other languages, there may be suffixes or auxiliary words that serve similar purposes), so we conclude that in this area syntax and morphology are closely interwoven.

In descriptions of individual languages, the construction of paradigms is traditionally treated in connection with the language's morphological units (or auxiliaries), more specifically, with its inflectional affixes. (This is the proceeding chosen by Gregores and Soares in their (1967) grammar.) (this is the case of the grammar of Gregores & Soares (1967), too). In the integrational framework, a method for describing such affixes (beyond stating their forms and morphological combinatory patterns) has apparently not been considered yet. It would be very useful if we had instruments in our hands that would allow for stating the effects of affixes.

When it comes to the morphological, in particular, to the lexicographical description of inflectional affixes, there has traditionally been no doubt that it is exactly their role in syntax (viz., in the construction of paradigms) that has to be accounted for. This idea should be reconstructed in IL. In comparative studies, especially when comparing the morphology of genetically related languages, such instruments would allow for well-defined statements like "in language *A*, affix  $m^L$  serves to mark categories  $x_1$  and  $x_2$ , whereas in language *B*, the (cognate) affix  $n^L$  marks the category *y*"; at the same time, a syntactic comparison of the Syntactic Unit Orderings of both languages and their relevant system connections would become possible. The same holds of auxiliary words (diachronically, often the source of later affixes).

The complex interplay of syntax and morphology seems to be a first major obstacle for obtaining such instruments. Since my original interest was exactly a lexicographical description of Guaraní, based on the IL framework, as well as a

possible comparison with other Tupí languages, I took up the work to fill in this apparent gap in the integrational theory. It turned out that the system connection can serve as a basis for describing the corresponding prefixes (lexemes), once an adequate concept of ‘containing’ (to be developed in the next section) is on hand, as will be shown in the last section of this paper. Some consequences for the resulting descriptions in Guaraní and other languages are indicated in subsection 5.2.

It is to be expected that a resulting compact and unified description of inflectional morphology is useful also for typological statements. Some first steps in this direction will be made in the last subsection 5.3, illustrating the use of some additional terms developed in this paper by applying them to Guaraní.

## 4 Inflectional units contained in forms

### 4.1 *Inflectional units*

Lieb’s (1992) paper introduces a conceptual framework that can easily be extended in order to describe the marking effects of inflectional units. By *inflectional units* we understand here all inflectional affixes as well as all auxiliary words of an idiolect system. Therefore, the concept is mixed with respect to the domains of morphology and syntax, but homogeneous with respect to ontology: all inflectional units are conceived as pairs of a paradigm P and a concept b that is its meaning, just as stems and other lexical words are. (For the IL conception of lexical meaning see p. 7 of this book; also Lieb 1983: Sec. 13.)

It is a general property of inflectional units that they don’t have a ‘proper’ lexical meaning in a traditional sense — instead of this, their second component is *the empty concept*  $b^0$  (cf. p. 7). For this very reason, a grammatical or lexicographical description of inflectional units has to accomplish more than just to mention their forms: It is precisely their role in the construction of paradigms that has to be accounted for.

As we have seen in the last sections, an occurrence of a form of an inflectional unit ‘marks’ syntactic forms and thus provides a criterion for structural categories that, in turn, determine the functional categories the ‘marked’ form belongs to. Given the concept of the system connection of an idiolect system, we have the instruments at our hands to make these notions more precise.

