

Dans cette partie, nous allons traiter de certains morphèmes qui sont essentiels dans la langue et que l'on retrouve dans différents types de constituants (cf. Chapitre 1, II sur la morphologie commune entre les noms et les verbes). Nous tenons à présenter ces morphèmes en début de thèse d'une part, afin de ne pas constamment obliger le lecteur à se référer à des chapitres ultérieurs et d'autre part, afin de discuter de ces morphèmes omniprésents dans un cadre ouvert et non dans une partie traitant d'un autre thème, plus spécifique, où leur présentation déborderait forcément du thème traité. Ces morphèmes transcatégoriels ont souvent été mentionnés lors du premier chapitre, du fait que certains peuvent être trouvés à la fois sur des noms et des verbes, et que certains ne se combinent qu'avec des arguments ou qu'avec des prédicats, départageant ainsi ces deux fonctions (cf. Chapitre 1, II et III).

Les morphèmes que nous traiterons dans les trois chapitres de cette partie sont les suivants :

- les indices de personne (Chapitre 3) : il s'agit d'un thème essentiel, car les indices de personne sont omniprésents dans la langue. Ils sont obligatoires sur les verbes et les noms dépendants. Il est important de présenter la série d'indices dite II de manière regroupée : en effet, elle n'est pas seulement présente sur les prédicats, où elle s'oppose à la série dite I (la série I marquant le sujet, la série II l'objet), mais elle marque aussi les possesseurs sur les noms ainsi que les objets de postpositions. On voit donc ici un parallélisme entre trois structures impliquant une tête de syntagme et son complément : le nom possédé, le groupe postpositionnel et le verbe transitif avec un indice d'objet.

- les morphèmes référentiant *-a* et relationnel */-* (Chapitre 4), tous deux utilisés dans le contexte où un complément est fortement régi par une tête de syntagme. Ces deux morphèmes sont très fréquents dans la langue.

- les morphèmes de nombre (Chapitre 5) : on présentera ici deux morphèmes indiquant le pluriel du sujet (*-(o)η* et *kupa*), et un autre morphème (*-kom*) dont la distribution est très large (sur les têtes de syntagmes nominaux, les noms possédés, les groupes postpositionnels, les verbes, et les groupes nominaux complexes).

Les deux premiers types de morphèmes (indices de personne, référentiel et relationnel) sont caractéristiques de la famille tupi-guarani, et nous nous attacherons à décrire les spécificités de leur usage en émérillon. Quant aux morphèmes de nombre, aucun n'a été reconstruit en proto-tupi-guarani, étant donné que l'expression de cette catégorie varie considérablement au sein de la famille. Le chapitre 5 décrira donc les outils spécifiques que l'émérillon a développé à cet effet. Ainsi, en plus de présenter des morphèmes incontournables en émérillon, cette partie permettra de situer solidement l'émérillon dans la famille tupi-guarani, à la fois dans son conservatisme et dans son originalité.

Chapitre 3 : Le système d'indices de personne⁵⁵

Ce chapitre a pour objet la description du système des marques de personne en émérillon. Ces indices de personne sont toujours des préfixes en émérillon et ne peuvent pas fonctionner comme tête de syntagme nominal. Ils se répartissent en deux principaux paradigmes appelés "séries". La présence d'un indice de personne est obligatoire sur les verbes (indépendants) alors que les pronoms libres sont utilisés uniquement dans un but d'emphase : on peut dire que l'émérillon est une langue à argument pronominal ("pronominal argument language" d'après Jelinek, par exemple 1990). Les pronoms libres seront donc présentés séparément (au Chapitre 6, II). Nous nous concentrons ici uniquement sur les indices. Les indices émérillon n'apparaissent pas que sur les verbes, mais aussi sur les noms et les postpositions : c'est pourquoi nous trouvons important de les décrire dès le début de ce travail.

En I, nous donnons les formes de ces indices et décrivons le système pronominal, notamment en relation avec le système proto-tupi-guarani. En II, nous observons la distribution des différentes séries d'indices, ce qui nous permet de dégager une structure complément-tête où la série II peut tenir le rôle du complément. Enfin, en III, l'étude plus spécifique de la distribution des indices de personne sur les verbes transitifs nous permet de déterminer le type de système d'accord sur le prédicat en émérillon : un système d'accord hiérarchique basé sur une hiérarchie des personnes et une hiérarchie des rôles sémantiques.

⁵⁵ Ce chapitre est une version remaniée d'une présentation faite dans le cadre de la rencontre WAIL/SSILA du 6 au 8 juillet 2001 à Santa Barbara. Nous remercions F. Queixalós, A. Kibrik et S. DeLancey pour leurs commentaires.

