

Coreferential marking in Tupí-Guaraní languages

Paper presented at the XIII International Conference of Historical Linguistics

Düsseldorf, Germany August 10-15,1997

Cheryl Jensen

Summer Institute of Linguistics, Belém, Brazil

1. Introduction

Tupí-Guaraní is one of the major language families in lowland South America. This family is very diverse geographically, with member language groups located all across Brazil, and reaching into French Guiana on the north, Paraguay and Argentina on the south, and Bolivia on the west. The family has also been highly documented, allowing a significant amount phonological (Lemle, 1971) and morphological (Jensen, 1989) reconstruction. Languages of this family have been tentatively divided by the foremost Tupinologist, Aryon Rodrigues (1984/1985), into 8 subgroups, based systematically on their phonological history, but considering other factors as well, as in Table 1:

1 (Guaraní)	Argentina, Bolivia, Brazil, Paraguay	Chiriguano cluster (Ava, Izoceño). Guayaki (Aché). Kaiwá. Mbyá. Nhandéva.
2	Bolivia	Guarayu. Sirionó.
3	Brazil (coastal)	Tupí (extinct). Tupinambá (extinct). Nheengatú. Kokáma.
4	Brazil (GO,MA,MT,PA)	Avá. Akwawa cluster (Tocantins Asuriní, Suruí of Tocantins, Parakanã). Tapirapé. Tenetehára cluster (Guajajara, Tembé).
5	Brazil (MT,PA)	Xingu Asuriní. Kayabí. Araweté??
6	Brazil (AC,AM, MT,RO)	Apiaká. Kawahíb cluster (Parintintín, Tenharim, and others)
7	Brazil (MT)	Kamaiurá.
8	Brazil (AP, MA, PA), French Guiana	Anambé. Emerillon. Guajá. Urubu- Kaapor. Wayampi. Zo'é.

Table 1: Subgroups of Tupí-Guaraní family

2. Typological features

Typologically this language family is characterized by head-marking with no dependent-marking to distinguish subjects from objects. The word order in independent clauses is flexible, the most likely basic order being *SOV or *SVO (Jensen, 1998). In independent clauses there is a split-S cross-referencing system in intransitive verbs as well as a split-ergative system in transitive verbs. The latter is governed by a person hierarchy (1>2>3) which indicates whether A or O will be cross referenced. A nominative set (set 1) of prefixes which refers to Sa and A is limited to independent verbs: **a-* '1SG', **oro-* '1EX', **ja-* '1IN', **ere-* '2SG', **pe-* '2PL', **o-* '3'. This set occurs with A only when O is third person, and the A prefix is followed by the third-person object prefix from set 2, as in **a-i-potár* 'I like him'. Person markers from set 2 are used with both independent and dependent (subordinate and serial) verbs: **čé* '1SG', **oré* '1EX', **jané* '1EX', **né* '2SG', **pé* '2PL', **i-* '3'. In the split-S system they refer to So, in the split-ergative to O. In dependent verbs there is an ergative-absolutive cross-referencing system in which set 2 person markers refer to S or O. The same set of person markers is used to refer to a pronominal genitive on nouns and the object of postpositions. A set of coreferential prefixes (set 3) occurs in basically the same syntactic contexts as the person markers of this set (2), except for independent verbs.

3. Coreferential marking in a nutshell

In English, coreferential marking is only explicit and obligatory when a pronominal object, either of the verb (ex. 1) or of a preposition (ex. 2), is identical with the subject.

(1) He killed him. > He killed *himself*.

(2) He bought the car for him. > He bought the car for *himself*.

In this case a special set of coreferential pronouns replaces the accusative pronouns. A complete paradigm exists: *myself*, *ourselves*, *yourself*, *yourselves*, *himself*, *herself*, *itself*, *themselves*.

When a pronominal genitive is identical with the subject, no special form exists, making the referent ambiguous. This is clarified by the optional addition of the word *own*:

(3) He returned to his house. > He returned to *his (own)* house.

The word *own* may occur with any possessive pronoun: *my (own)*, *our (own)*, *your (own)*, *his (own)*, *her (own)*, *its (own)*, *their (own)*.