### 4.2 *An example*

Consider the prefix  $re^{-L}$  of a given idiolect system S of Guaraní.  $re^{-L}$  is an inflectional unit and could be defined as follows (in this part, reference to idiolect systems S is reinstated in the definitions and assumptions):

(28) *Definitions.*



- a.  $re^{-L}(S) =_{df} \langle re^P(S), b^0 \rangle$   
 b.  $re^P(S) =_{df} \{ \langle re^1, \{ \text{Affix Form}(-, S) \} \} \}$

*Notes:* (i) “Affix Form” is the name of a general relation that holds between forms of affixes and idiolect systems (“*f is an affix form of S*”). This relation may be defined in a general theory of language, but the *category*  $\text{Affix Form}(-, S)$  (the set of all affix forms of  $S$ ) is to be *identified* in a theory of Guaraní idiolect systems. (ii) As can be expected (cf. Lieb 1992, p. 171),  $re^{-L}(S)$  is a pair of an improper paradigm and the empty concept  $b^0$ . (iii)  $re^1$  is a morph (a morphological form)  $f$ , that is, a sequence of morphological base forms  $w$ . In the present case, it is the unit sequence  $re^1$  of a base form that consists of the sequence of the phonemes /*re*/ and a phonological structure of this sequence. (iv) A different theory of Guaraní could evaluate the nasal vowels as phonemes, and thus,  $r\bar{e}^1$  as another allomorph of  $re^{-L}$ . In this case, the paradigm  $re^{-P}$  would contain two equivalent forms but would still be improper.

Usually,  $re^{-L}$  is called a ‘second person singular prefix’. — How is this to be understood? I think an adequate explication can be made informally as follows: Forms that ‘contain’  $re^{-L}$  ‘are marked’ by this affix as belonging to the second person singular. Using the concepts presented in the last sections, we can explicate ‘are marked’ as follows: Forms that ‘contain’  $re^{-L}$  belong to the structural category [*re-*] (cf. subsection 3.1). As [*re-*] figures as a (in fact, the only) element of a first component of an element of the system connection that has the set  $\{-1s, +2s, -3s\}$  as its second component, any form in [*re-*] belongs to each functional category in that set. Traditionally speaking: it is a ‘second person singular’ form (cf. discussion below (19)). Thus, assigning forms to these categories is an ‘effect’ of  $re^{-L}$ .

In developing a precise concept of such effects (that will eventually be called “marking effects”) I proceed as follows: First, a concept of ‘being contained’ has to be provided for both types of inflectional units (rest of this section). In the next section, the relevant structural categories will be singled out and characterized. Finally, the system connection will be used to come to a notion of marking effect and some auxiliary concepts.

### 4.3 *Being contained*

#### *Being morphologically contained*

First, consider the verb form *re guata*<sup>1</sup> ‘you(sg) go’ of Guaraní idiolect systems that ‘contains’ the Guaraní affix  $re^{-L}$ . The precise formulation of relation  $R$  between  $re^{-L}$  and *re-guata*<sup>1</sup> could be stated as follows:

- (29) There is a form  $f_1$  (*re*) and a syntactic base form  $w$  (*re guata*) in the idiolect system such that:
- a.  $f_1$  is a form of (the paradigm) of  $re^{-L}$  (in fact, the only form)

- b.  $w$  is a member (a second component of an element, see above section ) of *re-guata*<sup>1</sup> (in fact, the only member)
- c.  $f_1$  occurs before a form of the stem of *re-guata*<sup>1</sup> in a first component of a morphological version of  $w$ . (A morphological version of a syntactic base form consists of (i) a sequence of morphological base forms and (ii) a morphological structure.)

$R$  is a relation of ‘being morphologically contained’, to be more precise, the relation of ‘being morphologically contained as a prefix in the main part’ that could be defined using (29). There are many different relations between lexemes (either stems or affixes) of an idiolect system and syntactic word forms of the system that are relations of being morphologically contained. These relations can be defined as follows:

(30) *Definition.* Any relation  $R$  is a *relation of being morphologically contained* (is an *m-containing relation*, for short) with respect to an idiolect system  $S$  iff (if and only if):

- a.  $R$  holds between lexemes  $\langle P, b \rangle$  of  $S$  and syntactic word forms  $f$  of  $S$ ,
- b. for any pair  $\langle \langle P, b \rangle, f \rangle \in R$ , there is a form  $f_1$  and a syntactic base form  $w$  such that
  - (i)  $f_1$  is a form of  $P$ ,
  - (ii)  $w$  is a member of  $f$ ,
  - (iii)  $f_1$  occurs in the first component of a morphological version of  $w$ .