I- Les indices de personne

Les indices de personne se répartissent en sept "personnes" différentes, et en trois séries (selon le classement en séries de Jensen, cf. par exemple Jensen 1998a). Dans la section I, nous nous intéressons spécifiquement aux distinctions de "personne" (c'est à dire les lignes du Tableau 8). Les distinctions de séries seront examinées dans la section II qui se focalisera sur leur distribution et leur fonction.

Dans cette section I, nous verrons dans l'ordre les formes des indices de personne, puis l'organisation du système pronominal, et enfin comment ce système a évolué à partir du système pronominal du proto-tupi-guarani.

I- 1. Les formes des indices de personne

Le tableau suivant présente les indices de l'émérillon triés en séries selon le classement utilisé par Jensen. Synchroniquement, il est peu intéressant de parler de "série" pour la troisième colonne du tableau mais nous utiliserons ponctuellement cette dénomination pour faciliter la comparaison avec d'autres langues tupi-guarani, ou avec le proto-tupi-guarani.

	série I	série II	série III
1 ^{ère} SG	a-	e-	
2 ^{ème} SG	ele-	de-	
1 ^{ère} INCL	si- ⁵⁶	nōde- ⁵⁷	
1 ^{ère} EXCL	olo-	ole- (~ olone-)	
2 ^{ème} PL	pe-	pe- (~ pene-) ⁵⁸	
3 ^{ème}	o-	i- (~ Ø- ~t-) ⁵⁹	o-
indéterminé, général humain	za-	zo- / polo-	

Tableau 8 : Séries des indices de personne de l'émérillon

Tous ces indices sont préfixés aux racines flexionnelles (nom, verbe ou postposition). Aucun ne peut être considéré comme un pronom libre, à la fois du fait de sa proximité phonologique avec la racine et de par son incapacité à servir de tête de syntagme comme le font les pronoms indépendants. Les indices de série II doivent plus précisément être considérés comme des clitiques, étant donné qu'ils sont compatibles avec les trois types de racines flexionnelles : noms, verbes et postpositions.

I- 2. Le système pronominal

L'émérillon est une langue où le système pronominal connaît les distinctions de personne (1^{ère}, 2^{ème} et 3^{ème}) selon les deux corrélations proposées par Benveniste (1966) – la corrélation de personnalité (je/tu vs il) et la corrélation de subjectivité

⁵⁶ *si-* est réalisé /se/ devant des radicaux à voyelle initiale /i/. ex : la séquence *si-* 1^{ère} INCL.I + *ijunŋ* "mettre" est réalisée /sejnuŋ/ "nous mettons".

⁵⁷ *Nōde-* connaît une variante dialectale *kōde-*.

⁵⁸ La variation entre *ole-* et *olone-* et entre *pe-* et *pene-* est discutée au chapitre 4, II- 2.

⁵⁹ Les formes entre parenthèses apparaissent sur les racines prenant le relationnel, morphème qui sera introduit en III-4, et discuté plus en détail au chapitre 4, II.

(je vs tu). Elle possède de plus la distinction de pluriel, et la distinction inclusif/exclusif. Sa morphologie ne distingue par contre ni de duel, ni de genres, ni différents degrés de respect, ni le sexe du locuteur. Nous précisons maintenant le contenu des appellations "inclusif/exclusif" et "indéterminée", en précisant à l'avance que leur distinction est brouillée par l'usage.

I- 2.1. L'opposition inclusif/exclusif

L'émérillon possède deux marques de personne du pluriel impliquant le locuteur. La différence fonctionnelle entre les deux semble correspondre à la distinction bien connue en typologie entre une marque d'inclusif (incluant l'interlocuteur) et une marque d'exclusif (excluant l'interlocuteur), avec un écart mentionné ci-dessous. Comparons les deux phrases suivantes qui sont tirées de dialogues. Dans la première, *si-*, et *nōde* renvoient au locuteur (Dzasol) et à son frère à qui il s'adresse. Les inclusifs *si-* et *nōde-* font donc bien référence au locuteur et à l'interlocuteur. Dans la deuxième, une fille (locutrice) dit à sa mère ce qu'elle-même et sa sœur vont faire dans la journée. *olo-* réfère au locuteur et à une troisième personne, mais pas à l'interlocuteur : c'est un exclusif.

(100) "**si-ɕa-o-ne**⁶⁰ siʔe nōde, **nōde-l-aihi-pa-ɕi-la.**"
 1INCL.I-INDET.I-aller-CONTR même PRO 1INCL 1INCL.II-RELN-aimer-COMPL-NEG-TAM
 "Allons-y nous-même, parce qu'on ne nous aime pas."

(101) "Mama, **olo-ho-tal** baipuli **olo-zopodɕ.**"
 maman 1EXCL.I-aller-FUT tapir 1EXCL.I-nourrir
 "Maman, nous allons nourrir le tapir."