In Portuguese, special forms exist only for third-person pronouns, *se* (object of verb) and *si* (argument of the preposition), as in examples 4 and 5, respectively, and are unaltered

by gender or number. The existence of forms only for third person is not unusual since only third person presents any potential ambiguity.

(4) Ele o matou.> Ele *se* matou.
 he him killed > he himself killed

(5) Ele comprou o carro para ele. > Ele comprou o carro para *si* (*mesmo*).
he bought the car for him > *he bought the car for himself*

However, to emphasize the coreferential relationship in prepositional phrases, regardless of the person or number, it is possible to add the word *mesmo* 'specifically' after the pronoun.

In both languages the trigger for coreferential marking is the subject, and the targets, i.e. the potential recipients of such marking, are pronouns (Wiesemann, 1986:442).

In Tupí-Guaraní languages there are two type of morphemes for indicating coreferential marking: a reflexive morpheme **je-* (and its reciprocal counterpart **jo-*) and a complete paradigm (in various languages and in the protolanguage) of what I refer to as coreferential prefixes: **wi-* '1SG', **oro-* '1EX', **jere-* '1IN', **e-* '2SG', **peje-* '2PL', and **o-* '3'.

When the direct object is identical with the subject, the reflexive prefix **je-* is combined with the transitive verb stem, resulting in an intransitive verb:

(7) **o-i-potár* > **o-je-potár*
 3A-3O-like > 3A-REFL-like
 'he likes him/her/it' 'he likes himself'

This prefix occurs with all subject person markers: **a-je-potár* '1SG', **oro-je-potár* '1EX', **ja-je-potár* '1IN', **ere-je-potár* '2SG', **pe-je-potár* '2PL', **o-je-potár* '3'.

The coreferential object of a postposition is formed by the combination of one of the set of coreferential prefixes, together with the reflexive prefix, and the postposition.

Compare the normal and the coreferential forms of the postposition **kotý* 'toward':

(8) Normal	Coreferential	
<i>*ché kotý</i>	<i>*wi-je-kotý</i>	'toward me'
<i>*oré kotý</i>	<i>*oro-je-kotý</i>	'toward us EX'
<i>*jané kotý</i>	<i>*jere-je-kotý</i>	'toward us IN'
<i>*né kotý</i>	<i>*e-je-kotý</i>	'toward you SG'
<i>*pé kotý</i>	<i>*peje-je-kotý</i>	'toward you PL'
<i>*i-kotý</i>	<i>*o-je-kotý</i>	'toward him/her'

The coreferential pronominal genitive is (9b) expressed by one of the set of coreferential prefixes:

- (9) a) *i-čý o-s-epják ‘He/she saw his/her mother [nonCOR].’
 3-mother 3A-3O-see
- b) *o-čý o-s-epják ‘He/she saw his/her own mother.’
 3COR-mother 3A-3O-see

Like English and Portuguese, the trigger for coreferential marking is the subject. However, the range of targets is more extensive: it includes any structure which would normally receive set 2 person markers (noun, dependent verb, or postposition) where the referent is coreferential with the subject. In the case of verbs (subordinate or serial) the cross referencing is absolutive, referring to S or O. Thus the presence of the coreferential prefix on the dependent intransitive serial verb indicates its shared subject with the independent verb. In example 10, the independent verb form of ‘I go’ is *a- kjér, whereas the dependent serial verb form is *wi-kjé-a.

- (10) *a-có wi-kjé-a ‘I went and slept’
 1SG-go 1SG.COR-sleep-SER

Likewise, the subordinate verb receives coreferential marking when its referent is identical with the subject of the independent clause. When this verb is transitive the referent is O, as in example 11:

- (11) a) *o-’ár [i-pycýk-VmV] ‘He fell when X grabbed him [nonCOR].’
 3S-fall 3O-grab-WH
- b) *o-’ár [o-pycýk-VmV] ‘He fell when X grabbed him [COR].’
 3S-fall 3COR-grab-WH

Whereas it is not unusual cross-linguistically that the trigger should be nominative-accusative, it is probably quite unusual that the verbal constructions which are the recipients, or targets, of coreferential marking follow an ergative-absolutive cross-referencing system:

TRIGGER	Nominative	A	Ergative	
TRIGGER	Nominative	S	Absolutive	TARGET
	Accusative	O	Absolutive	TARGET

3. Versions of coreferential system

In the Tupí-Guaraní family, some languages have a complete paradigm of coreferential prefixes, even though only the third person referents are potentially ambiguous, and these are used in a wide variety of syntactic contexts. Others have only the third person prefix,

which is used in a reduced number of syntactic contexts. Individual Tupí-Guaraní languages fit into one of four possible categories, as described in Table 2. In this paper I will give an examples each type.