Any *m-containing* relation has to meet the conditions in (30 b) and possibly imposes further ones. For instance, it might be required that  $f_1$  occurs at a certain position or that  $w$  is a specific member (the main, the finite, the second, etc.) of  $f$ .

Consequently we may need relations of morphological containing as a suffix, infix, prefinal suffix and so on. Which relations must be used in a description of a language is an empirical question.

#### *Being syntactically contained*

In the definition of relations of being morphologically contained we had to cross the border between morphology and syntax. The syntactic equivalent, relations that hold between *words* (typically, auxiliary words), and word *forms*, can be defined much easier.

Consider the verb form *is going* that ‘contains’ a form of the auxiliary verb (to) *be*<sup>W</sup>. Spelled out: there is a form of the auxiliary verb (*is*<sup>1</sup>) that occurs in the verb form. Generally we can define the most unspecific relation of being syntactically contained as follows:

- (31) *Definition.* Let  $\langle P, b \rangle$  be a lexical word of an idiolect system  $S$  and  $f$  a form of  $S$ . Then  $\langle P, b \rangle$  is syntactically contained in  $f$  with respect to  $S$  iff there is a form  $f_1$  of  $S$  such that:
- a.  $f_1$  is a form of  $P$ ;
  - b.  $f_1$  occurs in  $f$ .

Trivially, every word of an idiolect system  $S$  is syntactically contained in each of its forms. It is primarily auxiliary words that may also be contained in forms of *other* lexical words.

Generally, the  $s$ -containing relations are defined in strict analogy to morphologic ones (substituting “syntactic” for “morphologic” in (30) and using the definiendum of (31) instead of (30 b)):

- (32) *Definition.* Any relation  $R$  is a *relation of being syntactically contained* (is an *s-containing relation*) with respect to an idiolect system  $S$  iff:
- a.  $R$  holds between lexical words  $\langle P, b \rangle$  of  $S$  and syntactic word forms  $f$  of  $S$ ,
  - b. for any pair  $\langle \langle P, b \rangle, f \rangle \in R$ , there is a form  $f_1$  such that
    - (i)  $f_1$  is a form of  $P$ ;
    - (ii)  $f_1$  occurs in  $f$ .

There may be many different relations of syntactic containing, too, that impose further and different conditions on the form  $f_1$  of the lexical word or its place in  $f$ . Which specific relation underlies a structural category depends on the facts to be accounted for; often the most general one defined in (31) will suffice.

The concept of containing in its morphological variant accounts for the complex connection between morphology and syntax mentioned above (sec. 3.4). This is a first important step for describing the effects of inflectional affixes. With the syntactic containing relation we are able to treat auxiliary words in arbitrary languages analogously to inflexional affixes. Together, these concepts are a starting point for describing the contribution of arbitrary inflectional units to the paradigmatic formation of word forms in arbitrary languages.

## 5 The functional effects of inflectional units

### 5.1 Containing sets and inflectional-unit-categories

#### *Morphological containing sets*

Starting from any relation of being morphologically contained in  $S$ , say, ‘is morphologically contained as a prefix in the main part’ ( $m$ -contained<sub>pre</sub>) that holds in idiolect systems  $S$  of Guaraní between  $re^{-L}$  and  $re-guata^1$  (cf. (29)), we may easily identify the set of syntactic word forms that contain (in the specific manner)  $re^{-L}$ :

$$(33) \text{re-form}_{\text{pre}}(-, S) = \{f \mid \text{re}^{-L} \text{ is m-contained}_{\text{pre}} \text{ in } f \text{ with respect to } S\}$$

For idiolect systems  $S$  of Guaraní, the forms *re-guata*<sup>1</sup>, *te-re-je-juhu*<sup>1</sup> and many others would belong to this class. If *re*<sup>-L</sup> appears as a nominal prefix, nominal forms would belong to this set, too. In languages such as English or German that have *re*<sup>-L</sup> as a derivational affix (given that the paradigms are the same as in Guaraní — which may well be the case) all forms with these prefixes belong to the corresponding classes of English or German idiolect systems (definition (30) is not restricted to inflectional affixes).

