

Tense, Aspect and Mood in Awetí Verb  
Paradigms:  
Analytic and Synthetic Forms

*Suggested short title:*  
Tense, Aspect and Mood in Awetí Paradigms

*Sebastian Drude*

Museu Paraense Emílio Goeldi (Brazil)

Sebastian.Drude@gmail.com

Sebastian Drude  
a/c Lingüística, CCH — Museu P.E. Goeldi  
Caixa Postal 399 — BELÉM — PA — CEP: 66040-170 —  
BRAZIL

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Verbs in a Word-and-Paradigm Approach</b>	<b>3</b>
2.1	Lexical Words and Word Forms . . . . .	3
2.2	Classification Systems . . . . .	6
<b>3</b>	<b>Awetí Verb Paradigms: The Person Part</b>	<b>8</b>
3.1	Three Awetí verb types . . . . .	8
3.2	Active intransitive verbs and basic person categories . . . . .	9
3.3	Transitive verbs and person hierarchy . . . . .	11
3.4	Stative intransitive verbs . . . . .	14
<b>4</b>	<b>Tense-Aspect-Mood Category Affixes</b>	<b>15</b>
4.1	Permissive mood prefixes . . . . .	15
4.2	Aspect suffixes . . . . .	17
4.3	Mood suffixes . . . . .	19
<b>5</b>	<b>Tense-Aspect-Mood Auxiliary Particles</b>	<b>23</b>
5.1	A particle for the negated permissive . . . . .	23
5.2	‘Temporal’ and ‘modal’ particles: factuality . . . . .	27
<b>6</b>	<b>The Verb Form System and System Link of Awetí</b>	<b>31</b>
6.1	The functional system . . . . .	31
6.2	The structural system . . . . .	32
6.3	The system link . . . . .	34
<b>7</b>	<b>Elements of Awetí Verb Paradigms: Examples</b>	<b>40</b>
7.1	Synthetic form <i>atupeju</i> <sup>1</sup> : Problems with traditional glossings . . . . .	40
7.2	Analytical form <i>tut etoka</i> : analysis of discontinuous occurrence . . . . .	41
7.3	Syncretism for <i>pejtup</i> <sup>1</sup> (synthetic form of a transitive verb) . . . . .	43
7.4	Syncretism for <i>etup tepe</i> (analytical form of a transitive verb) . . . . .	44
7.5	Synthetic form <i>ito</i> : no specific marking at all . . . . .	45
7.6	Final remarks . . . . .	45
	<b>References</b>	<b>46</b>

# **Tense, Aspect and Mood in Awetí Verb Paradigms: Analytic and Synthetic Forms**

Sebastian Drude

Museu Paraense Emílio Goeldi

## **Abstract**

This paper describes the verbal Tense-Aspect-Mood system of Awetí (Tupian, Central Brazil) in a Word-and-Paradigm approach.

One classification of Awetí verb forms contains clear aspect categories. A second set of independent classifications renders at least four moods and contains a third major TAM classification, factuality, that has one mainly temporal category Future, while others are partially or wholly modal.

Structural categories reflect the formal composition of the forms. Some forms are synthetic, ‘marked’ only by means of affixes, but many are analytic, containing auxiliary particles.

With selected sample forms we demonstrate in detail the interplay of structural and functional categories in Awetí verb paradigms.

## **1 Introduction**

This paper describes the core of the Tense-Aspect-Mood (T-A-M) system of verbs in the Awetí language, using a Word-and-Paradigm approach.<sup>1</sup>

Awetí is a Tupian language spoken by ca. 140 people in central Brazil in the Upper Xingú area. Awetí does not belong to, but is arguably the closest relative of the well-known Tupí-Guaraní subfamily, the largest branch of the Tupí stock.

---

<sup>1</sup> I am very grateful to the people that contributed to this paper by discussion and comments on earlier versions, especially to H.-H. Lieb, M. Budde, H. v.d. Voort, S. Meira and David Rood and the participants of the IL-Forum and of the DOBES congress 2004. Of course this does not imply that any of these is responsible for any flaws of this paper.

The following analysis is a preliminary result of seven years of work on the language, including a total of some eight months of fieldwork. From 2001 to 2005 the project was supported by the Volkswagen Foundation within the DOBES (Documentation of Endangered Languages) program, focusing on documenting the language and aspects of the culture.

To describe and gloss the data for this language, we rely on the framework of Integrational Linguistics. This is a neo-structuralist, declarative, axiomatic approach developed mainly by Hans-Heinrich Lieb. It can be characterized as a Word-and-Paradigm framework, within the Western tradition of language description. Although most of the descriptive conceptions are intuitively accessible, some space is dedicated to presenting an outline of the underlying concepts of the Integrational approach to paradigms and their forms. Since the framework will be unknown to most readers, who are likely used to some flavor of “Item-and-Arrangement” or “Item-and-Process” approaches,<sup>2</sup> it seems to be a justified secondary aim of the paper to demonstrate some advantages of applying this approach in language description, perhaps of interest to general readers, independent of the facts in Awetí.

As will be shown, the Integrational approach allows direct description of paradigms the forms of which are construed by morphological *and* syntactic means. In particular, inflexional affixes and auxiliaries play analogous roles in construing word forms, simple (synthetic) and analytical (‘periphrastic’) ones. The latter are notoriously problematic in approaches that conceive word paradigms as morphological and not syntactic, especially if they occur as a discontinuous constituent. At the same time, no zero morphemes are needed as would be the case in other current frameworks. Furthermore, the model is ‘declarative’, that is, its core aim is to formulate true statements (hypotheses) about linguistic entities, starting from actually observed (surface) units and assigning a meaning to them based entirely on their structural properties. So, this approach is ‘hearer-oriented’: in identifying units, categories and structures it proceeds from form to meaning.

In section 3 we give an overview over the person marking systems of the different types of Awetí verbs. The relevant prefixes in the indicative and other moods contrast with those for imperative or permissive mood,

---

<sup>2</sup> These and the label “Word-and-Paradigm” refer to a well-known paper of Charles Hockett (1954) where he identifies three major types of linguistic theories.

which are presented in section 4 together with the other affixes discussed in this paper: a set of two suffixes for progressive and imperfective aspect, and another set of suffixes which mark the subjunctive and the so-called 'gerund' mood.

In section 5 we turn to auxiliary particles used to build analytical verb forms. These include a particle used in the negated permissive forms (there are also negative forms in other moods) and a series of mutually exclusive particles occurring in the 'second' position of a clause that mark future tense and other categories such as irrealis, which are better characterised as factuality categories.

We then turn to the core components of the "basis" for verbal paradigms in Awetí in section 6. First we summarize the functional classification system which yields the functional verbal categories identified so far. Then we present the structural classification system whose categories are based on the formal composition of verb forms (synthetic and analytical ones). Both systems are linked by the "system link": structural categories are the basis for assigning a form to its relevant functional categories.

Finally, in section 7 we apply this framework by discussing a sample of verb forms and demonstrate how the respective elements of the paradigms are obtained.

## 2 Verbs in a Word-and-Paradigm Approach

### 2.1 *Lexical Words and Word Forms*

In the framework used in this paper, a fundamental distinction is made between lexical words on the one hand and their grammatical forms ("word forms") on the other.<sup>3</sup> They are ontologically different. *Lexical words* consist of two components:

- a word paradigm (see below), and
- a concept which is the lexical meaning of the word.<sup>4</sup>

<sup>3</sup> For details on paradigms, see Lieb (2005, 1992b). For an overall view of the framework, see Lieb (1983, 1992a), and in particular on syntax Lieb (1993). The website [www.germanistik.fu-berlin.de/il](http://www.germanistik.fu-berlin.de/il) gives information on the approach, including a comprehensive bibliography.

<sup>4</sup> The Integrational approach to lexical semantics cannot be explained here. It must suffice to say that concepts usually have a "content", which is a property of, or an intensional relation between, real-world entities. For a general presentation of the approach see

For instance, the lexical word *totu*<sup>W</sup> (an Awetí intransitive verb) is a pair of the paradigm *totu*<sup>P</sup> and the concept ‘to go’, as shown in (1).<sup>5</sup>

$$(1) \textit{totu}^W = \langle \textit{totu}^P, \textit{‘to go’} \rangle$$

On the other hand, we have word forms, forms of lexical words. For instance, *went*<sup>1</sup> (see next paragraph for the superscript “1”) and *has gone* are two forms of the English verb *to go*<sup>W</sup>. Word forms, not lexical words, occur in sentences and belong to the usual grammatical categories. For instance, *went*<sup>1</sup> belongs to the category **Past Tense**, among others, and *has gone* belongs to (is an element of) **Third Person, Singular** etc.

Ontologically, word forms are conceived as sequences of phonological words. This may come as a surprise, but it allows right away for the proper treatment of analytical forms such as *has gone*. This form is the sequence of the two phonological words *has* and *gone*, which are the members of

<sup>5</sup> The following *conventions* are used: (1) lexical words and paradigms are cited with superscript “W” or “P”, respectively. As **citation form** of verbs we do not use the bare stem but the form which Awetí speakers use for referring to the word. This form usually contains the suffix *-(t)u* (see section 4.3, especially footnote 24 for variants) and possibly a prefix such as *n(ā)-* (a third person prefix), *t-* or *to-* (see next section). In this paper, in order to facilitate the identification of the stems of Awetí verbs, these are usually put in bold face. (2) **Raised dots** ‘...’ distinguish names of concepts (lexical meanings), different from **single quotation marks** ‘...’ which are used to give all other sorts of meanings, especially meanings of syntactic units such as word forms or sentences. (3) Ordered pairs are represented by putting brackets ⟨...⟩, not parentheses, around the names of their components, which are separated by commas. ‘⊃’ reads: ‘contains as element(s)’. (4) Names of syntactic form categories have word-initial capital letters, names of categories of lexical words are capitalized throughout. See footnote 8. In formal contexts, both may be in a sans serif font. (5) Abbreviations (not self-explicatory or given by the context) used are: (a) **general**: **LWO**: Lexical Word Ordering, **SL**: system link, **SUO**: Syntactic Unit Ordering; (b) **constituent categories** (see sections 5.1 and 7.2): **Pf**: Particle Form, **Vf**: Verb Form, **VGr**: Verb Group; (c) **verb form categories**: **Fact**: Factual, **Frust**: Frustrative, **Fut**: Future, **Ger**: Gerund, **lpfv**: Imperfective, **Ind**: Indicative, **irr**: Irrealis, **Neg**: Negative, **O.ct / Obj.-ctrd**: Object-centred, **Pfv**: Perfective, **Pms**: Permissive, **Pos**: Positive, **Prog**: Progressive, **Reass**: Reassured (Future), **Recp**: Reciprocal, **Rfl**: Reflexive, **S.ct / Subj.-ctrd**: Subject-centred, **Subj**: Subjunctive. See also section 3.2 for numbers as person categories. (d) **lexical word categories** (see section 3.1): **TR.VB**: TRANSITIVE VERB, **ITR.VB**: INTRANSITIVE VERB, **ACT.ITR.VB**: ACTIVE ITR.VB, **STV.ITR.VB**: STATIVE ITR.VB.

the sequence.<sup>6</sup> In this view, it is not analytical forms but rather synthetic word forms that are the borderline case. These are often one-member or unit sequences such as the form *went*<sup>1</sup>, a sequence with only one member, the phonological word *went*. The superscript “1” is applied to distinguish unit sequences from their only member; in this case, synthetic word forms from phonological words.

All forms of a lexical word, each together with its categorization (the relevant grammatical categories a form belongs to), build the word’s first component, the paradigm. This will be further explained and illustrated in the next sections.

Please note that this approach to lexical words holds for content words and ‘function words’ alike. This is achieved by allowing for the limiting case of so-called “improper paradigms”, which are often paradigms with only one element, i.e., one form together with its classification.<sup>7</sup> Also, the so-called “empty concept” *b*<sup>0</sup>, which is the only concept without content, may occur as the second component of lexical words.

(2)  $tut^w = \langle tut^p, b^0 \rangle$  ( $tut^w$  is a ‘future marker’, see below sec. 5.2.)

For instance, the lexical word *tut*<sup>w</sup> in (2) is an Awetí auxiliary particle. As such, it has no lexical semantic content, that is, its lexical meaning is the empty concept *b*<sup>0</sup>. Also, the paradigm *tut*<sup>p</sup> in (2) is an improper one – indeed, it has only one form, *tut*<sup>1</sup>. This form occurs in verb forms, for instance within *ato tut* ‘I will go’, a form of the verb *totu*<sup>w</sup>. Both paradigm-meaning-pairs *totu*<sup>w</sup> in (1) and *tut*<sup>w</sup> in (2) are lexical words, with different functions and internal complexity, but of the same ontology.

Just as word forms belong to grammatical categories such as ‘Past Tense’, lexical words belong to syntactic categories, but of a different type. In particular, each lexical word belongs to a part of speech (such as the category **VERB**), and possibly to subcategories (such as **INTRANSITIVE VERB**). So there are two types of syntactic categories: a) sets of syntactic

<sup>6</sup> Sequences are named by writing the members one behind the other, as in *has gone*, the sequence of *has* and *gone*, in that order. — The term “phonological word” used in this paper has a different sense than that with which it occurs in recent functional literature, e.g. in Dixon and Aikhenvald (2002): phonological words in our sense (including, for instance, clitics) are precisely *defined* as the members of word forms; they may or may not be *identifiable* by phonological criteria alone.

<sup>7</sup> A paradigm with several forms may be improper, too, if its forms are all equivalent, that is, if they have all the same categorization.

units, in particular, of word forms (tense, aspect and mood categories belong to this type), and b) sets of lexical words (among these, the parts of speech).<sup>8</sup>

## 2.2 Classification Systems

Categories of any type always belong to some classification. A classification of a set A contains (is a set of) subsets A<sub>1</sub>, A<sub>2</sub> etc. of A. For instance, the classification (*main*) *parts of speech* classifies the set of all lexical words of a language into categories such as **VERB**, **SUBSTANTIVE** etc. (if these exist in the language), which are also sets of lexical words, subsets of the starting set which is classified by (*main*) *part of speech*. The categories (or classes) of a classification together exhaust the starting set. That is, each element of A belongs to at least one element (=class, category) of the classification of A. In the example, every lexical word belongs to at least one part of speech.