Maximal	- complete paradigm of coreferential prefixes - coreferential referencing extended to subordinate clause
In transition	- reduction in the number of coreferential prefixes and/or - coreferential referencing not extended to subordinate clause
Minimal	- only one coreferential prefix (third person) - limited or no coreferential prefixing on verbs
Nonexistent	- complete elimination of coreferential prefixing

Table 2: Coreferential systems in Tupí-Guaraní languages

4. Maximal version of coreferential system: Tocantins Asurini

In Tocantins Asurini (Nicholson 1978), member of subgroup 4, the full set of coreferential prefixes occurs: *we-* '1SG', *oro-* '1EX', *sere-* '1IN', *e-* '2SG', *pese-* '2PL', *o-/w-* '3'. They occur with transitive subordinate verbs, intransitive serial verbs, nouns, and postpositions. In subordinate clauses, when the O of a transitive verb is identical with the subject of the main verb, the coreferential prefix is used, as in example 13.

- (13) [*we*-nopo -ramo] a-ha-pota (ise)
 1SG.COR-beat-if 1SG-go-FUT 1SG
 'If (someone) beats **me** [COR], **I** will go away.'

In some languages intransitive subordinate verbs may also receive coreferential marking. This requires that the subject of the two clauses be identical. In Tocantins Asurini, however, subordinate clauses are only permitted when the subjects of the two clauses are different, thus precluding the possibility of coreferential marking on the intransitive subordinate verb.

Likewise the coreferential system appears in intransitive serial verbs, as in example 14:

- (14) *ere*-ha e-seegat-a
 2SG-go 2SG.COR-sing-SER
 'You (sg) went singing.'

Nouns receive coreferential prefixes when their possessor is coreferential with the subject. Compare the non-coreferential prefix in 15 a) with the coreferential one in b).

(15) a) *ere-ha-pota s-aga pype* ‘You will go to his house.’
 2SG-go-FUT 3-house to

b) *a-ha-pota w-aga pype* ‘He will go to his own house.’
 3-go-FUT 3COR-house to

These prefixes occur for other persons as well, as in example 16.

(16) *we-tyroa a-kotog* ‘I sew my own clothes.’
 1SG.COR-clothes 1SG-sew

Postpositions also receive coreferential prefixes, as in example 17.

(17) *a-se’eg we-se-ope* ‘I sang to myself.’
 1SG-sing 1SG.COR-REFL-to

The coreferential system used by Tocantins Asurini is illustrated by Table 3:

	Subordinate TV	Serial IV	Noun	Postposition
1 and 2 person	x	x	x	x
3 person	x	x	x	x

Table 3: Tocantins Asurini - Maximal use of coreferential prefixing

Languages other than Tocantins Asurini which make maximal use of coreferential prefixing include Tapirapé (Leite 1987), also from subgroup 4, and Kayabi (Dobson 1988) and Xingu Asurini (da Silva 1995), both from subgroup 5.

5. Languages in transition

Several languages of the family show signs of being in transition from the more extended system of coreferential marking to a reduced system. This is evident in two different ways: a reduction in the paradigm to only a third-person form, and/or a reduction in the number of syntactic contexts in which the cross referencing takes place.

When the coreferential prefixing on nouns and postpositions is reduced to only a third person form, the first and second person prefixes from set 3 are substituted by those from set 2, as illustrated in Table 4:

	Set 2	Set 3
1SG	*čé >	*wi-
1EX	*oré >	*oro-
1IN	*jané >	*jere-
2SG	*né >	*e-
2PL	*pé >	*peje-
3		*o-

Table 4: Paradigm replacement on nouns and postpositions

When the coreferential prefixing on subordinate and dependent serial verbs is reduced to only third person, there are two possibilities as to which set of person markers will substitute the first and second person coreferential prefixes. Either they will be replaced by the absolutive set of person markers (set 2), or they may be replaced by the nominative set (set1), as indicated by the arrows in Table 5. When the first and second person coreferential prefixes are replaced by set 2 person markers, the third person *o-* continues to be recognized as a coreferential prefix, in contrast with the normal third person prefix *i-*. If, however, the first and second person coreferential prefixes are replaced prefixes of set 1, the prefixes *o-* and *oro-* undergo reanalysis as belonging to the nominative rather than the coreferential set, as there is no contrast between the two forms. This is indicated by an = sign in the table.