Dans les narrations, l'exclusif indique toujours clairement l'exclusion de l'auditeur. Dans l'exemple suivant, *olo-* réfère à la narratrice et sa famille, et n'inclut pas l'auditeur.⁶¹

⁶⁰ La cooccurrence de *si-* et *za-* (~ *ɕa*) sera discutée en I-3.2.

⁶¹ Grenand (1980) restreint le sens du *olo-* wayampi à "1^{ère} personne duel". Notre corpus indique que ce n'est pas le cas en émérillon.

- (102) kōʔem **olo**-ho-tal esag t-**olo**-wikiᵑᵑ.
 demain 1EXCL.I-aller-FUT voir BUT-1EXCL.I-pêcher
 Demain nous irons voir pour pêcher.

Par contre, dans les textes narratifs ou expositifs, il existe un sens où la personne inclusive n'inclut en réalité pas l'auditeur (le linguiste qui enregistre), car elle réfère alors à l'ensemble de la communauté émérillon. Il semble que dans ce cas, la notion de "communauté" soit plus forte que la présence d'un auditeur extérieur.⁶²

- (103) aʔe-a-te nōde teko-kom **si**-ᵑᵑu.
 DEM-A-FOC PROINCL émérillon-PL 1INCL-être
 Ceux-là, c'est nous, nous sommes les Emérillons.

I- 2.2. Les personnes indéterminées

Les formes dites "indéterminées" ne renvoient pas à un référent spécifique, mais toujours à un référent général et humain.

Comme indice de série I, *za-* (~ *ᵑᵑa* après un /i/) indique un sujet général non déterminé souvent traduit par "on".

- (104) nan-āhā-te aᵑ baʔek^wəl **za**-idu.
 ainsi-seulement-FOC DEM histoire INDET.I-entendre
 C'est comme ça qu'on raconte cette histoire.
- (105) i-puli-we-ʔe baʔe i-pope pazalu **za**-iᵑuᵑ-a-maʔē.
 3.II-à.côté.de-aussi-INTENS chose 3.II-dans cachiri INDET.I-mettre-a-REL
 Et à côté il y a la chose dans laquelle on met du cachiri (= une marmite).

Mais en fait, dans plusieurs textes, *za-* réfère à l'ensemble de la communauté émérillon.

- (106) wane-maʔē **za**-ikiᵑᵑ-tal apam-a-wi.
 bien-REL INDET.I-prendre-FUT étranger-a-ABL
 On (les Emérillons) va prendre ce qui est bien aux étrangers.

⁶² La même remarque a été faite pour la 1^{ère} personne inclusive du sikvani par Queixalós (1998). Il propose l'explication suivante, transposable à l'émérillon : "Evidemment la grammaire sikvani fonctionne en général dans des situations où un Sikvani parle à un autre Sikvani. Sans doute, et de ce fait, la 4^{ème} personne [1^{ère} personne inclusive] "adhère" à la forme du discours propre aux textes traditionnels et/ou prescriptifs au point qu'elle continue à y figurer quand la nature de l'interlocuteur devrait en principe l'en exclure."(p. 258)

Avec ce sens-là, il alterne souvent avec l'inclusif :

- (107) nani-pamẽ **si-ɕu** **nõde** les jeunes, nani-pamẽ **za-ɕu**.
 ainsi-tous 1INCL.I-être PROINCL les.jeunes ainsi-tous INDET.I-être
 Nous sommes tous comme ça, les jeunes, on est tous comme ça.

Ces deux sens légèrement différents de *za-* s'approchent de la description de *ya-* faite par Rodrigues (1990) pour le tupinamba : sur un verbe intransitif, il marque un sujet de 1^{ère} personne inclusive et sur un verbe transitif un sujet non spécifié. L'explication proposée est que dans les deux cas, on a la configuration "1^{ère} et 2^{ème} personnes focalisées, 3^{ème} personne non focalisée". En émérillon, *za-* avec les deux sens (générique ou "nous, les Emérillons") se trouve à la fois sur les verbes transitifs et intransitifs.

Dans la série II, *zo-* est utilisé comme objet de postposition et comme possesseur pour un référent général et humain que l'on peut souvent traduire par "les gens". Contrairement à *za-* (série I), on n'observe pas avec *zo-* le sens de "nous les Emérillons". L'indice *zo-* est utilisé sur un nom dont le possesseur n'est pas déterminé, dans une phrase à caractère général (108). C'est aussi la forme utilisée par défaut sur les noms obligatoirement possédés pour ne pas spécifier le possesseur (109), ou sur une postposition pour exprimer un référent indéterminé.