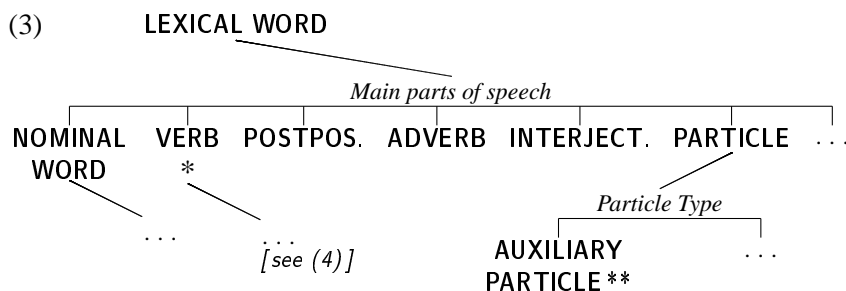
Some relevant classifications in the domain of lexical words of Awetí are indicated in (3). When several classifications are combined, we speak of a classification system on a (starting) set A. The classification system on the set of lexical words is called the lexical word ordering (LWO).<sup>9</sup> Universally, it contains the classification of lexical words into the main parts of speech. In (3) it is combined with subclassifications, in particular with one which subdivides the particle words of Awetí. The existence of other classifications is indicated by dots. The signs \* and \*\* identify categories to which *totu*<sup>w</sup> and *tut*<sup>w</sup> respectively belong.

Grammatical categories such as ‘Singular’ or ‘Progressive’ are sets not of lexical words but of word forms. Word forms (including analytical ones)

<sup>8</sup> In order to avoid confusion, names of categories of lexical words use only upper case letters, while names of categories of word forms or other syntactic units contain upper-case and lowercase letters. In informal contexts (especially before the categories are introduced) the names may be used without capital letters.

<sup>9</sup> Budde (2000) is an in-depth general study on Word Classes. — The conventions in graphical representations of classification systems should be easily understood: the slanted lines connect the name of a classification (below) with the name of the set (above) which is subdivided by the classification, while below the vertical lines the names of the resulting subclasses are given. They are elements of the classification which is named above the connecting horizontal line. Dots indicate that there are one or more classes or classifications which are left unspecified.





are simple syntactic units. Concatenated they form more complex syntactic units, “word groups” (including ‘phrases’ and ‘sentences’). All these units are sequences of phonological words.

The set of all syntactic units is the starting point of a second classification system of every language. This classification system is called the syntactic unit ordering, or SUO, of the language.<sup>10</sup> The SUO provides categories such as the set of word forms and its subsets, in particular, the grammatical categories.<sup>11</sup>

Much of the following sections is devoted to that part of the SUO of Awetí which concerns the tense, aspect and mood categories to which Awetí verb forms belong (cf. especially section 6). These categories form a central part of the “basis” for paradigms (such as *totu*<sup>P</sup>) of Awetí verbs (such as *totu*<sup>W</sup>, cf. (1)).

Together with other components, both classification systems, the LWO and the SUO, belong to the syntactic part of each language system. Indeed, all entities discussed so far belong to syntax—they are all constructs built on phonological words. This holds for word paradigms which are constructs of these: word forms and categories of word forms. It holds even for lexical words although they contain concepts besides word paradigms.

So, in the view taken here neither paradigms nor parts of speech be-

<sup>10</sup> Indeed, all linguistic entities have to be related in some way to the respective language (or more precisely, in order to cope with inner language variability, to *idiolect systems*): the set of verbs of English is obviously different from the set of verbs of Awetí. In this paper, however, explicit reference to languages (idiolect systems) will as a rule be omitted.

<sup>11</sup> In other terminologies, the term ‘(grammatical) category’ is used for what we call *classifications*, such as ‘number’ or ‘tense’. *Categories* (or *classes*, elements of classifications) in the traditional sense used here correspond in such terminologies to the ‘values of categories’, such as ‘Singular’ or ‘Present Tense’.

long to morphology. Morphology is mainly concerned with the inner composition of the phonological words, where they are analysed into *morphological units*: these in turn are sequences of morphs, in analogy to syntactic units (sequences of phonological words). Simple morphological units are forms of stems or affixes. Stems and affixes are both lexemes (including morphemes), which are, parallel to lexical words, pairs of a morphological paradigm and a concept (in the case of affixes, always the empty concept).

In this approach, morphology and syntax are largely analogous, ontologically and in other details. So, morphology is related to word paradigms only indirectly, by providing criteria for classes that represent the formal structure of word forms, see especially section 6.2. But the existence of morphologically determined categories does not affect the syntactic status of word forms, their classes, or word paradigms.

In the following section, we will characterize in more detail the conception of syntactic paradigms such as *totu*<sup>P</sup> by giving an overview of the person system, preparing the ground for categories specific to the tense-aspect-mood system presented in sections 4 and 5.

### 3 Awetí Verb Paradigms: The Person Part

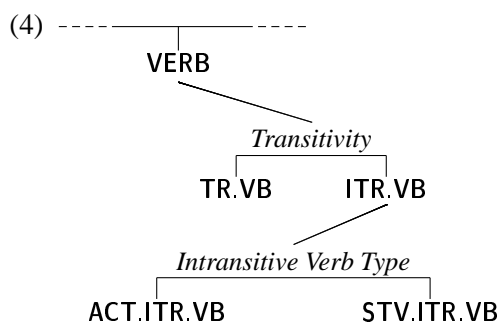
#### 3.1 Three Awetí verb types

When characterising Awetí verb paradigms, we have to distinguish three types of verbs. First, there is a distinction between transitive and intransitive verbs in Awetí: The former have forms that agree with or relate to the subject, “subject-centred forms”, as well as forms that relate to the object, “object-centred forms”. Intransitive verbs lack the latter.

The intransitive verbs fall into two different classes by morphological criteria which (with some exceptions) reflect semantic properties. The ‘active’ intransitive verbs have a form (that of the ‘first person singular’) that resembles the subject-centred form of transitive verbs. The corresponding form of the ‘stative’ verbs is similar to the object centred form. hence, at least as for the ‘first person singular’ forms, Awetí may be said to be a ‘active’ language or to have a ‘split-S’ or ‘split intransitive system’.<sup>12</sup>

<sup>12</sup> For this conception see, e.g., Mithun (1981); for a similar claim, in Klimov’s (e.g., 1977) terms, for the Tupí-Guaranian language Kamaiurá, see Seki (1990).

We can summarize these verb classes and their relations graphically as in (4), continuing the Lexical Word Ordering of Awetí, the classification system outlined in (3).<sup>13</sup>



### 3.2 Active intransitive verbs and basic person categories

Let's start by describing Awetí person marking by different forms of our example *totu*<sup>w</sup>, cf. (1). This word is an active intransitive verb. Its six basic forms are shown in (5).<sup>14</sup>

(5)	<b><i>ato</i></b> <sup>1</sup>	'I went'	:	1st P. Singular		+1	-2	-3
	<b><i>eto</i></b> <sup>1</sup>	'you <sub>(sg)</sub> went'	:	2nd P. Singular		-1	+2	-3
	<b><i>oto</i></b> <sup>1</sup>	'(s)he/it/they went'	:	3rd P.		-1	-2	+3
	<b><i>kajto</i></b> <sup>1</sup>	'we <sub>(incl)</sub> went'	:	1st P. Plural Incl		+1	+2	+3
	<b><i>ozoto</i></b> <sup>1</sup>	'we <sub>(excl)</sub> went'	:	1st P. Plural Excl		+1	-2	+3
	<b><i>e'ito</i></b> <sup>1</sup>	'you <sub>(pl)</sub> went'	:	2nd P. Plural		-1	+2	+3

In the first column, the forms are given orthographically.<sup>15</sup> The sec-

<sup>13</sup> The LWO as part of the Awetí language system may still differ from (3) and (4). For one thing, there have to be syntactical / functional criteria for establishing that a word class belongs to the LWO. Morphological or other formal criteria alone are not sufficient, not even combined with semantic criteria. — It would also have been possible to first separate the stative verbs from all 'active' ones, and divide these into transitive and intransitive (active) verbs.

<sup>14</sup> In this paper, in tables such as (5), where sets of forms are presented, raised dots ':' help to identify the different morphs inside the (main) phonological word. The stem usually stands out in boldface.

<sup>15</sup> Observe their ontological status: although synthetic forms, *ato*<sup>1</sup>, *eto*<sup>1</sup>, *oto*<sup>1</sup> etc. are sequences of phonological words – sequences with only one member (unit sequences), hence the superscript "1".

ond column contains a usual translation of the preceding form uttered in isolation. In the third column, the forms are labelled with traditional person and number categories. Note that there is no number distinction for third p(erson). However, there is an additional distinction in the first person plural between ‘incl(usive)’ (including the hearer) and ‘excl(usive)’ (excluding the hearer). I presented elsewhere (cf. Drude 2007) an analysis for the related Tupí-Guaranian language Guaraní that by and large carries over to Awetí. According to this analysis, these two facts are related and can be easily accounted for by assuming *three* classifications for participation of speaker, hearer and others, respectively, instead of *one* person classification and one number classification as in standard European languages.

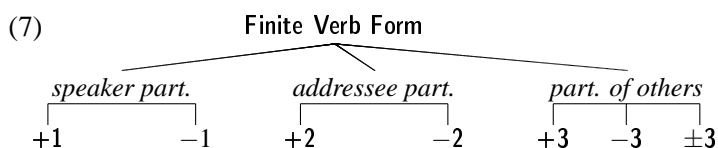
The relevant categorizations that result from this analysis are given in the last columns. They contain three categories each, the names of which are abbreviated using the numbers 1, 2 and 3. For instance, “+1” reads: “**Speaker Participates**”, “-2” “**Hearer Does Not Participate**”, and “±3” “**Non-specific As To Participation Of Others**”.

The table in (5) characterizes a section of the paradigm  $totu^P$ , the first component of the lexical word  $totu^W$ . Now, what exactly do we understand by a paradigm? In the framework used here any paradigm is formally conceived as a set of pairs (a relation). Each pair has one form of the word as its first component and a set of functional syntactic categories (‘values of grammatical categories’) as its second component. In each pair, the form (first component) belongs to each of the categories in the second component. The second component of each pair is called a categorization of the form which is the first component. In sum, a paradigm assigns or relates categorizations to forms.

Given this explanation of the notion of paradigm, each line in the table in (5) corresponds to an element of the paradigm  $totu^P$ . In a more formal notation, we represent the corresponding part of the paradigm  $totu^P$  in (6). The dots indicate that there are more categories in the categorizations, and the final dots in the last line that there are more elements in the paradigm. Many of these will be added in the next sections.

$$(6) \quad totu^P = \{ \langle ato^1, \{+1, -2, -3, \dots\} \rangle, \langle eto^1, \{-1, +2, -3, \dots\} \rangle, \\ \langle oto^1, \{-1, -2, +3, \dots\} \rangle, \langle kajto^1, \{+1, +2, \pm 3, \dots\} \rangle, \\ \langle ozoto^1, \{+1, -2, +3, \dots\} \rangle, \langle e'ito^1, \{-1, +2, +3, \dots\} \rangle, \dots \}$$

As said in the last section, the functional categories presented in (6) are given by the SUO, a classification system on the set of all syntactic units. In (7) we show a first part of that branch of this classification system which subdivides the set of all (finite) verb forms.<sup>16</sup> Because three person categories (instead of one category for person and one for number) are combined in each categorization in (6), at least three independent classifications are needed. We assume the three classifications shown in (7): one each for part(icipation) of speaker, addressee and others.



### 3.3 Transitive verbs and person hierarchy

We turn next to the description of the person marking in the forms of transitive verbs. As an example, we choose *nātipu*<sup>w</sup> ‘to see’. There is again a set of six basic forms which refer mainly to the person of the subject, such as *atup*<sup>1</sup> ‘I saw you or him/her/it/them’, or ‘I saw 2/3’, for short, where the *a-* ‘marks’ the ‘First Person Singular’, as above.<sup>17</sup> I call these forms subject-centred. They are characterised in the table in (8). Note that the prefixes in these forms are quite different from those of active intransitive verbs, except for those of the ‘first and second person singular’ (the *-3*-Forms, on our account).

As the glosses indicate, the use of these forms implies that the object is inferior in a hierarchy of reference where first person is superior to second person and both are superior to third person. (In the case of ‘third person’, the object is a *different* third person, or ‘fourth person’.) There is a second set of forms that are used when the object is superior to the subject on the

<sup>16</sup> Only the three person-classifications have the set of finite verb forms as their basis; all others classify the set of all verb forms, including the non-finite ones. For reasons of space, we do not consider the latter here.

<sup>17</sup> In morphology, we will not use ontological labels on names of lexemes or affix and stem forms, and use formulations as “*a-* ‘marks’ ...” instead of “occurrences of the form *a*<sup>-1</sup> of the affix(-lexeme) *a*<sup>-1</sup> ‘mark’ ...”. Wherever necessary, these notions can be recovered from the more informal expressions used here.

- (8) *atup*<sup>1</sup> ‘I saw 2/3’ : +1 -2 -3 (‘1st P. Sing.’)  
*etup*<sup>1</sup> ‘you<sub>(sg)</sub> saw 3’ : -1 +2 -3 (‘2nd P. Sing.’)  
*wejtup*<sup>1</sup> ‘(s)he/it/they saw 3’ : -1 -2 +3 (‘3rd P.’)  
*titup*<sup>1</sup> ‘we<sub>(incl)</sub> saw 3’ : +1 +2 ±3 (‘1st P. Pl. Incl.’)  
*ozojtup*<sup>1</sup> ‘we<sub>(excl)</sub> saw 2/3’ : +1 -2 +3 (‘1st P. Pl. Excl.’)  
*pejtup*<sup>1</sup> ‘you<sub>(pl)</sub> saw 3’ : -1 +2 +3 (‘2nd P. Pl.’)

same hierarchy, such as *itup*<sup>1</sup> ‘you/he/she/it/they saw me’, or below for short ‘2/3 saw me’. We present these forms informally in (9).

- (9) *itup*<sup>1</sup> ‘2/3 saw me’ : +1 -2 -3 (‘1st P. Sing.’)  
*etup*<sup>1</sup> ‘3 saw you<sub>(sg)</sub>’ : -1 +2 -3 (‘2nd P. Sing.’)  
*kajtup*<sup>1</sup> ‘3 saw us<sub>(incl)</sub>’ : +1 +2 ±3 (‘1st P. Pl. Incl.’)  
*ozotup*<sup>1</sup> ‘2/3 saw us<sub>(excl)</sub>’ : +1 -2 +3 (‘1st P. Pl. Excl.’)  
*e’itup*<sup>1</sup> ‘3 saw you<sub>(pl)</sub>’ : -1 +2 +3 (‘2nd P. Pl.’)

Awetí can be said to have person-based split ergativity (cf., e.g., Dixon 1994), combined with a split-S system as most prefixes in these forms are different from those used in the subject-centred forms in (8) but the same as those in the case of intransitive active verbs, cf. (5). The only exception is in the First Person Singular where the prefix *a-* for the subject of transitive and active intransitive verbs is different from the prefix for object, *i-*, (*it-* before vowels). Informally, the marking of the First Person Singular follows an active or nominative pattern.