	Set 1	Set 3	Set 3	Set 2
	A/S	COR O/S	COR O/S	O/S
1SG	*a- >	*wi-	*wi- <	*čé
1EX	*oro- =	(*oro-)	*oro- <	*oré
1IN	*ja- >	*jere-	*jere- <	*jané
2SG	*ere- >	*e-	*e- <	*né
2PL	*pe- >	*peje-	*peje- <	*pé
3	*o- =	(*o)-	*o-	

Table 5: Paradigm replacement on subordinate and dependent serial verbs

Comparative paradigms of intransitive serial verbs suggest that paradigm replacement is a gradual process. The data in Table 6 are arranged by the degree of paradigm replacement.

	Set 3	Kamaiurá	Parintintín	Tupinambá	Guarayú	Guajajara
1SG	*wi-	we-	i-	wi-	vi-	he-
1EX	*oro-	oro-	oro-	oro-	oro-	ore-
1IN	*jere-	jere-	jane-	ja-	ja-	zane-
2SG	*e-	e-	e-	e-	ere-	ne-
2PL	*peje-	peje-	pe-	pe-	pe-	pe-
3	*o-	o-	o-	o-	o-	o-

Direction of change: O/S A/S A/S O/S

Table 6: Paradigm replacement in transition - serial verbs

As can be seen from the table, Kamaiurá (Seki 1989), from subgroup 7, has the full set of coreferential prefixes, whereas Guajajara (Harrison 1986 and p.c.), from subgroup 4, retains only the third-person form. Parintintín, member of subgroup 6, has undergone a partial substitution by person markers from set 2 (Betts 1981). According to Helen Pease (p.c.) there are some remnants of *peje- in Parintintín dependent serial verbs. Both Parintintín and Guajajara have undergone replacement of coreferential forms by those of set 2, retaining the absolutive cross-referencing system along with the third person coreferential prefix *o-*. In Tupinambá (Rodrigues 1953), from subgroup 3, and Guarayu (Newton 1978), from subgroup 2, the coreferential prefixing on dependent intransitive serial verbs has been partially replaced by forms from the nominative set (set 1), with Guarayu being further along in the substitution process than Tupinambá. The *o-* and *oro-* prefixes are not distinguishable from those of the replacement set and can no longer be considered coreferential markers.

6. Minimal version of coreferential system: Wayampi

Minimal use of the coreferential system, with only a third-person form, is illustrated by data from the Jari dialect of Wayampi, a member of subgroup 8. In addition to the paradigm reduction in this language, there has been a reduction in the syntactic environment in which coreferential marking can occur, due to the extension of the split-

ergative cross-referencing system of the independent verbs to dependent verbs which originally had absolutive cross referencing (Jensen 1990).

Person markers from set 2 are used to express the first or second person in combination with nouns and postpositions, even when the referent is identical with the subject of the main clause, as can be seen in the possessive form used in examples 19 and 21. The third person coreferential prefix is used in 18 and 20.

(18) o-o o-koo pe
 3-go 3COR-garden to
 'He/she went to his/her (COR) garden.'

(19) ere-o ne-koo pe
 2SG-go 2SG-garden to
 'You went to your garden.'

(20) jety o-jo'o o-upe
 sweet.potato 3-dig.up 3COR-for
 'She dug up sweet potatoes for herself.'

(21) jety a-jo'o e-upe
 sweet.potato 1SG-dig.up 1SG-for
 'I dug up sweet potatoes for myself.'

Note that when the argument of the postposition is identical with the subject of the sentence, the third person coreferential prefix is used without the reflexive prefix *je- in this language, as in example 20.