- (108) ...i-moda-mãʔẽ-kom **polo-pihig** **zo-kalakuli-kom** o-ikiɕ...
 3.II-vol-REL-PL INDET.II-attraper INDET.II-argent-PL 3.I-prendre
 les voleurs attrapent les gens et prennent leur argent

- (109) mĩn-a-we **zo-ĩpi** o-pa.
 jadis-a-aussi INDET.II-ancêtre 3.I-être.fini
 Autrefois, les anciens sont décédés.

- (110) kol a-potal zãdam-am t-o-maʔẽ **zo-ehe**.
 puis 1SG.I-vouloir gendarme-TRANSL BUT-3.I-regarder INDET.II-POSTP
 Donc je voudrais des gendarmes pour surveiller les gens.

Quant à *polo*, comme on le voit en (108) et en (111), il est aussi utilisé pour un référent général et humain mais dans la position d'objet du verbe.

- (111) **polo-ʔu-pa** **boɟ.**
 INDET.II-manger-COMPL serpent
 Le serpent a mangé tout le monde.

Le fait que les distributions de l'inclusif, de l'exclusif et de l'indéterminé se chevauchent va maintenant être expliqué par l'étude de l'évolution du système pronominal de l'émérillon à partir du système proto-tupi-guarani.

I- 3. Conservatisme et évolution du système

Le système pronominal de l'émérillon est plutôt typique des langues tupi-guarani. Dietrich, dans un ouvrage sur la classification interne des langues tupi-guarani, note la chose suivante :

"One of the most prominent characteristics of a Tupi-Guarani language is its manner of person marking. Nearly all these languages have the same semantic distinctions." (Dietrich, 1990, p.64)

Le tableau des indices de l'émérillon peut donc facilement être comparé au tableau suivant (repris de Jensen 1990, p.120) qui donne les formes des indices reconstruits pour le proto-tupi-guarani.

	série I	série II	série III	série IV
1 ^{ère} SG	† a-	† če (r-)	† wi-	
2 ^{ème} SG	† ere-	† ne (r-)	† e-	† oro- (1→2SG)
1 ^{ère} INCL	† ja-	† jane (r-)	† jere-	
1 ^{ère} EXCL	† oro-	† ore (r-)	† oro-	
2 ^{ème} PL	† pe-	† pe (n-)	† peje-	† opo- (1→2PL)
3 ^{ème}	† o-	† i-, c-	† o-	

Tableau 9 : Reconstruction des séries d'indices de personne du proto-tupi-guarani

On observe tout d'abord une grande similarité des systèmes et des formes entre le proto-tupi-guarani et l'émérillon (l'évolution des formes particulières correspond

à l'évolution phonologique attestée dans le reste de la langue). Certaines différences subsistent et peuvent être résumées ainsi :

- les indices de série II du proto-tupi-guarani sont parfois considérés comme des morphèmes indépendants qui portent un accent. Dans certaines langues, ils sont devenus des préfixes (Jensen 1998b). En émérillon, les indices de série II sont des préfixes phonologiquement liés à la racine, comme les indices des autres séries.
- la disparition de la quasi totalité des formes de série III en émérillon (discutée au chapitre 15, II et surtout au chapitre 17)
- la reconstruction de deux formes de série IV dites "porte-manteaux" là où nous ne voyons qu'une réutilisation de formes d'autres séries (d'où l'absence d'une série IV dans le Tableau 8). Les reconstructions de cette série IV seront remises en question dans la partie III de ce chapitre.
- l'apparition de formes indéterminées
- l'évolution des formes de première personne inclusive *ja-* et *jane-* du proto-tupi-guarani en *si-* et *nõde-* en émérillon

La plupart de ces évolutions seront traitées séparément dans les parties de la thèse citées entre parenthèses, car les indices en question participent à des systèmes plus larges qui évoluent de manière complexe. Nous traiterons maintenant uniquement de l'origine des formes indéterminées en émérillon, et de l'évolution des marques d'inclusif car ces deux changements affectent peu le reste de la langue.

I- 3.1. L'apparition des formes indéterminées de l'émérillon

Le tableau des indices de personne émérillon (Tableau 8) donne trois formes dites "indéterminées" : *za-*, *zo-* et *polo-*. Or, cette catégorie de "personnes" n'apparaît pas dans le système proto-tupi-guarani (Tableau 9). Dans cette section, nous nous interrogeons sur l'histoire de ces trois formes.