For the the Second Person Singular, the prefix is always *e-*, and this makes the corresponding forms in many transitive verbs ambiguous. In the sample case, *etup*<sup>1</sup> can mean ‘he/she/they saw you<sub>(sg)</sub>’ as well as ‘you<sub>(sg)</sub> saw him/her/them’ (usually the ambiguity is resolved only by context). So this form can be assigned both to the set of subject-centred forms and to this second set which I call ‘object-centred’. When a single inflected form is assigned to several categorizations in a paradigm (in other words, has several different functions, occupies different structural places), one speaks of syncretism.<sup>18</sup> (Cf. section 7.4.)

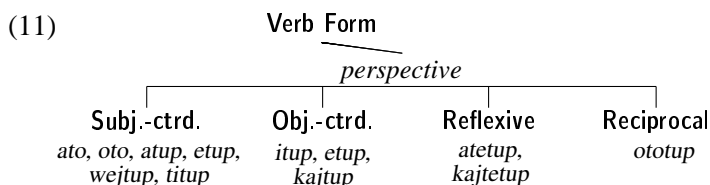
<sup>18</sup> In the words of Peter Matthews (in the Concise Oxford Dictionary of Linguistics, sv.), Syncretism is “the relation between words which have different morphosyntactic features but are identical in form. [...] Used especially when the identity is regular across all paradigms.” We understand ‘word’ as ‘word form’ and ‘having morphosyntactic features’ as ‘belonging to (a set of) syntactic categories’. The terms ‘homonymy/homophony’

There is only one third person prefix, *wej-*, for transitive verbs, distinct from *o-* in intransitive verbs. If one third-person referent acts upon another, forms with this prefix are used, which we here classify as subject-centred.

In the case that the referent of subject and object is the same, a different set of reflexive forms are used. In these forms there is an additional prefix *te-* between the stem and the person prefix, which in this case is of the set for active intransitive verbs. In the case of several referents acting reciprocally upon one another, the prefix *to-* is used instead of *te-* (not possible for –3-Forms). Compare the three sample forms in (10).

- (10) *ate-tup*<sup>1</sup> ‘I saw myself’  
*kajte-tup*<sup>1</sup> ‘we<sub>(incl)</sub> saw ourselves’ (everyone saw himself/herself)  
*oto-tup*<sup>1</sup> ‘they saw one another’

In terms of categories, we can account for these forms, as well as for the difference between subject-centred and object-centred forms, by assuming one more functional classification in the SUO in addition to those in (7). This classification, which I call ‘perspective’ here, is outlined in the diagram in (11).<sup>19</sup> Below the names of the classes I give some sample elements mentioned before (the superscript “1” has been omitted).



As the perspective classification applies to all (finite) verb forms, each form in a verb paradigm has one perspective category in its categorization. For illustration, some elements of the paradigm of the verb *nātupu*<sup>w</sup> are given in (12). Note that the syncretism involving the form *etup*<sup>1</sup> is reflected by two entries (elements) for this form in the paradigm. This implies that

---

should be reserved for forms (esp. paradigms) with different *lexical* meanings. — If the verb stem starts with a vowel, this syncretism does not occur, because then all prefixes in the subject-centred forms, but not of the object-centred forms, appear with an allomorph that ends in an additional *t*. For instance, *etētup*<sup>1</sup> ‘you<sub>(sg)</sub> hear 3’ is formally different from *eētup*<sup>1</sup> ‘3 hears you<sub>(sg)</sub>’.

<sup>19</sup> This classification corresponds in some respects to the voice or genus verbi classification in other languages.

the two classes **Subj.-ctrd.** and **Obj.-ctrd.** in (11) overlap.<sup>20</sup>

$$(12) \text{ nãtupu}^P = \{ \langle \text{atup}^1, \{+1, -2, -3, \text{Subj.-ctrd.,...}\} \rangle, \\ \langle \text{etup}^1, \{-1, +2, -3, \text{Subj.-ctrd.,...}\} \rangle, \\ \langle \text{wejtup}^1, \{-1, -2, +3, \text{Subj.-ctrd.,...}\} \rangle, \\ \langle \text{titup}^1, \{+1, +2, \pm 3, \text{Subj.-ctrd.,...}\} \rangle, \\ \dots \\ \langle \text{itup}^1, \{+1, -2, -3, \text{Obj.-ctrd.,...}\} \rangle, \\ \langle \text{etup}^1, \{-1, +2, -3, \text{Obj.-ctrd.,...}\} \rangle, \\ \langle \text{kajtup}^1, \{+1, +2, \pm 3, \text{Obj.-ctrd.,...}\} \rangle, \\ \dots \\ \langle \text{atetup}^1, \{+1, -2, -3, \text{Reflexive,...}\} \rangle, \\ \langle \text{kajtetup}^1, \{+1, +2, \pm 3, \text{Reflexive,...}\} \rangle, \\ \dots \\ \langle \text{ototup}^1, \{-1, -2, +3, \text{Reciprocal,...}\} \rangle, \dots \}$$

### 3.4 Stative intransitive verbs

There is another class of intransitive verbs: the stative verbs. Most prefixes of stative verbs align with those of the object-centred forms of transitive verbs. Languages with two such different classes of intransitive verbs are often called ‘split-S’ or ‘active-stative’ languages. But note that most prefixes (except the forms for the ‘First Person Singular’ and ‘Third Person’) are identical to those of active intransitive verbs anyway, so Awetí, differently from many Tupí-Guaranian languages, is mainly ergative (the ‘Second Singular’ forms have again the unspecific prefix *e-*). The nominative-accusative sub-pattern of ‘split ergativity’ is restricted to the first person singular. However, the paradigms of active and stative intransitive verbs differ in more respects than just with regard to the person prefixes.

For some Tupian (especially Tupí-Guaranian) languages, there is an ongoing discussion (cf., e.g., Queixalos 2001) on the status of stative verbs which sometimes are seen as nouns in copulaless predicative use. For Awetí, this view could be considered, at least for the female variety. But for reasons that cannot be explained here, the analysis as verbs is to be preferred.

<sup>20</sup> Two categories that are elements of a classification may overlap (share elements). It is excluded, however, that one category be a true subset of another category in the same classification.



To conclude this section, we give an informal account of the basic forms of a stative intransitive verb, *topetyj<sup>w</sup>* ‘to be sleepy’, in (13), without presenting the corresponding part of the paradigm formally. We assume that all forms of intransitive verbs, also of the stative type, are subject-centred, despite the similarity to the object-centred forms of transitive verbs.

- (13) *itopetyj<sup>1</sup>* ‘I am sleepy’ : +1 –2 –3 (‘1st P. Sing.’)  
*eopetyj<sup>1</sup>* ‘you<sub>(sg)</sub> are sleepy’ : –1 +2 –3 (‘2nd P. Sing.’)  
*topetyj<sup>1</sup>* ‘3 is/are sleepy’ : –1 –2 +3 (‘3rd Person’)  
*kajopetyj<sup>1</sup>* ‘we<sub>(incl)</sub> are sleepy’ : +1 +2 ±3 (‘1st P. Pl. Incl.’)  
*ozoopetyj<sup>1</sup>* ‘we<sub>(excl)</sub> are sleepy’ : +1 –2 +3 (‘1st P. Pl. Excl.’)  
*e’iopetyj<sup>1</sup>* ‘you<sub>(pl)</sub> are sleepy’ : –1 +2 +3 (‘2nd P. Pl.’)

Having presented the person prefixes and thus exemplified the conception of paradigms and functional categories used in this paper, the ground is prepared so that we now can turn to the description of tense, aspect, and mood categories in Awetí. We begin with categories marked by morphological means.

#### 4 Tense-Aspect-Mood Category Affixes

##### 4.1 Permissive mood prefixes

In Awetí as well as in many other languages, there are special verbal forms used to express a command, or rather, in the case of Awetí, a permission, by the speaker. Most of these forms differ from the forms presented in the previous section only with regard to the person prefixes.

In the singular, the prefix *i-* occurs in ‘permissive’ forms of intransitive verbs and the prefix *jo-* in forms of transitive verbs, instead of the usual prefix *e-*. It is unclear whether permissive forms for stative intransitive verbs do not exist or whether they are just difficult to elicit for semantic or pragmatic reasons. Luckily, there are some verbs that formally are stative verbs although they semantically describe actions, such as *ti’āpu’yp<sup>w</sup>* ‘to whistle’. However, acceptability of these forms is still unclear. Consider the forms in (14).

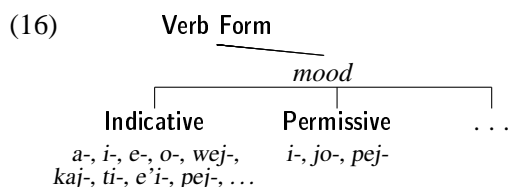
As exemplified in (15), in the plural analogous forms have the prefix *pej-*, for intransitive and transitive verbs alike. For transitive verbs, these forms are ambiguous (syncretism).

- (14) *ito*<sup>1</sup> ‘(you<sub>(sg)</sub> may) go!’ (cf. *eto*<sup>1</sup> ‘you went’)  
*jo-tup*<sup>1</sup> ‘(you<sub>(sg)</sub> may) look (it)!’ (*etup*<sup>1</sup> ‘y. saw/looked (it)’)  
*iti’āpu’yp*<sup>1</sup> ‘(you<sub>(sg)</sub> may) whistle!’ (*eti’āpu’yp*<sup>1</sup> ‘y. whistled’)

- (15) *pejto*<sup>1</sup> ‘(you<sub>(pl)</sub> may) go!’ (cf. *e’ito*<sup>1</sup> ‘you went’)  
*pejtup*<sup>1</sup> ‘(you<sub>(pl)</sub> may) look!’ (*pejtup*<sup>1</sup> ‘you saw/looked’)  
*pejti’āpu’yp*<sup>1</sup> ‘(you<sub>(pl)</sub> may) whistle!’ (*e’iti’āpu’yp*<sup>1</sup> ‘y. whistled’)

It seems indisputable that the forms presented in (14) and (15) belong to a ‘permissive’ (or imperative) mood category. That implies that we have a classification ‘mood’ of the set of verb forms, in addition to the three person- and one perspective classifications introduced above (cf. (7) and (11)). One class in this classification is **Permissive**, which is distinct from the set of forms presented in the last section, which simply are forms of the **Indicative** mood.

In later sections we will identify further mood categories. For the time being, the mood classification appears as in (16). Instead of complete sample verb forms we show only some relevant prefixes below the respective classes to which forms with these prefixes belong.



Note the overlap between Indicative and Permissive due to the syncretism of forms of transitive verbs with the prefix *pej-*. Note also that both of these categories contain forms with the prefix *i-*, too. But (except for stative verbs) this is NOT a case of syncretism because Indicative forms with *i-* are (Object-centred) forms in paradigms of *transitive* verbs, while Permissive forms with *i-* are forms of *intransitive* verbs only. Syncretism, however, implies that the SAME form is categorised more than once in the SAME paradigm.

A single form with *i-* can be ambiguous (and belong to both categories) if a transitive and an intransitive verb have the same stem, as in the case of *itan*<sup>1</sup> which can mean ‘(you<sub>(sg)</sub> may) run!’ as well as ‘2/3 painted me’. This is nevertheless not a case of syncretism, because the differ-

ent form-categorization pairs with the same form (first component) belong to different paradigms:  $\langle itan^1, \{-1, +2, -3, S.ct, Perm, \dots\} \rangle$  belongs to the paradigm of the active intransitive verb *tantu<sup>W</sup>* ‘to run’ and  $\langle itan^1, \{+1, -2, -3, O.ct, Ind, \dots\} \rangle$  belongs to that of the transitive verb *nãtantu<sup>W</sup>* ‘to paint something’.

With the prefixes that ‘mark’ the person and perspective categories as well as either Indicative or Permissive mood we have presented almost all verbal inflexional prefixes of Awetí. We may now turn to the suffixes.

#### 4.2 Aspect suffixes

There are only a few inflexional (form-building or non-derivational) verbal suffixes in Awetí. Two of them are *-(e)ju* and *-(z)oko*.<sup>21</sup> They always occur immediately after the verbal stem. The semantic effects and aspects of the allomorphy of *-(e)ju* are exemplified for active verbs in (17) and for stative verbs in (18).

- |      |                                    |   |  |
|------|------------------------------------|---|--|
| (17) | <i>o<b>tan</b>ju<sup>1</sup></i>   | ‘he is/was running’                         | ( <i>tantu<sup>W</sup></i> ‘to run’)       |
|      | <i>o<b>te</b>ju<sup>1</sup></i>    | ‘he is/was sleeping’                        | ( <i>tetu<sup>W</sup></i> ‘to sleep’)      |
|      | <i>e<b>tupe</b>ju<sup>1</sup></i>  | ‘you <sub>(sg)</sub> are/were looking (it)’ | ( <i>nãtupu<sup>W</sup></i> ‘to see/look’) |
|      | <i>a<b>tempe</b>ju<sup>1</sup></i> | ‘I am/was going out’                        | ( <i>tempu<sup>W</sup></i> ‘to go out’)    |
| (18) | <i>taku<b>w</b>eju<sup>1</sup></i> | ‘it is getting hot’                         | ( <i>takup<sup>W</sup></i> ‘(to be) hot’)  |
|      | <i>ijeme<b>j</b>ju<sup>1</sup></i> | ‘it is stinking / starting to stink’        | ( <i>ijem<sup>W</sup></i> ‘to stink’)      |

As can be seen in the examples, forms with *-(e)ju* usually present an event as in progress, similar to the progressive in English. In the case of stative verbs, a second meaning effect can be observed, where the state is being presented as incipient or being achieved. The exact conditions for one or the other meaning are still to be determined. It seems that *ijemeju<sup>1</sup>* can in isolation mean both ‘it is stinking’ and ‘it is beginning to stink’.

It is in any case clear that both meaning effects are aspectual, so the forms with *-(e)ju* belong to an aspect category that I label **Progressive**, as this seems to be the most general meaning effect (the inceptive meaning is more restricted).<sup>22</sup>

<sup>21</sup> Note again that we neglect in this paper the ontological distinctions in morphology. Parentheses in names of affixes (morphemes such as *-(e)ju<sup>L</sup>*) indicate that different allomorphs exist (e.g., *-ju* and *-eju*).

<sup>22</sup> For the names of aspect categories, we generally used Comrie (1976) for orientation.