Agentive intransitive serial verbs take the same set of prefixes as are used with the main verb of an independent clause, i.e. the nominative prefixes from set 1. In example 22 a-a has replaced *wi-có-B o. The prefix o- in example 23 is no longer identifiable as the coreferential prefix which occurred in *o-có-B o.

(22) a -jywy a-a 'I returned, going'
 1SG-return 1SG-go

(23) o-jywy o-o 'he returned, going'
 3-return 3-go

Likewise subordinate verbs are cross referenced in the same way as independent verbs, with a split-S system for intransitive verbs and the a split-ergative system for transitive verbs. In example 24, the subordinate intransitive verb takes a nominative prefix instead of a coreferential prefix.

(24) [myakyty a-a remě] tapi'i a-jawy jawy evovi pupe
 downriver 1SG-go WH tapir 1SG-miss miss gun with

‘When I [COR] went (hunting) downriver, I repeatedly missed (shooting) a tapir with my revolver.’

The minimal use of the coreferential system in Wayampi is illustrated by Table 7. Mbyá Guaraní (subgroup 1) also shows minimal use of coreferential prefixing.

	Subordinate V	Serial IV	Noun	Postposition
1 and 2 person				
3 person			x	x

Table 7: Wayampi - Minimal use of coreferential prefixing

The motivation for this elimination of coreferential marking on verbs is the elimination of the morphosyntactic environment in which the prefixing originally occurred. As mentioned above, the coreferential prefixing occurred in the context of absolutive cross referencing. In the original system, the potential environment for coreferential prefixing includes subordinate verb and intransitive serial verbs, as indicated by italics in Table 8. Transitive serial verbs, which cross reference O, do not receive coreferential prefixing because it is A that is identical with the subject of the main verb.

	Independent	Subordinate	Dependent serial
Intransitive	Split-S	<i>Absolutive (S)</i>	<i>Absolutive (S)</i>
Transitive	Split ergative	<i>Absolutive (O)</i>	<i>Absolutive (O)</i>

Table 8: Cross-referencing in Proto-Tupí-Guaraní

In Wayampi subordinate verbs, the same system of verbal agreement is used as with independent verbs. That is, a person hierarchy is used with transitive verbs, subject prefixes from Set 1 are used for agentive intransitive verbs, and person markers from Set 2 are used for nonagentive transitive verbs. This system has also been extended to intransitive serial verbs. These cross-referencing changes have resulted in the elimination of the environments in which the coreferential marking originally occurred, as in Table 9.

	Independent	Subordinate	Dependent serial
Intransitive	Split-S	Split-S	Split-S (Sa / So)
Transitive	Split ergative	Split ergative	<i>Absolutive (O)</i>

Table 9: Cross-referencing in Wayampi

7. *Eliminated system: Urubu-Kaapor*

In Urubu-Kaapor, another member of subgroup 8, even the third person coreferential prefix has been eliminated. Thus, no distinction is made between the coreferential possessor in example 26 and the non-coreferential one in example 27, where the same prefix *h-* (<*c-) 'third person', is used in both occurrences of 'his wife' *h-akehar*.

(26) pe kuja pytun mokõi pytun pe h-akehar rehe o-ho ã
and like.this night two night then 3-wife LK-for 3-go also
'And after this many nights, two nights, he went for his wife also.'

(27) ere-rur aja je h-akehar pandu i-pe
2SG-bring thus HRSY 3-wife 3+say 3-to
'"Did you bring it?" thus, it is said, his wife said to him.'

8.0 Conclusion

To summarize the coreferential cross referencing in Tupí-Guaraní languages, those which make maximal use of the system have a high degree of grammaticalized cohesion, but are also redundant, since coreferential markers for first and second person are not necessary to disambiguate referents. On nouns and postpositions the first and second persons are eliminated through the replacement by person markers from set 2. When they are replaced on verbs, there is a choice between the nominative person markers from set 1 (A/S) and the absolutive ones from set 2 (S/O). This choice is partly related to a decrease in the extent of the ergative-absolutive cross-referencing system. If the set 3 prefixes are replaced by the forms from set 2, the coreferential prefix for third person is usually retained. If they are replaced by the set 1 forms, the third person prefix *o-* is reanalyzed as the homonymous prefix from that set for lack of contrast. The underlying nominative-accusative system, as evidenced on the syntactic level by the trigger of the coreferential marking system, is no doubt a principal motivator in the systematic replacement of absolutive person markers by nominative prefixes in the languages of subgroups 1 and 8 (see Jensen 1990). Where this has happened the number of structures in which coreferential marking can occur is reduced, making them "minimal use" languages. Languages which fall in the transitional category allow us to get a glimpse of these changes in process. They serve as a reminder that changes from one system to another are not made overnight, but rather are a gradual process. And the data from Urubu-Kaapor is a reminder that ultimately a system, such as set 3, can be eliminated altogether.