A category such as (33) may be called an *m-containing set for re*<sup>-L</sup> in  $S$ , in the following sense:

$$(34) \textit{Definition.} \text{ Let } \langle P, b \rangle \text{ be an lexeme of an idiolect system } S. \text{ A set } K \text{ of word forms of } S \text{ is an } m\text{-containing set for } \langle P, b \rangle \text{ iff there is a relation } R \text{ of being morphologically contained in } S \text{ such that:}$$

$$K = \{f \mid \langle \langle P, b \rangle, f \rangle \in R\}$$

This and the following definitions are not restricted to inflectional units, so we can speak of *m-containing sets* for stems, too. This may be useful when considering all words (for instance, all compound words) that contain a certain stem in a certain manner (for example, in front of an other stem). However, in the following discussion, only the inflexional units will be considered.

#### *Restricted containing sets*

In order to be a possible category in the structural classification system on a given basic constituent category (here: **Verb Form**), any set has to be a subset of that basic constituent category. We can meet this further condition by intersecting an *m-containing set* for an affix with the constituent category. (Such a restriction is indeed necessary since a given affix may play a role in different areas — think of the English suffix *-s*<sup>L</sup> that ‘marks plural’ with nominal stems and ‘marks third person singular present tense’ with verb stems.)

Thus the structural category [*re*-] that appeared in subsection 3.1 is identified as follows:

$$(35) [\text{re-}](-, S) = \{f \mid f \in \text{Verb Form}(-, S) \wedge \text{re}^{-L} \text{ is m-contained}_{\text{pre}} \text{ in } f \text{ with respect to } S\}$$

(This is no definition — what could be defined is a general relation [*re*-] that holds between forms and idiolect systems.) This set, which relies on an *m-containing relation*, should be called an *m-containing set for re*<sup>-L</sup> *restricted to Verb Form in S*. Generally:

$$(36) \textit{Definition.} \text{ For any idiolect system } S, \text{ lexeme } \langle P, b \rangle \text{ of } S \text{ and basic syntactic constituent category } K_0 \text{ of } S, K \text{ is an } m\text{-containing set for } \langle P, b \rangle \text{ restricted to } K_0 \text{ in } S \text{ iff there is a relation } R \text{ such that:}$$

- a. R is an m-containing relation with respect to S;
- b. for every  $f \in K$ ,
  - (i) R holds between  $\langle P, b \rangle$  and  $f$ ;
  - (ii)  $f \in K_0$ ;
- c. there is no  $f \notin K$  that satisfies (bi) and (bii).

(Again, this definition is not restricted to inflectional units. An investigation in, say, derivational morphology would make use of the term for dealing with derivational affixes  $\langle P, b^0 \rangle$ .) Restricted m-containing sets for inflectional affixes are one of three major types of structural categories of idiolect systems. Put differently, using occurrence of forms of inflectional affixes (in the way explained here) is one of three main manners for defining structural categories. The second type relies on occurrence of forms of auxiliary words (see below). The third type refers to lexeme categories of stems. It underlies structural categories such as [*che-stem*] and should be made more precise in an other occasion.

In the syntactic case, a set of syntactic word forms can analogously be called an s-containing set (compare (34)):

- (37) *Definition.* Let  $\langle P, b \rangle$  be a lexical word of an idiolect system S. A set K of word forms of S is an *s-containing set for*  $\langle P, b \rangle$  iff there is a relation R of being syntactically contained in S such that:
- $$K = \{f \mid \langle \langle P, b \rangle, f \rangle \in R\}$$

Analogously to the morphological case we can restrict these containing sets to a given basic constituent category. Provided that the relevant words  $\langle P, b \rangle$  are auxiliary words, we again obtain in this way sets that are potentially structural categories.