A second suffix, *-(z)oko*, is mutually exclusive with *-(e)ju*. There are systematically two different possible meaning effects connected with forms that have *-(z)oko*, cf. (19) for the *-(z)oko*-forms of the same active verbs as above in (17). There are almost no occurrences of forms of stative verbs with *-(z)oko*. The few cases I have been able to elicit show the first of the semantic effects above. This effect can be labelled as ‘durative’ or ‘habitual’ or simply ‘imperfective’. An example is given in (20).<sup>23</sup>

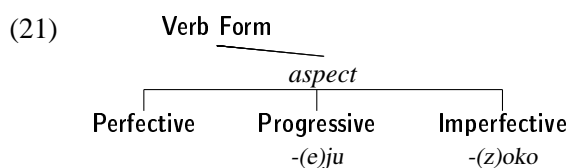
- (19) *otan-oko*<sup>1</sup> ‘he always ran’ / ‘he is about to run’  
*otezoko*<sup>1</sup> ‘he sleeps always’ / ‘he is going to sleep’  
*etuw-oko*<sup>1</sup> ‘you<sub>(sg)</sub> look always’ / ‘you are about to see it’  
*atem-oko*<sup>1</sup> ‘I used to leave’ / ‘I am about to leave’

- (20) *iti’apu’ywezoko*<sup>1</sup> ‘he always whistles’ (*ti’āpu’yp*<sup>w</sup> ‘to whistle’)

In sum, forms with *-(z)oko* should be analysed as belonging to another aspect category. The two meaning effects associated with this category can be labeled as ‘imperfective’ and ‘inchoative’. Of these we again choose the more general one as a label for the identified category and call it **Imperfective**.

Obviously, both categories, **Prog** and **Imp**, belong to the same classification, ‘aspect’, in the SUO of Awetí idiolect systems. They contrast with verb forms that have neither *-(z)oko* nor *-(e)ju* and that belong to a third aspect category. As indicated by examples above, unmarked forms usually have a perfective aspectual reading (hence the most natural translation in past tense). This and the secondary effect of Imperfective forms similar to an immediate future may give the misleading impression that these categories are tenses.

Aspect in Awetí can be summarised as the classification in (21).



However, he labels ‘inceptive’ as ‘ingressive’.

<sup>23</sup> For reasons of space, we cannot explain the rules for different allophones and allomorphs here. If it was an active verb, the form would be *\*oti’apu’ywoko*<sup>1</sup>.

4.3 *Mood suffixes*

There are two more suffixes that can occur in verb forms:  $-(t)u$ <sup>24</sup> and  $-aw$ . If of one of the aspect suffixes discussed above is present,  $-(t)u$  or  $-aw$  occurs after it. However, whether they are inflexional or derivational is debatable.

In Awetí, substantives in general have the same person prefixes (marking the ‘possessor’) as stative verbs (or object-centred forms of transitive verbs). Only in its male variety do they differ, in the Third Person.<sup>25</sup> Interestingly, the forms with  $-(t)u$  or  $-aw$  always have the same prefixes as substantives. This makes one suspect that they are nominalizations.

Compare the forms of substantives in (22) with those with  $-(t)u$  in (23), which belong to the active intransitive verbs  $totu^w$  ‘to go’ (it’s stem **to** is consonant-initial) and  $tagetu^w$  (vowel-initial stem **age**) ‘to shout’.<sup>26</sup>

(22)	<i>ity</i> <sup>1</sup> ‘my mother’ female: <i>ĩty</i> <sup>1</sup> ‘his mother’ male: <i>nãty</i> <sup>1</sup> ‘his mother’		<i>itok</i> <sup>1</sup> ‘my house’ female: <i>tok</i> <sup>1</sup> ‘his house’ male: <i>nok</i> <sup>1</sup> ‘his house’
------	--	--	---

(23)	<i>itotu</i> <sup>1</sup> ‘(that) I go’ / (‘my going’?) female: <i>ĩtotu</i> <sup>1</sup> ‘(that) he goes’ / (‘his going’?) male: <i>nãtotu</i> <sup>1</sup> ‘(that) he goes’ / (‘his going’?)  <i>itagetu</i> <sup>1</sup> ‘(that) I shout’ / (‘my shouting’?) female: <i>tagetu</i> <sup>1</sup> ‘(that) he shouts’ / (‘his shouting’?) male: <i>nagetu</i> <sup>1</sup> ‘(that) he shouts’ / (‘his shouting’?)
------	---

In the case of transitive verbs, the bare person prefixes in the  $-(t)u$ -forms can refer only to the semantic object, not the subject, even if the latter would be higher on the hierarchy of reference.<sup>27</sup> In order to explicitly

<sup>24</sup> Allomorphic variants:  $-tu$  (after vowels and  $n, j$ ),  $-pu$  (after  $m$ ),  $-ku$  (after  $ng$ ) and  $-u$  (after  $p, t, k$ ).

<sup>25</sup> The different prefixes for Third Person are one of the major features that differentiate the two genderlects of Awetí, i.e. the variety used by men and that used by women. See Drude (2002) for details.

<sup>26</sup> Only the male pronoun is given instead of ‘his/her’ or ‘(s)he’. No stative verb is used in these examples. A suffix  $-tu$  (with an allomorph  $-ytu$ ) exists with stems of stative verbs. But this clearly serves to derive nouns from the verb, with the meaning: ‘what/who has (the property named by the stative verb)’. For reasons of space, stative verbs will generally not be treated in the remaining sections of this paper. Therefore ‘intransitive’ will usually refer to ‘active-intransitive’.

<sup>27</sup> This ‘absolute’ alignment of the subject of intransitive verbs with the object of transi-

‘mark’ the subject by a person prefix, an extra ‘antipassive’ prefix *po(r)-* has to be added before the stem.<sup>28</sup> Compare the forms in (24).

- (24) *itup-u*<sup>1</sup> ‘(that) ... see(s) me / (‘my being seen’?)  
*ipo-tup-u*<sup>1</sup> ‘(that) I see ...’ / (‘my seeing’?)  
 female: (<sup>ʔ</sup>*i*)*ipo-tup-u*<sup>1</sup> ‘(that) he sees ...’ / (‘his seeing’?)  
 male: *nãpo-tup-u*<sup>1</sup> ‘(that) he sees ...’ / (‘his seeing’?)

The difference between the male and the female variety in this context is too complex for the limited scope of this paper, which will be restricted to the male variety.<sup>29</sup>

The bare *-(t)u*-forms without any of the person prefixes can also occur if they are immediately preceded by a nominal; see (25) for an example. This again is similar to a possessed substantive in a ‘genitive’ construction in juxtaposition after a possessor.<sup>30</sup>

Nevertheless, a number of properties suggest an analysis as verb forms and hence, their inclusion in the verbal paradigms:

- The *-(t)u*-forms are not combined with most affixes that occur with nouns.
- As said above, *-(t)u* can occur after one of the aspect suffixes *-(e)ju* and *-(z)oko*, and this is indeed a very frequent combination. It is at least unusual that a derivation take a stem together with an inflexional affix as its basis.<sup>31</sup>
- The *-(t)u*-forms rarely occur in typical nominal functions, e.g., as the complements of postpositions or as the complements of verbs (see the exceptions below). In particular, they do not occur as the subject, nor have they been found in equational sentences.
- Rather, the forms with *-(t)u* almost always function as a predicate.

---

tive verbs in *-(t)u*-Forms may be taken as a further evidence for ergativity in Awetí.

<sup>28</sup> For a different view on *po(r)-*, see Monserrat (2002).

<sup>29</sup> The male variety has more different forms. The facts of the female variety differ, with possible consequences for the paradigm bases and even the treatment of *-(t)u*-forms as deverbal nouns or verbal mood.

<sup>30</sup> Verb forms in these constructions can be classified as non-finite. They will not be taken into account here.

<sup>31</sup> One must admit, however, that this can also be observed in the case of two other affixes, *-at* and *-ap*, which should clearly be analysed as nominalisers of agent and circumstance/instrument, respectively.

Especially due to the last property we see these forms as verbal. The option to analyse the *-(t)u*-forms as nominalizations exists but will not be pursued here any further. It may turn out that an analysis as nominal predicates is also viable (especially semantically) and might even be preferable. At any rate, the inclusion of the *-(t)u*-forms increases the complexity of the verbal paradigms that are of interest here.

It is difficult to ascribe a meaning to *-(t)u*. Two sentences that differ only in the form of the verb (with vs. without *-(t)u*) have propositions that are identical or at least completely equivalent. If there is a difference in meaning, it is rather subtle and belongs to pragmatics. Possibly, use of *-(t)u*-forms makes the situation expressed by the verb less salient in comparison to other components of the sentence. One indication for this may be the fact that *-(t)u*-forms tend to occur together with adverbial expressions, in particular when these are at the beginning of the sentence or even in ‘cleft’-constructions.<sup>32</sup>

It is a very frequent stylistic figure, especially in narratives, to present the same situation at least twice, changing only the predicate using forms with and without *-(t)u*. The next example, (25), is a sequence of slightly simplified sentences from a historical narrative. The predicates, *-(t)u*-forms or not, stand out in bold face. (The glosses below the words are to be helpful but have no theoretical status.)

(25)	<i>kara'iwa</i>	<b><i>pokỹjokotu</i></b>	<i>nanype tsã</i>		<i>tsã</i>	<b><i>tokỹjtu</i></b>	<i>nanype</i>			
	the_white	kept_killing	there them		they	fought	there			
	‘the white kept killing them there’				‘they fought there’					
	<b><i>otokỹj</i></b>	<i>nanype tsã</i>		<i>nãtezak-ti</i>	<i>tsãn</i>	<b><i>utu</i></b>		<i>ajkulula</i>	<i>pe</i>	<b><i>out</i></b>
	3_fought	there they		afraid_of_it	they	came		Ajkulula	to	3_came
	‘they fought there’			‘afraid of this they came’				‘to Aj. they came’		

There are three occurrences of the transitive verb *tokỹjtu*<sup>W</sup> ‘to kill’, which can also mean ‘to fight’, especially the reciprocal forms with *to-*. Next we have two occurrences of the (active) intransitive verb *tutu*<sup>W</sup> ‘to come’. Already the first form, *pokỹjokotu*, illustrates several features discussed above, for instance juxtaposition of a person-prefix-less *-(t)u*-form

<sup>32</sup> In Tupí-Guaranian languages, forms with a suffix *-i* (several variants after vowels) that are analogs to the *-(t)u*-forms are reported to be obligatory in this position. The forms are usually seen as verbal and are called ‘indicative II’, ‘circumstantial’ or ‘oblique-topicalised’. For an overview, see Jensen (1998, § 6.1.)

after a nominal. *Kara'iwa* is here the subject (not object) due to the prefix *po-*. Also, this form combines the suffixes *-(z)oko* and *-(tu)*.

The *-(tu)*-forms also (rarely) occur as the object in certain sentences, if forms of certain verbs function as the (main) predicate. This holds at least for *nākwawapu*<sup>W</sup> 'to know' and *nākwakupu*<sup>W</sup> 'to want', as main verbs, as shown in (26).<sup>33</sup>

- (26) *wejkwawap mimō etotu* 'h. knows that y. went yesterday (*mimō*)'  
*ekwakup atit itotu* 'you want that I<sup>34</sup> go'

In sum, *-(tu)*-forms occur most often in main clauses as the predicate and may also occur as the predicates of subordinate clauses.<sup>35</sup> Both uses indicate that these forms belong to a modal category.<sup>36</sup> We label this mood category (the set of forms with *-(tu)*) **Subjunctive**.<sup>37</sup> This category contrasts with the **Indicative** mood (verb forms without *-(tu)*).

A similar situation holds for forms with the suffix *-aw*. These forms show the same resemblances to nouns as the Subjunctive forms but they never function as referential expressions. Instead, they only cooccur with another verb, often a verb of motion. Both verb form occurrences always share the same subject. See the examples in (27), with the verbs of motion *totu*<sup>W</sup> 'to go', *tutu*<sup>W</sup> 'to come' and the verbs *tatuku*<sup>W</sup> 'to (take a) bath' and *tantu*<sup>W</sup> 'to run'. The stems again stand out in bold face.

As indicated by the translations, the *-aw*-forms often semantically serve to indicate purpose or manner of the event expressed by the main verb. A closer look shows that the semantic relations are more complicated. Sometimes, the two verb forms seem to express two aspects or components

<sup>33</sup> This use of the *-(tu)*-forms is not necessarily in contradiction with the nominalization analysis, the equivalent meaning could be imitated in English as 'he knows your going' and 'you want **my** going'.

<sup>34</sup> *I* receives emphasis, e.g. with contrastive meaning, due to the use of the pronoun form *atit*, which functions as subject in the subordinate clause.

<sup>35</sup> As said before (footnote 5), the suffix *-(tu)* also serves to form the citation form used only in meta-linguistic contexts. As this is not related to the T-A-M system, we will not discuss these forms any further.

<sup>36</sup> Negative forms (see below sec. 5.1) appear only in subordinate clauses. If there were *-(tu)*-forms that appear only in main clauses, this might justify even two different categories.

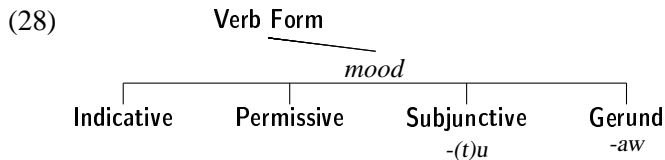
<sup>37</sup> The function of building subordinate predicates is the most distinctive one. The term 'circumstantial' has been used, e.g., by Seki (2000) for forms in Kamayurá that are functionally similar to the *-(tu)*-forms in main clauses. Again, see Jensen (1998, § 6.1).



- (27) a. *ato itatukaw*                      b. *e'itupaw ozout*  
       1\_go 1\_bath                            23\_see 13\_come  
       'I went to take my bath'            'we<sub>excl</sub> came to see you<sub>pl</sub>'
- c. *ajut itantaw*                        d. *ajatuk itoaw*  
       1\_come 1\_run                            1\_bath 1\_go  
       'I came running?'                    'I went to take my bath'

of a single action or event, as in example d. in (27), where it is the motion verb that carries the suffix *-aw*. The propositions of (27.a) and (27.d) seem to be almost the same. This makes the construction similar to serial verbs.

We analyse these forms again as constituting a mood category which we call, again following Tupinist tradition, **Gerund** (which is an unusual name for a mood). The two modal categories **Subj** and **Gerund** contrast with the modal categories identified before, **Ind** and **Pms** (without the suffixes *-(t)u* or *-aw* and using different sets of person prefixes). So it seems possible to conceive all of these categories as belonging to one and the same classification, mood, cf. (16), above. The extended classification can be represented as in (28) (the distinctive suffixes are indicated below).



These are almost all the T-A-M categories that contain basic synthetic forms, that is, those which are 'marked' by affixes only. There are other categories in the T-A-M system of Awetí and that involve the use of particles. These are presented in the following section.