Pour la forme indéterminée de série I *za-* de l'émérillon, nous posons l'hypothèse que c'est un réflexe du † *ja-* de série I du proto-tupi-guarani. Le changement du

proto-tupi-guarani † /j/ à l'émérillon /z/ (réalisé [ɟ ~ z]) est attesté ailleurs : † *jatxy*⁶³ "lune" a donné *zai*, † *pajé* "chamane" a donné *paze*, et † *jakaré* "caïman" a donné *ɟakale*. Si le † *ja-* proto-tupi-guarani doit vraiment être analysé comme un indice de 1^{ère} personne inclusive⁶⁴, on est donc confronté à une forme inclusive qui a perdu son utilisation comme indicateur de référents spécifiques ("moi, toi et lui" dans une situation particulière) et n'a gardé que le sens de "nous les humains, nous les Emérillons". Cette évolution ne paraît pas unique dans la famille : Jensen (1990, p.131) indique que le wayampi a deux préfixes d'inclusif : *ja-* avec les verbes intransitifs, et *si-* avec les verbes transitifs, mais que *ja-* apparaît avec les deux types de verbes quand son sens est non-marqué ou générique. Sur le plan typologique, cette évolution serait en contrepoint de l'évolution plus commune de pronoms indéfinis en marque de 1^{ère} personne (comme le "on" français utilisé pour "nous").

Quant à la forme *zo-*, son origine nous est inconnue. On peut juste noter qu'il existe en proto-tupi-guarani un préfixe † *jo-* de réciproque, qui aurait dû évoluer en émérillon en *zo-*. Or l'émérillon ne possède pas en synchronie de marque spécifique de réciproque. C'est le réfléchi *ze-* qui prend en charge les rôles du réciproque (cf. Chapitre 17, III- 1). L'évolution fonctionnelle d'une marque de réciproque en marque de situation générale ou collective est possible. Ainsi, Lichtenberk (2000) montre l'évolution dans les langues austronésiennes d'une marque réciproque en marque de situation collective. Rappelons cependant qu'en émérillon, *zo-* n'apparaît pas comme le proto-tupi-guarani † *jo-* sur les verbes et les postpositions, mais comme indice de personne sur les noms et les postpositions.

Enfin, en ce qui concerne *polo-*, son origine est attestée. Néanmoins, son statut pose problème car il reflète une évolution diachronique d'un nom indépendant "les

⁶³ Terme utilisé dans Rodrigues, 1984/85. "y" marque une voyelle haute centrale /i/.

⁶⁴ Rappelons que pour Rodrigues (1990), en tupinambá, *ya* marque un sujet de 1^{ère} personne inclusive sur un verbe intransitif, et un sujet non spécifié sur un verbe transitif. Sa classification comme 1^{ère} personne inclusive en proto-tupi-guarani n'est donc pas si évidente.

gens" (tupi-guarani), à un nom d'objet générique humain incorporé (tupi-guarani), puis à un indice d'objet indéterminé (émérillon). En effet, la forme *polo* est décrite par Jensen (1998a, p.536) comme un objet générique humain incorporé signifiant "personne", parallèle à un objet générique non-humain signifiant "chose" † *ma'e*. En émérillon, nous n'avons pas d'exemple où *ba?e* "chose" est incorporé. Quant à *polo*, il n'existe pas comme forme libre mais est seulement utilisé comme objet générique humain (cf. Chapitre 8, I-2).

(112) **polo-?u-pa** **boɟ.**
 INDET.II-manger-COMPL serpent
 Le serpent a mangé tout le monde.

(113) ...i-moda-mã?ẽ-kom **polo-pihig** zo-kalakuli-kom o-ikiɟ...
 3.II-vol-REL-PL INDET.II-attraper INDET.II-argent-PL 3.I-prendre
 les voleurs attrapent les gens et prennent leur argent

Ces exemples montrent que *polo-* est toujours antéposé à un verbe qui ne prend alors pas d'indice de personne. Cette position dans le groupe verbal n'en fait pas un objet ordinaire (qui précède ou suit un verbe porteur d'un indice de personne)⁶⁵, ni un objet incorporé (qui s'insère entre indice de personne et verbe, cf. Chapitre 10, III). Il a en fait toutes les caractéristiques d'un indice d'objet. Cette situation semble être la preuve d'une première étape de grammaticalisation d'un nom *polo* (depuis disparu en émérillon) comme indice d'objet générique humain (assimilable à la série II).

⁶⁵ C'est seulement sur les gérondifs que cette position de *polo* est normale (1). En effet, sur les gérondifs comme *ekal* en (2), l'objet *pulelu* précède le verbe qui ne nécessite pas d'indice de personne.

(1) o-?u(l)-tal kuwi **polo-mõ-ma?am.**
 3.I-venir-FUT un.jour INDET.II-faire-lever
 Il viendra un jour ressusciter les gens.

(2) o-ho-tal **pulelu-l-eka**
 3.I-aller-FUT crapaud-RELN-chercher
 Il va chercher le crapaud.

Peut-être est-ce à partir de cette structure objet-gérondif que *polo* a été utilisé par analogie comme indice d'objet antéposé à un verbe principal.