#### *Inflectional-unit-categories*

Now consider any inflectional unit that underlies a structural category in an idiolect system. (Note that all terms which are developed in the remaining sections of this paper are only relevant to inflectional units. Therefore the definitions will be restricted in the appropriate way.) We can characterize the structural category in the following way:

- (38) *Definition.* Let S be any idiolect system,  $\langle P, b^0 \rangle$  an inflectional unit of S, K a set of syntactic units of S, and  $K_0$  a basic constituent category of S. K is a  $\langle P, b^0 \rangle$ -category for  $K_0$  in S iff
- a. K is a category in the structural system for  $K_0$  in the SEO of S
  - b. (i) or (ii):
    - (i)  $\langle P, b^0 \rangle$  is an affix of S and K is an m-containing set for  $\langle P, b^0 \rangle$  restricted to  $K_0$  in S

- (ii)  $\langle P, b^0 \rangle$  is an auxiliary word of  $S$  and  $K$  is an  $s$ -containing set for  $\langle P, b^0 \rangle$  restricted to  $K_0$  in  $S$

For example, given an appropriate idiolect system  $S$  of Guaraní, we can say that  $[re-](-,S)$  as given in (35) is a  $re^{-L}$ -category for Verb Form in  $S$ . (This is no tautology despite the contrary impression that might arise from the name chosen for the structural category.) There might be several  $re^{-L}$ -categories in  $S$  (with varying  $K_0$ , the underlying relation  $R$ , or both).

The concept of inflectional-unit-category is an important achievement towards defining the *marking effects* of an inflexional unit.

## 5.2 Marking pairs of inflectional units and their description

Consider an inflexional unit of a given idiolect system, say again,  $re^{-L}$  of an idiolect system of Guaraní, and a corresponding category for a given basic constituent category of this idiolect system, say,  $[re-](-,S)$ .

We can now observe the system connection of the idiolect system and find all elements  $\langle J_1, J_2 \rangle$  whose first component has the category as an element. — In the case of  $[re-]$ , we find the two pairs  $\langle \{[re-]\}, \{-1s, +2s, -3s\} \rangle$ , cf. (18), and  $\langle \{[N-je-jo-], [re-]\}, \{-1o, -2o, +3o\} \rangle$ , cf. (22). We will call these elements *marking pairs* for  $re^{-L}$  relative to Verb Form and  $S$ . Generally:

(39) *Definition.* [ $S, \langle P, b^0 \rangle, K$ , and  $K_0$  are as in (38).] A pair  $\langle J_1, J_2 \rangle$  is a *marking pair* for  $\langle P, b^0 \rangle$  relative to  $K_0$  and  $S$  iff

- a.  $\langle J_1, J_2 \rangle$  is an element of the system connection for  $K_0$  in  $S$
- b. there is a  $K \in J_1$  that is a  $\langle P, b^0 \rangle$ -category for  $K_0$  in  $S$

We can identify the ‘marking content’ of an inflexional unit with the set of all of its marking pairs:

(40) *Definition.* [ $S, \langle P, b^0 \rangle, K$ , and  $K_0$  are as in (38).] The *marking content* of  $\langle P, b^0 \rangle$  relative to  $K_0$  and  $S$  =<sub>df</sub>  $\{ \langle J_1, J_2 \rangle \mid \langle J_1, J_2 \rangle \text{ is a marking pair for } \langle P, b^0 \rangle \text{ relative to } K_0 \text{ and } S \}$

It is the marking content that has to be described when treating an inflexional unit in descriptions of a language. In the case of  $re^{-L}$ , an informal description (using abbreviated traditional terminology for intersections of sets, cf. discussion below (19): ‘second person singular subject’ =  $\bigcap \{-1s, +2s, -3s\}$  etc.) could be made as follows:

- (41) Verb forms with prefix  $re^{-L}$  are always marked for second person singular subject. If  $re^{-L}$  stands in front of a transitive verb stem and neither  $je^{-L}$  nor  $jo^{-L}$  are present, the form belongs to third person object also.