## 5 Tense-Aspect-Mood Auxiliary Particles

### 5.1 A particle for the negated permissive

In section 4.1, we described the Permissive mood. In the case of negated permission, the forms presented above are not used. Instead, the particle *kware* occurs, usually sentence-initially, together with a verb form in the Subjunctive mood, if the verb is intransitive, or in the Indicative, if the verb is transitive. See the examples in (29), again with the verbs *totu<sup>w</sup>* 'to go'

and *nākỹjtu*<sup>w</sup> ‘to kill’.

- (29) *kware etotu* ‘don’t you<sub>sg</sub> go!’ (cf. *ito*<sup>1</sup> ‘(you<sub>sg</sub> may) go!’)  
*kware ekỹj mōj* ‘don’t you<sub>sg</sub> kill the snake (*mōj*)!’  
 (cf. *jokỹj mōj* ‘(you<sub>sg</sub> may) kill the snake!’)

Obviously, the particle *kware* causes simultaneously both negation and the semantic effect that is usually connected with the Permissive mood. In order to decide whether to include it in the verb paradigms, we should take a quick look at negation in Awetí in general. As the negated clauses in (30) show, in addition to the negator *an*, the verb forms in the Indicative are also marked as negated by the suffix *-(y)ka*. In the Subjunctive, the suffix *-e’ympu* is used, which probably originated from a combination of the suffix for nominal negation, *-e’ym*, and *-(t)u*. But note that the order is ‘wrong’ if these forms were indeed nominalizations, which is an argument for considering *-e’ympu* synchronically as one suffix, and reinforces the analysis of the *-(t)u*-forms as a verbal mood.

- (30) *an mimō otoka* ‘he/they didn’t go yesterday (*mimō*)’  
*an ekỹjyka mōj?* ‘didn’t you<sub>sg</sub> kill the snake (*mōj*)?’  
*akwawap etoe’ympu* ‘I know that you<sub>sg</sub> did not go’ (cf. (26))

So we have a distinction between Positive and Negative verb forms in Awetí (independently from the expressions with *kware*), and we also have a distinction between the Permissive and the Indicative and other moods. As it happens, there is no synthetic or one-word form that is a Negative AND a Permissive form. Instead, as exemplified in (29), there are analytical (in other terminologies: ‘periphrastic’) forms that combine the particle *kware* with a verb form which by itself would be a Positive form. Note that *kware* can be combined only with verb forms that belong to the Subjunctive, if the verb is intransitive, or to the Indicative, if it is transitive. However, the analytical forms as a whole, including *kware*, don’t belong to any of these categories, Subjunctive, Indicative or Positive.

In the overall presentation of the verbal functional part of the SUO of Awetí in section 6.1 (figure 1) we include the polarity classification introduced here, although it may not be a classification belonging to the T-A-M system *strictu sensu*. However, polarity closely interacts with mood in Awetí: the Negative category can be marked either by means of the suffix *-(y)ka* on the (finite part of the) verb form (only in the Indicative), or by

the suffix *-e'ympu*, which also marks the Subjunctive, or by the particle *kware<sup>w</sup>*, which also marks the Permissive. The latter categories belong to the mood classification shown in (28).

In approaches where paradigms belong to morphology only, analytical forms usually are an aberrant or at least a special or complementary case. As explained in section 2, in the view taken here there is no reason why we should not include analytical forms in verbal paradigms, since in our ontology all word forms are sequences of phonological words. We started by examining the core of synthetic forms (unit sequences, with only one member).

Applying what has been discussed so far, some elements of *tupu<sup>P</sup>*, the paradigm of the verb *tupu<sup>w</sup>* ‘to stay’ are shown in (31).

- (31) *tupu<sup>P</sup>*  $\ni$  (has among its elements:)
- $\langle \textit{ajupeju}^1, \quad \{+1, -2, -3, \text{S.ct, Ind, Pos, Prog, \dots}\} \rangle$
  - $\langle \textit{itupu}^1, \quad \{+1, -2, -3, \text{S.ct, Subj, Pos, Pfv, \dots}\} \rangle$
  - $\langle \textit{euwyka}^1, \quad \{-1, +2, -3, \text{S.ct, Ind, Neg, Pfv, \dots}\} \rangle$
  - $\langle \textit{kware eupejutu}, \quad \{-1, +2, -3, \text{S.ct, Pms, Neg, Prog, \dots}\} \rangle$

The members of the analytical form *kware eupejutu* do not always occur side by side. For illustration, we show in (32) a syntactic unit (together with a graphical representation of its constituent structure) which contains a discontinuous occurrence of *kware eupejutu*.

- (32)
- |  |  |      |                                  |
|--|--|------|----------------------------------|
|  | VGr  |      |                                  |
|  | Vf   |      |                                  |
|  |  |      |                                  |
|  | kware  | kype | <b>eupejutu</b>                  |
|  | NEG.PMS  | here | you <sub>(sg)</sub> -are_staying |
|  | ‘don’t you <sub>(sg)</sub> keep staying here!’ |      |                                  |

According to the constituent structure in (32) there are two simple constituents,<sup>38</sup> a verb form (*kware eupejutu*) and a particle form (*kype*, a form of an adverb<sup>39</sup>). The horizontal line which joins the two parts of the verb form is crossed by the vertical line which connects the label ‘VGr’ (assigned to the full verb group *kware kype eupejutu*) with its subconstituent

<sup>38</sup> ‘Constituents’ are roughly ‘meaningful parts of a unit among which grammatical relations’ hold.

<sup>39</sup> The constituent category **Particle Form** applies to the forms of words of many parts of speech, in particular adverbs, adpositions etc.

*kype*. Discontinuous constituents imply crossing lines, but these do not hinder the clear set-theoretical interpretation of the tree graph. Note that, although parts of the verb form, neither *kware* nor *eupejutu* is by itself a constituent of the sentence.

We also give glosses in (32) for each of the single phonological words (not for single morphemes). Crucially, the glosses alone do not indicate much about the structure of the sentence and the grammatical relations which hold among the words (these can only in part be inferred from the constituent structure). Therefore, they are by no means sufficient as an analysis of the sentence. In section 7.2 we exemplify our conception of grammatical relations.

Note that the glosses are of a very heterogenous character: In the case of *kype* the gloss does indicate the (lexical) meaning of this word, the concept ‘here’. The lexical meaning of the word *tupu*<sup>W</sup> and therefore of its analytical form *kware eupejutu* is ‘to stay’. The lexical meaning of *kware*<sup>W</sup> on its own is the empty concept, *b*<sup>0</sup> (cf. section 2.1; *b*<sup>0</sup> is the lexical meaning of all auxiliary particles as well as of all affixes). ‘NEG’ and ‘PMS’ designate the functional categories **Negative** and **Permissive** (which hold for the form as a whole, as discussed above). These are indirectly associated with *kware* via the ‘system link’, as will be shown below in section 6.3, cf. especially (51 c). The gloss below *eupejutu* indicates the syntactic meaning that this form would have in isolation.

The role that *kware*<sup>40</sup> has in verb forms is analogous to that of auxiliary verbs in European languages, with the difference that *kware*<sup>W</sup> itself does not inflect. Therefore, *kware*<sup>W</sup> is not an auxiliary verb, but an auxiliary particle, cf. (3). (As it is typical for particle words, it has an ‘improper’ paradigm; cf. the remarks above (2).)

We discussed *kware*-forms in considerable detail in order to illustrate analytical forms and discontinuous constituents. Analytical verb forms with *kware* should be readily admitted by many as belonging to Awetí verb paradigms because they fill the ‘gap’ of negative permissive forms which do not exist as synthetic forms. In the next subsection we show other analytic forms with auxiliary particles of Awetí.

---

<sup>40</sup> It would be more precise to say “occurrences of forms of (the word) *kware*<sup>W</sup>”. But here and elsewhere in this article we use the informal formulations.

### 5.2 ‘Temporal’ and ‘modal’ particles: factuality

Almost all examples presented so far had translations in the present tense or in the past tense (some as immediate future, cf. (19)), but, as was explained in section 4.2, these are side-effects of the aspect categories involved. There is only one clearly marked tense in Awetí: Future. Future forms are marked not by an affix, but by an auxiliary particle, *tut*<sup>W</sup> (cf. (2) for its components) which always occurs in the ‘second position’<sup>41</sup> of a clause and can be combined with verb forms in the Indicative or Subjunctive mood. Consider the examples in (33), with the verb *totu*<sup>W</sup> ‘to go’. A more detailed analysis of a sentence with *tut*<sup>W</sup> similar to (33 c) will be given below in section 7.2, cf. (59).

- (33) a. *ato tut* ‘I will go’  
 b. *ko’jem tut kajtotu* ‘tomorrow we<sub>incl</sub> will go’  
 c. *an tut etoka ko’jem?* ‘will you not go tomorrow?’

When referring to future events, use of the particle *tut*<sup>W</sup> is obligatory in main and, for most speakers, also in subordinate clauses. This holds even when temporal adverbials such as *ko’jem*<sup>W</sup> ‘tomorrow’ in (33 b+c) already indicate the temporal location of the event: bare *\*ko’jem kajtotu*, without *tut*, is not grammatical. This is the main reason to analyse *tut*<sup>W</sup> as an auxiliary particle and not as, say, an adverb.<sup>42</sup>

For the time being we assume a new verb form classification, tense, with the categories **Future** (marked by *tut*<sup>W</sup>) and the (unmarked) category **Present Tense or Non-Future**, whose elements are used for reference to past, present, or permanent events / states / situations.

There are only very few possibilities to refer to future situations without using *tut*<sup>W</sup>. One is by using the particle *ari*<sup>W</sup> instead of *tut*<sup>W</sup>, as in (34).

<sup>41</sup> The rules that underly the word order restrictions at hand are complex. Informally, in the first position, one usually finds a complete constituent or part of an analytical form, as in (33 a). In the second (‘Wackernagel’s’) position one finds function particles of several classes, with a fixed order of occurrence.

<sup>42</sup> Note, however, that *tut*<sup>W</sup> also occurs in sentences with nominal predicates and in fragments without a verb form, such as short answers. The latter are probably cases of ellipsis; whereas in the former case we are forced either a) to assume a second particle *tut*<sub>2</sub><sup>W</sup> (with non-empty lexical meaning), or b) to admit the possibility of *tut*<sup>W</sup> having a second syntactic function, not as auxiliary, or c) to allow a Future tense category also for nominal paradigms. Similar considerations exist for other auxiliary particles.

- (34) a. *ato ari*                    ‘(don’t you worry,) I will go’  
 b. *ko’jem ari kajtotu*    ‘(don’t worry,) tomorrow we<sub>incl</sub> will go’

With both sentences, the speaker assures that the event of his/their going will happen (as with *tut<sup>w</sup>*). Additionally, the speaker signals that he knows that this event is wanted by the hearer, and tells the latter that he need not be worried about the fulfilment of his wish. This pragmatic component of the meaning effect of *ari<sup>w</sup>* is typical of a mood category.

However, this particle has the same syntactic behaviour as *tut<sup>w</sup>* (occurrence in the same ‘slot’ among the ‘second position’ particles). Also, they cannot be combined, so they seem to interact in a way that ‘markers’ for different categories in the same classification usually do. Furthermore, *ari<sup>w</sup>* can be combined with verb forms in the **Ind** (34 a) as well as in the **Subj** (34 b), so it is neutral as to the mood categories identified in section 4.3.<sup>43</sup>

We conclude that there is a category marked by means of the auxiliary particle *ari<sup>w</sup>*, and we can label the category **Reassured Future**. However, it is unclear whether this category is to be considered a tense or a mood, although it belongs to the same classification as the pure **Future** category marked by *tut<sup>w</sup>*.

To complicate things, *tut<sup>w</sup>* and *ari<sup>w</sup>* are only two of a series of related particles most of which have predominantly modal semantic effects. Another particle in this group is *tutepe<sup>w</sup>*, illustrated in (35).

- (35) a. *an tutepe atoka*            ‘I wouldn’t go / I wouldn’t have gone’  
 b. *mimō tutepe kajtotu*    ‘yesterday we<sub>incl</sub> would have gone’

In (35 a), the speaker considers the possibility of the non-occurrence of the event of his going (note the negation by *an* and the suffix *-ka*). This event is not located somewhere on the time-line, but its non-occurring is presented as counterfactual. In other words, it is implied that the speaker does, did or will go indeed. Only adverbs such as *mimō* ‘yesterday’ in (35 b) may specify the time of the hypothetical event; *tut<sup>w</sup>* and *ari<sup>w</sup>* cannot co-occur. As the examples in (35) show, the particle *tutepe<sup>w</sup>* can be combined with the **Ind** and the **Subj** (cf. section 4.3).

The evidence suggests that *tutepe<sup>w</sup>* belongs in one group with *tut<sup>w</sup>*

<sup>43</sup> On the other hand, *ari<sup>w</sup>* does not occur with Permissive forms, but then, neither does *tut<sup>w</sup>*. This should not come as a surprise, as it seems to be a universal tendency for Future tense forms.

and *ari*<sup>W</sup>. Therefore, it is best analysed as an auxiliary particle that marks *Irrrealis*, which belongs to the same classification as **Present Tense**, **Future** and **Reassured Future**. This suggests that the classification is of a relatively heterogeneous character.

We will return to this observation after having introduced the last of the series of ‘second-position particles’, *tepe*<sup>W</sup>. Consider the examples in (36), in which the verbs *tetu*<sup>W</sup> ‘to sleep’ and once more *totu*<sup>W</sup> ‘to go’ and the adverb *mimō* ‘yesterday’ occur.

- (36) a. *atet tepe* ‘I slept in vain’  
 b. *mimō tepe nātotu* ‘yesterday he/she went to no avail’

The sentence in (36 a) could be used, for example, when the speaker is indicating that he indeed slept but is still tired. Similarly, (36 b) indicates that somebody went (to some place), but that he or she did not achieve some purpose (his or her own, or a purpose of the speaker).<sup>44</sup> Generally, sentences with *tepe*<sup>W</sup> indicate that an event did indeed happen, but that it seems as if not, for the desired or expected results did not occur.

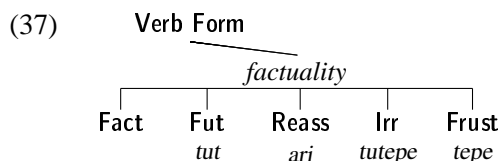
Again, the syntactic behaviour of *tepe*<sup>W</sup> is similar to that of the other particles presented in this section, with which it is mutually exclusive. (It seems probable that *tutepe*<sup>W</sup> even developed historically from a combination of *tut*<sup>W</sup>, which originally might have been a volitive particle, with the particle *tepe*<sup>W</sup>.) We conclude that *tepe*<sup>W</sup> is an auxiliary particle too, and that there is one more category in the same classification, to be called **Frustrative**. Semantically, **Frustrative** is, again, a category which has more a modal than a temporal character. So there is one tense category (**Fut**), two categories, (**Irr** and **Frust**) that are rather to be considered moods, and one category, **Reassured Future**, which is of a mixed character – and there is the unmarked category, **Present Tense (Non-Future)**.