Bibliography

Betts, La Vera. 1981. *Dicionário Parintintín-Português Português-Parintintín*. Brasília: Summer Institute of Linguistics.

da Silva, Heliana Maria. 1995. *Marcadores de Pessoa na língua Asurini do Xingu*. Senior monograph prepared for the Federal University of Pará (Brazil).

Dixon, R. M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press.

Dobson, Rose. M. 1988. *Aspectos da língua Kayabí*. *Série Linguística*. No. 12. Brasília: Summer Institute of Linguistics.

Dooley, Robert A. 1982. *Vocabulário do Guaraní*. Brasília: Summer Institute of Linguistics.

Dooley, Robert A. 1991. A double-verb construction in Mbyá Guaraní. *Work Papers of the Summer Institute of Linguistics, University of North Dakota*, 35: 31-66.

Harrison, Carl. 1986. Verb Prominence, Verb Initialness, Ergativity and Typological Disharmony in Guajajara. *Handbook of Amazonian Languages*. Vol. 1, ed. by Desmond C. Derbyshire and Geoffrey K. Pullum, pp. 407-439. Berlin: Mouton de Gruyter.

Jensen, Cheryl. 1989. *O desenvolvimento histórico da língua Wayampi*. Campinas: Editora da UNICAMP (Universidade Estadual de Campinas).

Jensen, Cheryl. 1990. Cross-referencing changes in some Tupí-Guaraní languages. *Amazonian Linguistics: Studies in Lowland South American Languages*, ed. by Doris L. Payne, pp. 117-158. Austin: University of Texas Press.

Jensen, Cheryl. 1998. The use of coreferential and reflexive markers in Tupí-Guaraní languages. *Journal of Amazonian Languages* (University of Pittsburgh, Department of Linguistics). 1(2):1-49.

Jensen, Cheryl. 1998. Comparative Tupí-Guaraní morphosyntax. *Handbook of Amazonian Languages*, Vol. 4. Ed. by Desmond Derbyshire and Geoffrey Pullum. Berlin: Mouton de Gruyter.

Kakumasu, James. 1986. Urubu-Kaapor. *Handbook of Amazonian languages*, Vol. 1. Ed. By Derbyshire and Pullum, 326-403. Berlin: Mouton de Gruyter.

Leite, Yonne. 1987. Referential Hierarchy and Tapirape Split Marking Systems. MS presented at the Working Conference on Amazonian Languages, University of Oregon.

Lemle, Miriam. 1971. Internal Classification of the Tupi-Guarani Linguistic Family. *Tupi Studies I*, ed. By David Bendor-Samuel, pp. 107-129. Norman, Oklahoma: Summer Institute of Linguistics.

Newton, Dennis. 1978. Guayaru Discourse. *Work Papers of the Summer Institute of Linguistics*. Riberalta: Summer Institute of Linguistics.

Nicholson, Velda. 1978. *Aspectos da língua Assurini*. Brasilia: Summer Institute of Linguistics.

Rodrigues, Aryon Dall'Igna. 1953. Morfologia do verbo Tupí. Curitiba: *Letras, No. 1*.

Rodrigues, Aryon Dall'Igna. 1984/1985. Relações internas na família linguística tupi-guaraní. *Revista de Antropologia* 27/28:33-53. São Paulo.

Seki, Lucy. 1989. Sistema de Marcação de Caso do Kamaiurá. MS presented at the Seminário de Tipologia. UNICAMP.

Wiesemann, Ursula. 1986. Grammaticalized Coreference. *Pronominal Systems*, ed. by U. Wiesemann, pp. 437-481. Tübingen: Gunter Narr Verlag.