A corresponding entry in a lexicon of Guaraní probably has to rearrange things a bit, but is ultimately based on the marking content as well (in case of any doubts, it is indeed a “useful practice” to treat affixes as entries of their own in dictionaries, cf. Zgusta (1989:304)):

- (42) **re-** ⟨Verb Prefix⟩    **1.** (with intransitive verb stems:) indicates that subject is second person singular    **2.** (with transitive verb stems but without *je-* or *jo-*;) indicates that subject is second person singular and that third person is object.

It can be observed that affixes like English  $-s^L$  or German  $-e^L$ ,  $-en^L$ ,  $-(e)st^L$  etc. have more than one marking content — at least one relative to Verb Form and an other one for Noun Form. This might have to be indicated by a dictionary as follows (this is, of course, not more than a first intuitive approximation):

- (43) **-s**    I ⟨Verb Suffix⟩ indicates third person singular  
              II ⟨Nominal Suffix⟩ indicates plural

Finally, some affixes have to be accounted for more than once even with the same basic constituent category due to variation of the relation of morphological containing. German  $-(e)t(-)^L$ , for instance, appears as a suffix immediately after the word stem indicating past tense (forming a so-called ‘inflectional stem’), but as a suffix at the very end of the word form it indicates second person plural (or third person singular in the present tense). In a form as (*ihr*) *red-et-et* (‘you(pl) talked’), it appears twice, to the effect that the difference of the m-containing relations can be seen.

### 5.3 *Marking effects of inflectional units — specificity of inflectional units and language types*

#### *Marking effects of inflectional units*

After this brief excursion into the theory of language description, let us consider the functional categories that forms are ‘marked for’ or ‘indicated for’ or ‘cross-referenced for’ by inflectional units. They are the *marking effects* of the inflectional units:

- (44) *Definition.* [S, ⟨P,  $b^0$ ⟩, K, and  $K_0$  are as in (38).] K is a *marking effect* of ⟨P,  $b^0$ ⟩ relative to  $K_0$  and S iff there is a  $J_1$  and a  $J_2$  such that
- a. ⟨ $J_1$ ,  $J_2$ ⟩ is a marking pair for ⟨P,  $b^0$ ⟩ relative to  $K_0$  and S;
  - b.  $K \in J_2$

So,  $-1s$ ,  $+2s$ ,  $-3s$ ,  $-1o$ ,  $-2o$ , and  $+3o$ , all are marking effects of  $re^L$  relative to Verb Form and Guaraní idiolect systems. Of course, for a form being

assigned to  $-1\mathbf{o}$ ,  $-2\mathbf{o}$  and  $+3\mathbf{o}$ , a second condition has to be satisfied (viz., being a form that contains a verb stem that occurs with both,  $a^{-L}$  and  $che^{-L}$ , and that does not contain  $je^{-L}$  nor  $jo^{-L}$ , since the relevant pair in the system connection is  $\langle \{[N-je-jo-], [re-]\}, \{-1\mathbf{o}, -2\mathbf{o}, +3\mathbf{o}\} \rangle$ ; cf. (22) and subsec. 3.1) while  $-1\mathbf{s}$ ,  $+2\mathbf{s}$  and  $-3\mathbf{s}$  can be ascribed unconditionally to any form that contains  $re^{-L}$ ; cf. (18). This leads to two further definitions (variables not introduced here are as given in (38)):

(45) *Definitions.* [S,  $\langle P, b^0 \rangle$ , K, and  $K_0$  are as in (38).]