L'évolution du tupi-guarani à l'émérillon constitue donc un exemple de grammaticalisation du mot "personne" en indice de personne indéfinie "on". Ce modèle de grammaticalisation est bien attesté. Dans leur dictionnaire de la grammaticalisation, Heine & Kuteva (2002, p.232) citent plusieurs exemples, dont le portugais et l'albanais où *a pessoa* "la personne"⁶⁶ et *njeri* "personne" doivent être compris comme "quelqu'un, on".

I- 3.2. L'évolution des indices inclusifs

Il nous faut maintenant expliquer la différence entre les formes inclusives du proto-tupi-guarani et de l'émérillon : † *jane* > *nōde* (série II) et † *ja-* > *si-* (série I).

Pour le passage du proto-tupi-guarani † *jane-* à l'émérillon *nōde-*, on a concernant l'évolution des consonnes le passage de /j/ à /n/ et celui de /n/ à /d/ (réalisé phonétiquement [d ~ nd]). Or, dans la famille tupi-guarani, /j/ et /ɲ/ alternant souvent, de même que les nasales et les pré-nasales (Dietrich, 1990, p.19-20). L'évolution de † *jane* à *nōde* est tout à fait acceptable diachroniquement : on peut affirmer que *nōde-* est un réflexe de † *jane-*.

Quant au fait que dans la série I, † *ja-* a donné *si-* en émérillon (réalisé [si ~ tsi]), l'évolution phonologique seule ne peut l'expliquer. Jensen (1987, p.50-52) propose une explication de l'origine d'un morphème † *ti-* qui aurait comme réflexe *si-* pour trois langues tupi-guarani dont le wayampi : ce morphème serait une forme dérivée de la séquence / t-ja-i /, où † *t-* est un préfixe de but, † *ja* la première personne inclusive de série I et † *i-* l'indice de 3^{ème} personne de série II.⁶⁷ Un fait semble confirmer cette analyse de manière probante : en émérillon, le marqueur de but/exhortation *t-* (~ *ta* ~ *te-*) peut apparaître devant toutes les personnes, sauf *za-* et *si-*.

⁶⁶ Au Brésil, c'est "a gente" qui prend ce sens.

- (114) "a-wu(l)-tal a-ho **t-a-poʔo** iŋa-m."
1SG.I-monter-FUT 1SG.I-aller BUT-1SG.I-cueillir pois.sucré-TRANSL
Je vais monter pour cueillir un pois sucré.
- (115) "e-k^wa sisig-a-kom-a-l-ehe **t-~~ele~~-maʔẽ**."
2SG.IMP-partir sœur-a-PL-a-REL-POSTP BUT-2SG.I-voir
"Pars voir tes sœurs".
- (116) o-ijnuŋ **t-o-kaʔẽ** ikeʔi.
3.I-mettre BUT-3.I-boucaner alors
Alors elles les mettent à boucaner.
- (117) kob ɕuliba e-l-apidɕ-a-pe **t-olo-ike**. exemple élicité
COP escalier 1SG.II-RELN-maison-a-dans BUT-1EXCL.I-entrer
Il y a un escalier pour que nous rentrions dans la maison.
- (118) wane **te-pe-ɕapiaka-õwã** wilakala-koti.
bien BUT-2PL.I-penser-un.peu dieu-POSTP
Il faut que vous pensiez un peu à Dieu.
- (119) *baʔezaʔu a-mumuŋ **ta-si-zopodɕ** pita-kom.
nourriture 1SG.I-faire.cuire BUT-1INCL.I-nourrir enfant-PL
Je fais à manger pour que nous nourrissions les enfants.
- (120) *kob ɕuliba e-l-apidɕ-a-pe **ta-za-ike**.
COP escalier 1SG.II-RELN-maison-a-dans BUT-INDET.I-entrer
Il y a un escalier pour qu'on rentre dans la maison.
- En même temps, une forme verbale avec *si-* peut être interprétée comme ayant un sens de but ou d'exhortation, sans morphème *t-* identifiable, ce qui semble indiquer que *t-* est déjà "présent" dans *si-*.
- (121) **si-ɕapiaka-õwã** i-koti.
1INCL.I-penser-un.peu 3.II-POSTP
Pensons un peu à lui.
- (122) kob ɕuliba e-l-apidɕ-a-pe **se-ike**.⁶⁸
COP escalier 1SG.II-RELN-maison-a-dans 1INCL.I-entrer
Il y a un escalier pour que nous rentrions dans la maison.
- (123) baʔezaʔu a-mumuŋ **si-zopodɕ** pita-kom.
nourriture 1SG.I-faire.cuire 1INCL.I-nourrir enfant-PL
Je fais à manger pour que nous nourrissions les enfants.