This represents a problem because, generally, different classes in one classification are defined by criteria that share one single ‘point of view’. We propose that the common point of view here is not time reference but rather ‘factuality’, which can be subsumed to modality. The unmarked forms, then, belong to a category **Factual** rather than **Present Tense**, and

<sup>44</sup> This is different from indicating that a person did something without any specific purpose. In that case a different particle, *tene*<sup>W</sup> is used. In other Tupian languages the difference between these particles has been blurred in favor of only one particle, with one or both functions (cognate with Awetí *tene*<sup>W</sup> or with *tepe*<sup>W</sup>).

the Future categories could indeed be identified as ‘Not-Yet-Factual’<sup>45</sup> and ‘Reassured-To-Become-Factual’ (but for short, we will use the terms **Future** and **Reassured**). **Factual** also has time-related meaning components or at least implications and hence possibly could also be named ‘Non-Future’.

The resulting factuality classification is graphically identified in (37).



It turns out, then, that Awetí does not have a tense classification at all although it has a category **Future** (which is a tense in other languages).<sup>46</sup> The factuality classification is opposed to the classification in (28) which comprizes moods that are of a more grammatical than semantic nature—with exception of the **Permissive**.

The **Perm** category does not fit well in this system anyway. As the classifications ‘mood’ and ‘factuality’ are independent classifications that subdivide the same set (i.e., cross-classifications), each verb form must belong to (at least) one class of both classifications. All **Perm** forms would belong to the **Fact** category because they cannot be combined with any of the particles presented in this section—but semantically this is at best problematic. On the other hand, it would technically be possible to include **Perm** in the factuality classification instead of in mood. But this would cause problems of consistency with respect to the factuality classification. Also, in that case, the **Perm** forms would at the same time belong to the **Indicative**, which hardly would make any sense.

In sum, the **Perm** forms are not combinable with any of the affixes that mark (other) moods nor with any of the particles that mark factuality categories. **Perm** forms are also formally marked by quite different means,

<sup>45</sup> The definition of this term, which ultimately relates to sentence-semantic facts, may well turn out to be equivalent to that of ‘Future’.

<sup>46</sup> This is not impossible even if the Future category should turn out to be ‘future’ in just the same sense as in other languages (i.e., applying the same sentential-semantic definition). The double nature of Future, having both temporal and modal/factual characteristics, has been often noted, see Comrie (1985, sec. 2.3); see also Palmer (2001) for a comprehensive discussion of modal categories. — There are other Tupian languages that are said not to have proper verbal tense (cf. Jensen 1998).



such as special person prefixes or a sentence-initial auxiliary particle, as opposed to suffixes for the marked moods and second-position particles for the marked factuality categories.<sup>47</sup> All this suggests that the Permissive class does not belong to *either* of the two classifications. Rather, this class should be set apart from all other forms *before* these classifications are applied to them. We thus obtain a part of the SUO classification system with three classifications, where adapted versions of those given in (28) and (37) apply only to the set of all the other forms which do not belong to *Perm*, a class obtained by a superior classification ‘permission-mood’ (similar to that in (16)). This solution seems to account best for the morphological, syntactic and semantic facts in Awetí. It is therefore included in the complete presentation in the following section 6.1 (figure 1).

There are several other possible candidates for auxiliary particles in Awetí. For reasons of space, these cannot be presented and discussed in this paper. We will now summarize the results so far and present how the formal properties and functional categories interact in building Awetí verb paradigms.

## 6 The Verb Form System and System Link of Awetí

### 6.1 *The functional system*

In the preceding sections, we introduced step by step the functional classifications that are relevant for the verbal Tense-Aspect-Mood system in Awetí. In figure 1, we summarize the results, presenting a possibly almost complete view of the functional classification system for verb forms, a major part of the Syntactic Unit Ordering (SUO) of Awetí (see section 2.2).

Of the nine classifications in the system, seven classify the origin of the classification system itself, the set of (Finite) Verb Forms.<sup>48</sup> These are all cross-classifications, which implies that any finite verb form belongs simultaneously to (at least) one class in each classification: to one perspective

<sup>47</sup> Also semantically, the main effect of the Permissive is in the ‘illocutionary act’, not in the proposition or in the involved speakers attitudes or presuppositions, as is the case of most other moods. For the Integrative conception of these concepts and of sentential semantics in general, see Lieb (1983, part E).

<sup>48</sup> As said above, Non-Finite verb forms are neglected here for reasons of space. If they are included, still all classifications except the three person (participation) classifications apply to the set of (all) Verb Forms.

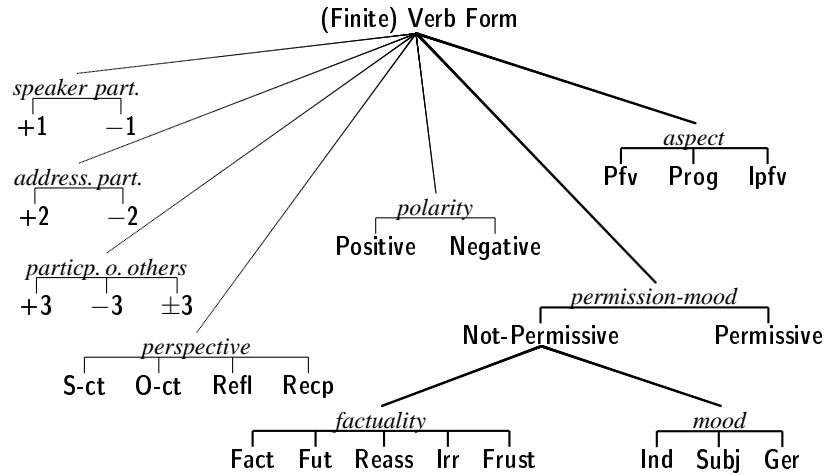


Figure 1: The functional verb form system of Awetí

category, to one aspect, one polarity, and so forth.

The Tense-Aspect-Mood subsystem in question here is marked by bold lines (the person subsystem, consisting of four classifications, is another major subsystem). It comprises the aspect classification and the three classifications of the modal(-temporal) branch, the only one that shows subclassifications. This means that any verb form *either* belongs to the **Permissive** or, if **Not-Permissive**, it belongs to both, a (grammatical) mood and a factuality class. The only temporal categories (**Future**, **Reassured** (**Future**), and arguably **Factual/Non-Future**) are located in the latter branch, but there is no homogeneous tense classification in Awetí.

A paradigm is structured with respect not only to functional but also to structural properties of its forms. The latter are treated in the next subsection.

## 6.2 The structural system

A given verb form that belongs to a certain functional category (in the system of functional classifications presented above) presents certain corresponding formal or structural properties, especially certain ‘markers’ that occur in its inner morphological and/or syntactic composition. In many cases it is only a combination of formal criteria that suffices to determine membership of a given forms in certain functional categories. Therefore

we need to separate the functional categories and the structural criteria and to establish the relationship between them in a separate step (see the next subsection).

The structural criteria also bring forth classifications of the set of verb forms. Jointly these classifications form a second subsystem of the SUO that has the set of verb forms as its basis: the ‘structural (or formal) verb form classification system’ of Awetí. The types of formal criteria that characterize Awetí verb forms have been implicitly presented above:

- The first (or person) prefix (if any) of the form, or more exactly, of the main part of the form, in the case of analytical forms.<sup>49</sup>
- The class of the stem, transitive, active-intransitive or stative-intr.
- The presence or absence of one second prefix (*te-*, *to-* or *po(r)-*). Note that these occur only with transitive verb stems.
- Occurrence or not of a first suffix, immediately after the stem (*-(e)ju* or *-(z)oko*).
- Presence or absence of a final suffix (in particular, *-(t)u*, *-aw*).
- Presence of a suffix *-ka* or *-e’ympu*, or of the particle *kware*<sup>w</sup>, or absence of these.
- Presence or absence of one of the ‘second-position-particles’ presented in section 5.2.

These criteria are straightforwardly translated into necessary and sufficient conditions that define the respective classes in the structural system. We need not show this here; for details, we refer the reader to Drude (2004, 2007). We present directly the resulting structural system, in figure 2.<sup>50</sup> Classifications or classes that are particularly significant for the TAM-system again stand out in bold lines or bold face.

<sup>49</sup> In Awetí, the main part of a verb form is easily identified: it is the smallest part which on its own also is a verb form. So, the main part of *ato tut* ‘I will go’ (cf. (33)) is *ato* ‘I go/went’ (cf. (5)). In other languages such as English, other criteria have to be used, because a form like *will go* is construed of two verb forms.

<sup>50</sup> The abbreviations and names of classes in figure 2 should be self-evident. The brackets basically serve to distinguish structural classes from functional ones, and to distinguish the classes connected with certain morphemes or particles from the morphemes or particles themselves. For instance, [*-(z)oko*] is the name of the class of verb forms defined by the occurrence of (an allomorph of) the suffix *-(z)oko*.

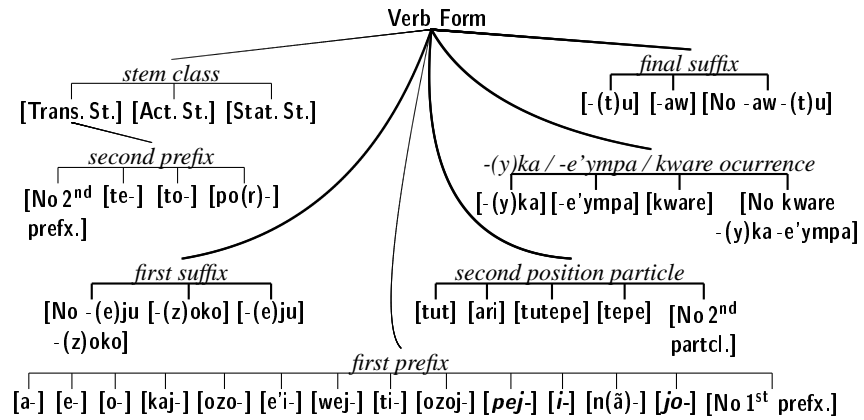


Figure 2: The structural verb form system of Awetí

### 6.3 The system link

Traditional paradigm tables relate the structural properties of forms to their relevant functional categories. In the framework used here, this relationship is captured by the so-called “*system link*” (SL, for short). In this section we present an almost complete but slightly simplified version of the SL for verb forms of Awetí.<sup>51</sup>

The system link formally is a relation between combinations (sets) of structural categories and combinations of functional ones. We need combinations of categories of each type rather than single categories because in many instances, a certain ‘marker’ (or, more general, structural property) alone is not sufficient for determining relevant functional categories. Often one can assign a form that has a certain structural feature (or combination of features) to several functional categories.

For instance, most person prefixes indicate all three person categories (participation or not of speaker, hearer, or others) simultaneously.<sup>52</sup> Furthermore, occurrence of *a-* and most prefixes of the subject-marking series

<sup>51</sup> The simplification again basically concerns the Non-Finite forms (without person prefix), and Subjunctive and Gerund forms with the prefix *o-*, which are not covered. Also, for reasons of presentation, there may still be some redundancy, e.g. in (41). The degree of technicality of this section is higher than that of the rest of this paper and certainly greater than it would be in a descriptive grammar, in order to introduce the framework and underlying concepts.

<sup>52</sup> Similarly person endings in many European languages indicate both person and number.

for transitive verbs also determine that it is a **Subj.-ctrd** form and belongs to the **Ind** mood. All this is represented by the elements of the SL in (38). (Forms with *e-* and with *pej-* will be treated separately.)

- (38) a. [a-] : +1, -2, -3, **Subj.-ctrd**, **Indicative**  
 b. [wej-] : -1, -2, +3, **Subj.-ctrd**, **Indicative**  
 c. [ti-] : +1, +2, ±3, **Subj.-ctrd**, **Indicative**  
 d. [ozoj-] : +1, -2, +3, **Subj.-ctrd**, **Indicative**

Each line in (38) can be translated into a sentence, a logical implication which is an empirical claim to be made in an (declarative, ideally axiomatic) grammar of the language. For instance, (38 a) may read: “If a form belongs to the structural category [a-], then it belongs to the functional categories +1, -2, -3, **Subj.-ctrd.**, and **Indicative.**” In (38), there is only one structural category (defined by the structural property of containing the named prefix) at the left of each colon. If there are several, the underlying structural properties form *together* a *sufficient* condition for assigning a form to the respective functional categories.<sup>53</sup> We claim that it is by reference to the system link that inflexional affixes are best described. For the general Integrational treatment of inflexional units, especially in lexicographical terms, see Drude (2004, 2006).

Most of the other person prefixes can occur, with different functions, in forms of intransitive and of transitive verbs in different moods. Therefore, neither the mood nor the perspective can be connected to the person prefix alone. But most person prefixes (including *pej-*) do determine unequivocally the person categories, as is shown in (39).

- |                           |                                  |
|---------------------------|----------------------------------|
| (39) a. [e-] : -1, +2, -3 | e. [pej-] : -1, +2, +3           |
| b. [kaj-] : +1, +2, ±3    | f. [o-] : -1, -2, +3             |
| c. [ozo-] : +1, -2, +3    | g. [n(ã)-] : -1, -2, +3          |
| d. [e'i-] : -1, +2, +3    | h. [t-] : -1, -2, +3, <b>Ind</b> |

In the male variety described here, the prefix *t-*, cf. the structural category in (39 h), occurs only in forms of stative verbs, cf. (13), and these are then always **Ind**.

<sup>53</sup> Formally, each line represents an ordered pair, an element of the SL. The first component of each pair is the set of structural categories left of the colon (with only one element each in (38)), the second is the set of functional categories to its right. So (38 a) represents:  $\langle \{[a-]\}, \{+1, -2, -3, \mathbf{Subj.-ctrd.}, \mathbf{Indicative}\} \rangle$ .

Most perspective categories are covered by the pairs in (40). All forms of intransitive verbs (stative and active) are **Subject-centred**. The same holds for the forms of transitive verbs which contain the prefix *po(r)-* (see above, section 3.3). The other two perspective categories, **Refl** and **Recp**, are uniquely connected with the prefixes *te-* and *to-*, respectively.

- |                    |               |          |              |
|--------------------|---------------|----------|--------------|
| (40) a. [Act. St.] | : Subj.-ctrd. | d. [te-] | : Reflexive  |
| b. [Stat. St.]     | : Subj.-ctrd. | e. [to-] | : Reciprocal |
| c. [po(r)-]        | : Subj.-ctrd. |          |              |

More criteria than the person prefix alone must be considered in order to determine to what mood and perspective forms of the structural categories in (39 a-d) belong. The same holds for forms with the prefix *i-*.