- a. Let K be a marking effect of  $\langle P, b^0 \rangle$  relative to  $K_0$  and S, and  $\langle J_1, J_2 \rangle$  a relevant marking pair of  $\langle P, b^0 \rangle$  relative to  $K_0$  and S as required by (44), and  $K_1$  any set of syntactic units.  $K_1$  is a *co-condition of*  $\langle P, b^0 \rangle$  for K relative to  $J_1, J_2, K_0$  and S iff
  - (i)  $K_1 \in J_1$
  - (ii)  $K_1$  is *not* a  $\langle P, b^0 \rangle$ -category for  $K_0$  in S
- b. Any K is an *unconditional marking effect of*  $\langle P, b^0 \rangle$  relative to  $K_0$  and S iff
  - (i) K is a marking effect of  $\langle P, b^0 \rangle$  relative to  $K_0$  and S
  - (ii) for any  $\langle J_1, J_2 \rangle$  that satisfies (44), there are no co-conditions of  $\langle P, b^0 \rangle$  for K relative to  $J_1, J_2, K_0$ , and S

Reformulating the above example, we can state that  $-1\mathbf{s}$ ,  $+2\mathbf{s}$  and  $-3\mathbf{s}$  (for short: ‘second person singular subject’) are unconditional marking effects of  $re^{-L}$  within Verb Form in Guaraní idiolect systems S, while [*a-che-stem*] is a co-condition for  $-1\mathbf{o}$ , as well as for  $-2\mathbf{o}$  and  $+3\mathbf{o}$  relative to the second pair in (18) and the set of verb forms in Guaraní idiolect systems.

#### *Specificity of inflectional units*

Consider again the system connection of Guaraní idiolect systems as given in subsection 3.3, and compare functional categories such as  $+1\mathbf{s}$  or  $-3\mathbf{o}$  with other ones such as *Desiderative* or *Reciprocal*. The first ones figure as marking effects of several inflexional units, whereas the second ones appear only once. They are *specific to* the inflexional units that underlie the structural categories in the first components of the respective pairs — usually, there is even only one such structural category (this is indeed the case for *Desiderative*, *Command*, *Reciprocal*, etc.). We define:

(46) *Definition.* [S,  $\langle P, b^0 \rangle$ , K, and  $K_0$  are as in (38).] Any K is a *specific marking effect of*  $\langle P, b^0 \rangle$  relative to  $K_0$  and S iff

- a. K is a marking effect of  $\langle P, b^0 \rangle$  relative to  $K_0$  and S
- b. there is exactly one  $\langle J_1, J_2 \rangle$  such that  $K \in J_2$



Applying this concept to the verb system of Guaraní, we see that **Command** is a specific marking effect of  $e^{-L}$ , and **Reflexive/Passive** and **Reciprocal** are specific effects of  $je^{-L}$  and  $jo^{-L}$ , respectively, and so is **Desiderative** of  $ta^{-L}$ , since these categories appear only once in the system connection presented above, and in the relevant pairs (elements of the system connection) categories of  $e^{-L}$  (viz., [e-]) etc. are elements in the first component.

Finally, we can call an inflexional unit itself *specific*: the categories of a specific inflexional unit do not occur together with any other structural category in the system connection.

(47) *Definitions.* [S,  $\langle P, b^0 \rangle$ , K, and  $K_0$  are as in (38).]  $\langle P, b^0 \rangle$  is a *specific* inflexional unit relative to  $K_0$  and S iff for all  $\langle J_1, J_2 \rangle$  in the marking content of  $\langle P, b^0 \rangle$  relative to  $K_0$  and S,  $J_1$  has exactly one element.  $\langle P, b^0 \rangle$  is *highly specific* if there exists only one such  $\langle J_1, J_2 \rangle$  and also  $J_2$  has only one element.

In order to avoid redundancy, it could be required of any system connection that there is only one marking pair in the marking content of any specific inflexional unit. This and other practical restrictions of system connections have been present in descriptions within the IL framework without being part of the definition of system connection.