⁶⁷ Remarquons que la présence d'un indice de série II (objet) à la source de ce morphème sous-entend que cette grammaticalisation a d'abord opéré uniquement sur des verbes transitifs, ce qui serait encore le cas en wayampi d'après Jensen (1987, p.51).

⁶⁸ /si/ est réalisé [se] devant un /i/. cf. note 56.

On peut noter que curieusement, la cooccurrence de *si-* et de *za-* (ou son allomorphe /ɟa/) est fréquente, sans qu'elle puisse être expliquée :

- (124) *nōde wiŋi si-ɟa-o-l-ehe, ...*
 PROINCL là-bas 1INCL.I-INDET.I-aller-REL-parce.que
 Parce que nous sommes allés là-bas, ...

Le tableau suivant récapitule les quelques évolutions d'indices de personne entre le proto-tupi-guarani et l'émérillon que nous venons d'expliquer.

proto-tupi-guarani	émérillon
† <i>ja-</i> "inclusif, série I"	<i>za-</i> "indéterminé, série I."
? († <i>jo-</i> "réciproque")	<i>zo-</i> "indéterminé, série II, comme possesseur du nom et objet de postposition"
† <i>polo-</i> "objet générique humain incorporé"	<i>polo-</i> "indéterminé, série II, comme objet de verbe"
† <i>jane-</i> "inclusif, série II"	<i>nōde-</i> "inclusif, série II"
? † <i>t-ja-i</i> "BUT-INCL.I-3.II"	<i>si-</i> "inclusif, série I"

Tableau 10 : Evolution d'indices entre le proto-tupi-guarani et l'émérillon

Nous avons montré les différentes catégories de "personnes" qui constituent le système des indices de personne de l'émérillon (les lignes du Tableau 8), en nous penchant plus particulièrement sur la distinction parfois vague entre les personnes inclusive, exclusive et indéterminée. Nous avons ensuite vu que ce système diffère peu du système proto-tupi-guarani, et avons essayé d'expliquer deux de ces différences : l'apparition des formes indéterminées en émérillon, et l'évolution des indices inclusifs.

Nous pouvons maintenant regarder l'autre paramètre de classement des indices de personne : les séries (les colonnes du Tableau 8). Nous allons observer la distribution de ces différentes séries, pour voir quelles sont leurs fonctions. Nous ne

ferons plus référence aux formes particulières, mais seulement aux séries dans lesquelles elles sont organisées.

II- La distribution des séries

Dans cette section, nous présentons la distribution des trois séries sur les trois types de racines flexionnelles de l'émérillon (noms, verbes et postpositions), ainsi que leurs fonctions. Nous laissons pour la section suivante la discussion plus spécifique sur le système d'accord sur le verbe en émérillon, discussion qui se basera sur les faits présentés ici concernant les personnes sur les racines verbales.

Voici un tableau annonçant la distribution de chaque série.

	sujet du verbe	objet du verbe	possesseur du nom	objet de postposition
série I	x			
série II		x	x	x
série III			x	x

Tableau 11 : Distribution des séries d'indices de personne en émérillon

Nous allons maintenant illustrer chacune de ces fonctions. Rappelons que l'important dans cette section n'est pas la forme particulière de l'indice, ou la personne à laquelle il renvoie, mais uniquement la série à laquelle il appartient.

La série I marque le sujet d'un verbe qu'il soit intransitif ou transitif. C'est sa seule fonction.

(125) **a-ʔita ʔi-pope.** sujet d'un verbe intransitif
 1SG.I-nager rivière-dans
 Je nage dans la rivière.

(126) **a-nupã.** sujet d'un verbe transitif
 1SG.I-frapper
 Je l'ai frappé.

La série II marque l'objet sur un verbe transitif, le possesseur sur un nom (qui une fois possédé, peut prédiquer), et l'objet d'une postposition.

- (127) *zawal e-suʔu.* objet du verbe
 chien 1SG.II-mordre
 Un chien m'a mordu.
- (128) *e-mebil.* possesseur du nom
 1SG.II-enfant
 mon enfant / J'ai un enfant.
- (129) *o-naʔaŋ-tal e-koti.* objet de postposition
 3.I-rassembler-FUT 1SG.II-chez
 Ils se rassemblent chez moi.

La conjonction de ces fonctions en un seul morphème n'est pas inattendue. En Amazonie, les affixes possessifs marquent typiquement aussi les objets de postposition, et correspondent souvent à la série d'indices qui marque les objets (Aikhenvald, 1986, Unit 3, p.4).⁶⁹

La série III du proto-tupi-guarani n'a qu'un seul réflexe en émérillon : la marque de 3^{ème} personne *o-*. On la trouve dans deux des contextes de la série II : comme possesseur du nom (130) et comme objet de postposition (131). Contrairement à la marque de 3^{ème} personne de série II *i-*, *o-* indique la coréférentialité avec le sujet de la proposition (131) *vs* (132). Le chapitre 17 sera consacré à l'utilisation de cette marque.