Negative forms are usually quite specific for mood (see (51), below). Usually, forms that belong to [**No kware-(y)ka-e'ympa**] (below abbreviated as [**No K-E-K**]) and without the suffixes *-(t)u* or *-aw* are **Indicative**, cf. (41). Forms of active intransitive verbs with the prefix *i-* cannot belong to the **Ind**. These facts make it necessary to refer not just to the absence of *-(t)u* or *-aw* but also to the single ‘first prefix’ structural categories (cf. figure 2). (Prefixes that occur *only* in Indicative Forms have been dealt with in (38).) Other elements for [**o-**], [**kaj-**], [**ozo-**] and [**e'i-**] instead of [**e-**] in (41 c) were omitted for reasons of space.

- |   |       |
|---|-------|
| (41) a. [i-], [ <b>No K-E-K</b> ], [ <b>No-aw-(t)u</b> ], [ <b>Trans. St.</b> ] | : Ind |
| b. [i-], [ <b>No K-E-K</b> ], [ <b>No-aw-(t)u</b> ], [ <b>Stat. St.</b> ]       | : Ind |
| c. [ <b>e-</b> ], [ <b>No K-E-K</b> ], [ <b>No-aw-(t)u</b> ]                    | : Ind |

In the case of transitive verbs only, Indicative forms with these prefixes are **Object-centred** if no other prefix (*te-* or *to-*) is present. Additionally, the forms with *e-* can also be **Subject-centred**. This is a first case of syncretism (cf. section 3.3 and 7.4), which is reflected by the fact that the same combination of structural categories is assigned to two different sets of functional categories in (42 b+c).<sup>54</sup> Again, other elements for [**kaj-**], [**ozo-**] and [**e'i-**] instead of [**i-**] in (42 a) were omitted.

Forms with *-(t)u* belong to the **Subj** mood. The perspective of forms with *po(r)-* (**Subj.-ctr.**) has already been determined in (40 c), but we must still account for the fact that forms of transitive verbs with *-(t)u* but without

<sup>54</sup> Note that [**No sec. prefix**] is a subset of [**Trans. St.**].

- (42) a. [i-], [No 2<sup>nd</sup> prefix], [No-aw-(t)u] : Obj.-ctrd.  
 b. [e-], [No 2<sup>nd</sup> prefix], [No-aw-(t)u] : Obj.-ctrd.  
 c. [e-], [No 2<sup>nd</sup> prefix], [No-aw-(t)u] : Subj.-ctrd.

*po(r)*- are always **Object-centred** ('ergativity', see sec. 3.3). This is accomplished in (43 b).<sup>55</sup> The Gerund forms are determinable by the occurrence of the suffix *-aw*, cf. (43 c). As the *-(t)u*-forms, the *-aw*-forms without a second prefix after the person prefix are also always **Object-centred**.<sup>56</sup>

- (43) a. [-(t)u] : Subj      |      c. [-(t)u], [No 2<sup>nd</sup> prefix] : O-ct.  
 b. [-aw] : Ger            |      d. [-aw], [No 2<sup>nd</sup> prefix] : O-ct.

Now we come to the prefixes that can mark **Permissive**. *jo-* has exclusively this function; this form is highly specific and poly-functional.

- (44) [jo-] : -1, +2, +3, Subj.-cntrd., Perm

A similar situation holds for *pej-* in the case of intransitive verbs. But *pej-* can also occur in forms of transitive verbs that belong to the **Ind** and to different factuality categories. Indeed, syncretism occurs only in the case of transitive verbs when there is no second position particle present. In this case *pej-* forms belong to the **Perm** as well as to the **Ind** and to the category **Factual** ( $\simeq$  **Non-Future**). The person categories of *pej-* forms have already been accounted for in (39 e). When *pej-* co-occurs with one of the factuality particles, the form is always in **Ind** mood (the factuality categories will be accounted for separately in (49), below). The relevant *pej-*-related elements are given in (45).<sup>57</sup>

- (45) a. [pej-], [Act. St.] : Perm  
 b. [pej-], [No 2<sup>nd</sup> prefix] : S-ct  
 c. [pej-], [No 2<sup>nd</sup> prefix] [No 2<sup>nd</sup> partcl] : Perm  
 d. [pej-], [No 2<sup>nd</sup> prefix] [No 2<sup>nd</sup> partcl] : Ind, Fact  
 e. [pej-], [tut] : Ind

<sup>55</sup> Note again that [No sec. prefix] is a subset of [Trans. St.]. All forms of intransitive verbs are determined to be Subject-centred; see (40 a+b).

<sup>56</sup> Apparently, *te-* and *to-* do not co-occur in one form with *-(t)u* or *-aw*.

<sup>57</sup> The other elements can be obtained by substituting [Stat. St.] for [Act. St.] in (45 a), and [ari], [tutepe], and [tepe] for [tut] in (45 e)

A complex case is also that of *i-*, which can mark the ‘first person singular’ (in stat. itr. verbs and in non-Indicative moods). It also can mark the Permissive ‘singular’ in act. itr. verbs. Syncretism is involved only in the case of stative verbs if no factuality-particle is present. For act. itr. verbs, *i*-forms are ‘first person singular’ prefixes only if one of the suffixes *-aw* or *-(t)u* is present, otherwise the form is **Perm**. Note that mood categories of most of the relevant combinations have been accounted for already, cf. (41 a+b), therefore they are not specified in (46).<sup>58</sup>

- (46) a. [*i-*], [**Stat. St.**], [**No-A-T**] : +1, -2, -3  
 b. [*i-*], [**Stat. St.**], [**No-A-T**], [**No 2<sup>nd</sup> ptc**] : -1, +2, -3, **Perm**  
 c. [*i-*], [**Act. St.**], [**No-A-T**] : -1, +2, -3, **Perm**  
 d. [*i-*], [**Act. St.**], [**-(t)u**] : +1, -2, -3

For transitive verbs, the prefix *i-* always marks the ‘first person singular’ (sometimes as subject, sometimes as object, depending of the presence of second prefixes and final verb suffixes, which has been accounted for above, see especially (42 a), above.). Nevertheless, the forms can sometimes belong to the **Perm**, when no factuality particles nor suffixes *-aw* or *-(t)u* are present, as the hierarchy of reference holds also for **Permissive** forms. For the same reason, forms with *ozo-* and possibly with *kaj-* (not included in (47)) can be **Perm**.

- (47) a. [*i-*], [**Trans. St.**] : +1, -2, -3  
 b. [*i-*], [**No 2<sup>nd</sup> prefix**], [**No 2<sup>nd</sup> ptc**], [**No -A-T**] : +1, -2, -3, **Perm**  
 c. [*ozo-*], [**No 2<sup>nd</sup> prefix**], [**No 2<sup>nd</sup> ptc**], [**No -A-T**] : **Perm**

With the elements of the SL given in (44) to (47) we have accounted for most conditions that involve the **Permissive** and the mood categories. The person and perspective categories have also completely been specified. The remaining elements are much simpler, involving mostly one structural category and one or two corresponding functional categories. This reflects the agglutinative character of the aspect system and the isolating character of the factuality system, as far as these are marked by prefixes and particles, respectively.

The easiest cases are the aspect categories in (48). Indeed, the classes in this classification (in the functional system) and those in the classification ‘first suffix’ (in the structural system) are extensionally identical.

<sup>58</sup> [**No -A-T**] stands for [**No -aw -(t)u**]; add a line as (46 d), but with [**-aw**] instead of [**-(t)u**].



- (48) a. [-(e)ju] : Prog | c. [No -(e)ju -(z)oko] : Pfv  
 b. [-(z)oko] : lpfv

Similarly, the forms containing second-position particles unambiguously mark factuality categories (hence ‘factuality particles’), as shown in (49). The difference from the aspect case is that not all verb forms without any of these particles belong to the ‘unmarked’ category (**Factual**), since some of them do (also) belong to **Perm**.

- (49) a. [tut] : Fut | c. [tutepe] : Irr  
 b. [ari] : Reass | d. [tepe] : Frust

However, almost all forms without any factuality particle do belong to **Fact**, except forms with the prefix *jo-* and some forms with the prefix *i-*. Therefore, we must refer to the prefixes one by one, following the scheme in (50) which represents eleven lines by substituting each of the categories [a-], [e-], [o-], [kaj-], [ozo-], [e’i-], [wej-], [ti-], [ozoj-], and [n(ã)-] for ‘X’ in (50 a). The forms belonging to [i-] have to be excluded from being **Perm** forms, except in act. itr. verbs.

- (50) a. X, [No 2<sup>nd</sup> ptc] : Fact  
 b. [i-], [No 2<sup>nd</sup> ptc], [Trans.St.] : Fact  
 c. [i-], [No 2<sup>nd</sup> ptc], [Stat.St.] : Fact

Finally, we must account for the effects of the classes of ‘-(y)ka / -e’ympa / kware occurrence’. These are unambiguous and involve polarity, and mood / permission, cf. (51).

- (51) a. [-(y)ka] : Negative, Ind  
 b. [-e’ympu] : Negative, Subj  
 c. [kware] : Negative, Perm  
 d. [No kware-(y)ka -e’ympa] : Positive

In the following, final section we discuss how single forms are assigned to their categorizations with which they classify as elements of a paradigm.

## 7 Elements of Awetí Verb Paradigms: Examples

### 7.1 Synthetic form *atupeju*<sup>1</sup>: Problems with traditional glossings

In the previous section, we presented three of the major components of the basis for verb paradigms in Awetí. The system link in particular is in the core of descriptions of forms and their properties in any language. Other components required for construing paradigms involve lexical meaning (there has to be an identical one for all forms of a paradigm) and formal relatedness of the stem (allowing, however, for suppletive forms).

Let us exemplify the interaction of these components, principally the functioning of the system link, in building elements of paradigms. Consider for instance the synthetic form *atupeju*<sup>1</sup> ‘I am/was looking’. It belongs to the structural categories in (52), cf. the structural system in figure 2 (it is irrelevant here that (52 a) implies [Trans. St.]):

- |   |  |  |
|---|--|--|
| (52) a. [No 2 <sup>nd</sup> prefix]<br>b. [-(e)ju]<br>c. [a-] |  | d. [No 2 <sup>nd</sup> ptc]<br>e. [No kware-(y)ka-e'ympa]<br>f. [No -aw -(t)u] |
|---|--|--|

Given the formal properties of a form, the system link (SL) determines to which functional categories the form belongs. For this, we consider all elements of the SL that have only the structural categories to which the form in question belongs. In (53), all relevant elements of the SL with categories in (52) are again listed.<sup>59</sup>

- |                                |   |                             |
|--------------------------------|---|-----------------------------|
| (53) (38 a) [a-]               | : | +1, -2, -3, Subj.-ctrd, Ind |
| (48 a) [-(e)ju]                | : | Prog                        |
| (50 a) [a-], [No sec. partcl.] | : | Factual                     |
| (51 d) [No kware-(y)ka-e'ympa] | : | Positive                    |

All the functional categories *atupeju*<sup>1</sup> belongs to (according to the SL) are from different classifications with no conflicts (see the functional system in figure (1) and section 7.3, below, for the notion of ‘conflict’). Therefore we can combine them all to one categorization of *atupeju*<sup>1</sup>. The combination of the form and its categorization qualifies as an element of the paradigm *nātupu*<sup>P</sup>, cf. (54), which we began to characterize in (12).

<sup>59</sup> It does no harm that (52 a and f) do not come into play; this is due to the highly specific prefix *a-* that determines several functional features of the form by itself.

(54)  $n\dot{a}tupu^P \ni \langle atupeju^1, \{+1, -2, -3, \text{S-ct, Ind, Fact, Prog, Pos}\} \rangle$

Note that in this case it is possible to assign most functional categories to one morph that occurs in the form (in particular, to the prefix *a-*). This makes it possible to propose glossings like that in (55).<sup>60</sup>

(55)  $a- \text{tup} -eju$  'I am/was looking'  
1.SG.SCT.IND look PROG

Although certainly useful to give a quick impression of the functional effects of each element, such glossings are problematic, at least in the Word-and-Paradigm view taken here: none of these functional categories (in other frameworks often: 'morphosyntactic features') is the or a 'meaning' of any morph or morpheme, not even is **Progressive** a meaning of *-eju*, although in this case the relation is very immediate (cf. (48 a), the only element in the system link relevant for *-(e)ju*). Also, neither **Factual** nor **Positive** can be related to any morpheme—unless one introduces 'empty morphemes' or something similar, a debatable step which is not necessary in this framework.

## 7.2 Analytical form **tut etoka**: analysis of discontinuous occurrence

Consider *tut etoka* (as in *an tut etoka* 'you will not go'), which is (different from *atupeju*<sup>1</sup>, a synthetic form) an analytical form again of the intransitive verb *totu*<sup>w</sup> 'to go', used already in (33 c). It belongs to the structural categories in (56).

(56) [Act. St.], [No-(e)ju-(z)oko], [e-], [tut], [-(y)ka], [No-aw-(t)u]

Applying the SL in the same way as above, *tut etoka* can be assigned to the functional categories in (57), where the relevant element of the SL is indicated for better orientation.

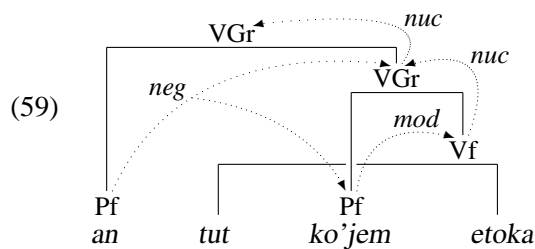
(57)	-1, +2, -3	for	[e-]	cf. (39 a)
	Subj.-ctrd	for	[Act. St.]	cf. (40 a)
	Pfv	for	[No -(e)ju -(z)oko]	cf. (48 c)
	Fut	for	[tut]	cf. (49 a)
	Negative, Ind	for	[-(y)ka]	cf. (51 a)

<sup>60</sup> Here, the informal traditional category 1st Person Singular is used, 'SCT' abbreviates Subject-centred.

Again, all categories belong to different classifications and are combinable, so we obtain the corresponding categorization and the element of the paradigm *totu*<sup>P</sup> in (58).

(58) *totu*<sup>P</sup>  $\ni$   $\langle$  *tut etoka*, {−1, +2, −3, S-ct, Ind, Fut, Pvf, Neg }  $\rangle$

In the sentence *an tut ko'jem etoka* ‘tomorrow you will not go’ this verb form occurs as a discontinuous constituent. This sentence can be analysed as having the constituent structure presented in (59), where also the main syntactic relations that hold among the constituents are indicated by dotted bowed arrows.