#### *Inflectional units and language types*

In the Guaraní verbal system, we encounter only three specific prefixes:  $ta^{-L}$ ,  $i^{-L}$ , and  $po^{-L}$ . Of these, only  $ta^{-L}$  is highly specific. This may be seen as a sign that (as far as the verbal prefix system is concerned) Guaraní is of the inflectional, not agglutinative type of languages (assuming a simple traditional typology), if we accept the following theorem that should be argued for on another occasion:

(48) *Theorem.* The more inflectional units in a language system are specific or have specific or unconditional effects, the more the language is of the ‘agglutinative type’ (if the inflexional units are mainly affixes) or of the ‘isolating type’ (if the inflexional units are mainly auxiliaries). The less such inflexional units exist, the more the language tends to the ‘inflectional type’.

Still, some words must be said about Guaraní: if there was no hierarchy of reference, Guaraní would tend significantly more to the agglutinative language type. Also, if we adhered to the traditional categories (that we have been able to reconstruct as intersections of sets of our categories), the degree of specificity of the person prefixes would be considerably higher. It might well be said that our somewhat scrupulous analysis gave Guaraní a less agglutinative and more inflecting appearance, but that traditional descriptions of Guaraní that saw it as an agglutinating language had their right, too. (This is supported by results from analyses of the suffixes and the nominals.)

It is interesting that the tendency of dissolving the reference hierarchy appears first with two prefixes, *po*<sup>-L</sup>, that is specific, and *ro*<sup>-L</sup>, that is the ‘least specific’ (its structural category appears six times in the system connection of Guaraní idiolect systems — more than those of any other prefix), due to syncretism. If it was not for *ro*<sup>-L</sup>, the reflexive and reciprocal marking prefixes *je*<sup>-L</sup> and *jo*<sup>-L</sup> would be highly specific prefixes, too.

Similar signs of dissolution of the proposed original reference hierarchy of proto-Tupí-Guaraní are found among other Tupí-Guaraní languages, and with the instruments developed here, a comparison might prove useful. Such a comparison could possibly even serve as a type of criterion in genetical language grouping of its own. Due however to time and space limits of this article, it will not even be commenced at this point.

### References

- Gregores, E. and Suárez, J. A. (1967). *A Description of Colloquial Guaraní*, volume XXVII of *Janua Linguarum, Series Practica*. Mouton, Den Haag, Paris.
- Hausmann, F. J., Reichmann, O., Wiegand, H. E., and Zgusta, L., editors (1989ff). *Wörterbücher: Ein internationales Handbuch zur Lexikographie*, volume 5 of *Handbücher zur Sprach- und Kommunikationsforschung*. de Gruyter, Berlin / New York. 3 Bände.
- Jensen, C. (1990). Cross-referencing changes in some tupí-guaraní languages. In Payne (1990), pages 117–158.
- Lieb, H.-H. (1983). *Integrational Linguistics, Bd. 1: General Outline*, volume 17 of *Current Issues in Linguistic Theory*. John Benjamins, Amsterdam, Philadelphia.
- Lieb, H.-H. (1992). Paradigma und klassifikation: Explikation des paradigmabegriffs. *Zeitschrift für Sprachwissenschaft*, 1(11):3–46.
- Monseñat, R. and Facó-Soares, M. (1970). Hierarquia referencial am línguas tupí. Museu Nacional do Rio de Janeiro.
- Payne, D. L., editor (1990). *Amazonian Linguistics: Studies in Lowland American Languages*. Univ. of Texas Press, Austin.
- Rodrigues, A. D. (1985/86). Relações internas na família lingüística tupí-guaraní. *Revista de Antropologia (São Paulo)*, 27/28.
- Rodrigues, A. D. (1990). You and i is neither you nor i: The personal system of tupinamba. In Payne (1990), pages 393–405.
- Zgusta, L. (1989). The influence of scripts and morphological language types on the structure of dictionaries. In Hausmann et al. (89ff), chapter 32, pages 296–304.