The conventions should be easy to understand: for instance, the occurrence of the adverb *ko'jem*<sup>w</sup> functions as a **modifier** of the occurrence of the verb form *tut etoka*. This occurrence is in turn the **nucleus** of the inner VerbGroup *tut ko'jem etoka*. In the case of a qualifier relation like negation, the domain has to be separated from the scope of the negation for semantic reasons. Therefore the negator *an* **negates** *tut ko'jem etoka* (domain) with respect to *ko'jem*. (The scope can vary with different intonations.) Of course, applying other theories could result in different analyses.

But in any framework, it is difficult to gloss this sentence adequately using only interlinear morphemic translations as formalised by Lehmann (1982) or Comrie *et al.* (2004). Especially the categories **Subj.-ctd** and **Pfv** are problematic to address in glossing formats based on Item-and-Arrangement and Item-and-Process approaches (cf. footnote 2), and the constituency and grammatical relations are difficult to render. For the purposes of language documentation, we proposed ‘Advanced Glossing’ (Lieb and Drude 2001), a glossing format that is also compatible with Word-and-Paradigm approaches as the one used here. We believe Advanced Glossing can solve most of the abovementioned problems.

We continue to exemplify the construction of elements of paradigms (assignment of categorizations to forms), turning now to more complex cases.

### 7.3 Syncretism for *pejtup*<sup>1</sup> (synthetic form of a transitive verb)

One relevant phenomenon is syncretism. Consider the form *pejtup*<sup>1</sup>, meaning ‘you<sub>(pl)</sub> saw’ or ‘you<sub>(pl)</sub> may look’. We again start by identifying the structural categories. As shown in (60), besides the prefix *pej-*, the form has no overt marking at all.

- (60) [Trans. St.], [No 2<sup>nd</sup> prfx], [No-(e)ju-(z)oko] [pej-],  
 [No 2<sup>nd</sup> ptc], [No-(y)ka-e’ympa kware], [No-aw-(t)u]

According to the SL, *pejtup*<sup>1</sup> can belong to the functional categories in (61).

(61)	funct. catg.:	due to:	see:
a.	-1, +2, +3	[pej-]	(39 e)
b.	Subj.-ctr	[pej-], [No 2 <sup>nd</sup> prfx]	(45 b)
c.	Perm	[pej-], [No 2 <sup>nd</sup> prfx], [No 2 <sup>nd</sup> ptc]	(45 c)
d.	Ind, Fact	[pej-], [No 2 <sup>nd</sup> prfx], [No 2 <sup>nd</sup> ptc].	(45 d)
e.	Pfv	[No-(e)ju-(z)oko]	(48 c)
f.	Positive	[No-(y)ka-e’ympa kware]	(51 d)

The functional categories in (61 a,b,e,f) are all from different cross-classifications and combinable as before. But this does not hold for (61 c+d)—these categories are in conflict. Informally, classes are in conflict if they belong to the same classification or if one of them or both belong to branches ‘below’ classes that belong to the same classification. In this case, **Ind** and **Fact** are classes ‘below’ **Not-Permissive**, and this class belongs to the same classification (permission-mood) as **Perm**, therefore these classes are in conflict and may not be combined in one categorization.

Note, however, that **Ind** and **Fact** are combinable with one another and also with the categories in (61 a,b,e,f), so the combination of all these is allowed. The same holds for **Perm** alone which can be combined with the categories in (61 a,b,e,f). In such a case, *both* possible combinations qualify as categorizations of *pejtup*<sup>1</sup>, and both resulting pairs belong to the same paradigm if they share the same lexical meaning as in this case. So we can

add two more elements to the paradigm  $n\dot{a}tupu^P$ , cf. (54), two elements that have identical first components. Exactly this is syncretism.

- (62)  $n\dot{a}tupu^P \ni$   
 $\langle pejtup^1, \{-1, +2, +3, S\text{-ct}, \text{Ind}, \text{Fact}, \text{Pfv}, \text{Pos}\} \rangle,$   
 $\langle pejtup^1, \{-1, +2, +3, S\text{-ct}, \text{Pms}, \text{Pfv}, \text{Pos}\} \rangle$

Syncretism is no coincidence but reflects the fact that in (45 c+d), the same combination of structural categories is assigned to different functional categories. At least whenever this is the case there will be syncretism.

#### 7.4 Syncretism for *etup tepe* (analytical form of a transitive verb)

Consider the form *etup tepe* ‘he saw you in vain / you saw him in vain’. Mood(s) and aspect are uniquely determined by the lack of suffixes and the particle *tepe*. But the form is both **Subject-centred** and **Object-centred**, as is already indicated by the translations. This is correctly determined by the paradigm base. Observe again the structural categories to which *etup tepe* belongs, listed in (63), and the categories that *etup tepe* belongs to are listed in (64) as above.

- (63) [Trans. St.], [No 2<sup>nd</sup> prfx], [No -(e)ju -(z)oko], [e-], [tepe],  
 [No kware-(y)ka -e’ympa] (= [No K-E-K]), [No -aw -(t)u]

(64)	funct. catg.:	due to:	see:
a.	-1, +2, -3	[e-]	39 a
b.	Ind	[e-], [No K-E-K], [No -aw -(t)u]	(41 c)
c.	Obj.-ctrd	[e-], [No 2 <sup>nd</sup> prfx], [No -aw -(t)u]	(42 b)
d.	Subj.-ctrd	[e-], [No 2 <sup>nd</sup> prfx], [No -aw -(t)u]	(42 c)
e.	Pfv	[No -(e)ju -(z)oko]	(48 c)
f.	Frust	[tepe]	(49 d)
g.	Pos	[No kware-(y)ka -e’ympa]	(51 d)

The two categories in conflict are indeed **Obj.-ctrd** and **Subj.-ctrd**, in (64 c+d). Both of them can be combined with all other classes, and so can all of them with each other. The result is again syncretism: the association of two categorizations for the same form. The two syncretistic elements of paradigm  $n\dot{a}tupu^P$  are given in (65), continuing (54) and (62).

- (65)  $\text{totu}^P \ni \langle \text{etup tepe}, \{-1, +2, -3, \text{O-ct, Ind, Frust, Pvf, Pos}\} \rangle,$   
 $\langle \text{etup tepe}, \{-1, +2, -3, \text{S-ct, Ind, Frust, Pvf, Pos}\} \rangle$

### 7.5 Synthetic form *ito*: no specific marking at all

Finally we present the results for the form *ito*<sup>1</sup>, possibly the smallest form that can also be a complete sentence in Awetí (the word happens to also mean ‘I’, in the female genderlect). The structural categories for *ito*<sup>1</sup> are listed in (66), and the functional categories assigned to *ito*<sup>1</sup> by the SL in (67).

- (66) [Act. St.], [No-(e)ju-(z)oko], [i-], [No 2<sup>nd</sup> ptc],  
 [No kware-(y)ka-e’ympa], [No-aw-(t)u]

(67) funct. catg.:	due to:	see:
Subj.-ctrd	[Act. St.]	(40 a)
-1, +2, -3, Perm	[i-], [Act. St.], [No-aw-(t)u]	(46 c)
Pfv	[No-(e)ju-(z)oko]	(48 c)
Pos	[No kware-(y)ka-e’ympa]	(51 d)

Note that many categories are assigned to the form *ito*<sup>1</sup> not because of the *occurrence* of a certain morpheme but rather for the *absence* of markers of a certain class. True, the assignment of several categories involves the prefix *i-*, the only overt ‘marker’—but *i-* is never sufficient to determine the functional categories in question, but rather only if combined with other structural properties (such as the stem class and again the absence of certain other markers). The only other ‘substantial marking’ is that the stem (and thus, the word) belongs to the active-intransitive class. In other words, none of the functional categories can uniquely be assigned to any morph that occurs in *ito*<sup>1</sup> (unless we use numerous ‘zero morphemes’), and nevertheless this form has the categorization in (68).

- (68)  $\text{totu}^P \ni \langle \text{ito}^1, \{-1, +2, -3, \text{Subj.-ctrd, Pms, Pvf, Pos}\} \rangle$

### 7.6 Final remarks

The procedure can be repeated with all verb forms in order to obtain the respective elements of paradigms (more than one element for a form in the case of syncretism). Using the categorizations of the individual forms, it is easy to present a paradigm (as conceived in this framework) in traditional

tables. In the case of Awetí, the tables are at least six-dimensional. For transitive verbs, the variables are person (if we count the three person categories as one, as the six resulting combinations can be put side by side), perspective, polarity, factuality, mood, and aspect. For the reduced subparadigm of Permissive forms, mood and factuality do not apply.

In the case of transitive verbs, we can estimate a total of 3200 possible combinatorial positions (valid categorizations without conflicts) of which about 1400 are actually ‘filled in’ (these are combinations that do result by applying the system link) by ca. 1000 different forms. For active intransitive verbs, these figures are considerably smaller: about 750 positions, 500 of which are filled in by ca. 400 forms. The numbers for stative verbs are even lower.

These figures are by no means unusual; many inflecting languages have similar results, especially if all analytic forms are considered. Polysynthetic languages are even to be expected to have much more complex paradigms. It is self-evident that these forms do NOT have to be memorized one-by-one. Rather, the speaker basically needs command of the system link (some 80 elements in the case of Awetí verb paradigms, where some 22 affixes and 5 auxiliary particles occur). The SL is highly structured and contains many repetitions and analogous regularities. Arguably, to acquire the inflectional system of a language can easily be described as learning parts of (adding elements to) the system link. Many overgeneralizations in learner varieties correspond to paradigm bases where certain elements of the SL are missing.

We hope to have demonstrated how a Word-and-Paradigm approach can be fruitfully used to describe grammatical categories such as tense, mood and aspect categories in Awetí.

### References

- Budde, Monika. 2000. *Wortarten: Definition and Identifikation*. Doctoral thesis, Freie Universität Berlin.
- Cabral, Ana Suely Arruda Câmara and Aryon D. Rodrigues (eds.). 2002. *Línguas Indígenas Brasileiras: Fonologia, Gramática e História. Atas do I Encontro Internacional do Grupo de Trabalho sobre Línguas Indígenas da ANPOLL*. Belém: Editora UFPA.



- Comrie, Bernard. 1976. *Aspect: an introduction to the study of verbal aspect and related problems*. Cambridge; New York: Cambridge University Press.
- Comrie, Bernard. 1985. *Tense*. Cambridge; New York: Cambridge University Press.
- Comrie, Bernard, Martin Haspelmath and Balthasar Bickel. 2004. The leipzig glossing rules: Conventions for interlinear morpheme-by-morpheme glosses. Electronic document, available at: <http://www.eva.mpg.de/lingua/files/morpheme.html>.
- Dixon, Robert M. W. 1994. *Ergativity*. Cambridge; New York: Cambridge University Press.
- Dixon, Robert M. W. and Alexandra Y. Aikhenvald (eds.). 2002. *Word: a cross-linguistic typology*. Cambridge; New York: Cambridge University Press.
- Drude, Sebastian. 2002. Fala masulina e feminina em awetí. In: Cabral and Rodrigues (2002), 177–190.
- Drude, Sebastian. 2004. *Wörterbuchinterpretation: Integrative Lexikographie am Beispiel des Guaraní*. [Lexicographica Series Maior 120]. Tübingen: Niemeyer.
- Drude, Sebastian. 2006. Die Lexikographie der Personenpräfixe im Guaraní. *Kolloquium Geschichte und Aktualität der Guaraní-Philologie in den deutschsprachigen Ländern*. Kiel University 2000.
- Drude, Sebastian. 2007. Inflectional units and their effects: The case of the personal prefixes of Guaraní. In: *Studies in Integrational Linguistics I*, edited by Sackmann, Robin. Amsterdam, Philadelphia: John Benjamins.
- Hockett, Charles F. 1954. Two models of grammatical description. *Word* 10: 210–234.
- Jensen, Cheryl Joyce S. 1998. Comparative Tupí-Guaraní morphosyntax. *Handbook of Amazonian Languages*, edited by Derbyshire, Desmond C. and Geoffrey K. Pullum, volume 4, 487–618. Berlin etc.: Mouton de Gruyter.
- Klimov, G.A. 1997. *Tipologija jazykov aktivnogo stroja*. Moskau: Nauka.
- Lehmann, Christian. 1982. Directions for interlinear morphemic translations. *Folia Linguistica (Acta Societatis Linguisticae Europaeae)* XVI: 199–224.
- Lieb, Hans-Heinrich. 1983. *Integrational Linguistics. Vol I: General Outline*, [Current Issues in Linguistic Theory 17]. Amsterdam etc.: Benjamins.
- Lieb, Hans-Heinrich. 1992a. Integrational linguistics: Outline of a theory of language. In *Prospects for a New Structuralism*, edited by Lieb, Hans-Heinrich, Current Issues in Linguistic Theory 96, 127–182. Amsterdam; Philadelphia: Benjamins.
- Lieb, Hans-Heinrich. 1992b. Paradigma und Klassifikation: Explikation des Paradigmenbegriffs. *Zeitschrift für Sprachwissenschaft* 11: 3–46.
- Lieb, Hans-Heinrich. 1993. Integrational linguistics. In: *Syntax: Ein internationales Handbuch zeitgenössischer Forschung / An International Handbook of Contemporary Research*, edited by Jacobs, Joachim et al. [Handbücher zur Sprach- und Kommunikationswissenschaft 9.1], 430–468. Berlin.: de Gruyter.

- Lieb, Hans-Heinrich. 2005. Notions of paradigm in grammar. *Lexikologie / Lexicology*, edited by Lutzeier, Peter R. [Handbücher zur Sprach- und Kommunikationswissenschaft 21.2]. Berlin: de Gruyter.
- Lieb, Hans-Heinrich and Sebastian Drude. 2001. Advanced Glossing: A language documentation format. Working Paper (1st version), available at: <http://www.mpi.nl/DOBES/documents/Advanced-Glossing1.pdf>.
- Mithun, Marianne. 1981. Active/agent case marking and its motivation. *Linguistics* 67.3: 510–46.
- Monserrat, Ruth Maria Fonini. 2002. Linguas Tupí e ergatividade. In: Cabral and Rodrigues (2002), 191–202.
- Palmer, F. R. 2001. *Mood and modality*. Cambridge; New York: Cambridge University Press.
- Queixalós, Francisco (ed.). 2001. *Des noms et des verbs en Tupi-Guarani: État de la question*. [Lincom Studies in native American Linguistics 37]. Munich: Lincom.
- Seki, Lucy. 1990. Kamaiurá (Tupí-Guaraní) as an active-stative language. *Amazonian Linguistics. Studies in Lowland South American languages*, edited by Payne, Doris L., Texas Linguistic Series, 367–392. Austin: University of Texas Press.
- Seki, Lucy. 2000. *Gramática do Kamaiurá, Língua Tupi-Guarani do Alto Xingu*. Campinas: Editora da Unicamp.