- (130) *o-ikiɕ o-iba o-poʔã-pope.*
 3.I-prendre 3.COREF-animal 3.COREF-main-dans
 Il_i prend son_i animal (une grenouille) dans sa_i main.⁷⁰

⁶⁹ Ceci n'est pas seulement vrai en Amazonie : la même série d'affixes personnels est utilisée sur les noms et comme complément d'adposition en arabe, en maya et en hongrois, par exemple. Cette même série indique aussi l'objet en arabe, le sujet ergatif en maya, et ne s'affixe pas sur les verbes en hongrois (Creissels, c.p.).

⁷⁰ Dans la traduction, l'indice "i" sur plusieurs participants indique qu'ils ont le même référent. Un indice "j" indique la non-coréférence avec un argument marqué par un indice "i".

- (131) o-ɪnuŋ o-wɪb.
 3.I-mettre 3.COREF-sous
 Elle_i le met sous elle-même_i.
- (132) o-bowɪg i-wɪb.
 3.I-charger.le feu 3.II-sous
 Elles_i mettent du bois dessous (i.e. sous le boucan_j).

Avant de passer à la description du système d'accord sur les verbes, remarquons que si la série I et IV sont spécialisées dans l'indexation des personnes sur le verbe, les séries II et III apparaissent sur différents types de racines. Il est notamment important de noter que la série II a trois fonctions selon sa distribution : objet du verbe, possesseur du nom et objet de postposition, et que c'est dans ces mêmes contextes que l'on retrouvera le morphème relationnel *l-* discuté dans le chapitre suivant. Si l'on veut caractériser l'ensemble de ces trois fonctions, elles ont en commun d'exprimer une dépendance forte entre une tête et son complément le plus fortement régi (verbe-objet, nom-possesseur, postposition-objet). Mais derrière le partage de ces trois fonctions par la même série d'indices, il ne faut pas forcément chercher une explication fonctionnelle ou cognitive. Si cette série est omniprésente, c'est plus probablement le résultat d'une cliticisation des pronoms en indices dans différentes positions, à une même étape chronologique que le résultat d'une affinité fonctionnelle⁷¹. Nous nous concentrons maintenant sur l'organisation de l'indexation des personnes sur le verbe, afin de caractériser le système d'accord de l'émérillon.

III- Le système d'accord hiérarchique

Le point de départ de notre description est le suivant : sur les verbes, la série I marque le sujet et la série II marque l'objet.

⁷¹ Nous remercions D. Creissels pour nous avoir éclairée sur ce sujet.

	sujet	objet
série I	x	
série II		x

Tableau 12 : Utilisation des séries d'indices de personne sur les verbes transitifs

Si l'on précise qu'il n'y a généralement qu'une position pour un indice de personne sur les verbes en émérillon, la question se pose de savoir dans une phrase précise, quel argument est marqué (donc quelle série est utilisée). Une question plus large est de savoir comment caractériser un tel système d'accord : actif/statif, nominatif/accusatif, absolutif, inverse... ? Puisqu'en émérillon les relations entre arguments et prédicat sont marquées essentiellement par des marques d'accord sur le prédicat (c'est une langue à argument pronominal), et qu'il n'y a pas sur les constituants nominaux de marque morphologique de leur fonction (autre qu'oblique), c'est uniquement le choix des séries de personne sur le prédicat qui indique les relations arguments/prédicat. Ainsi, étudier ces relations arguments/prédicat revient d'abord à étudier le système d'accord sur le verbe, qui repose lui-même sur le choix des séries.

Dans cette section, nous commençons par présenter en termes larges les spécificités du système d'accord émérillon parmi les langues tupi-guarani (III-1). Puis nous montrons dans quels types de configurations des arguments A et P⁷² on trouve les différentes séries sur un verbe transitif émérillon (III-2). Nous discutons ensuite deux points qui sont au centre des débats sur les systèmes d'accord dans la linguistique tupi-guarani : d'abord la hiérarchisation des personnes et des rôles sémantiques (III-3), ensuite l'hypothèse d'une analyse de ces systèmes en termes de "système inverse" (III-4). Après avoir examiné ces points, nous concluons que l'émérillon connaît un système d'accord hiérarchique.

⁷² Nous utilisons par commodité les abréviations suivantes : S pour le sujet d'une proposition intransitive, A pour le sujet d'une proposition transitive, P pour l'objet d'une proposition transitive.

indépendants, et un système "absolutif" pour les formes verbales dépendantes. Il sera montré au chapitre 15, II que les verbes dépendants⁷³ de l'émérillon sont dans leur grande majorité marqués comme les verbes indépendants (seuls les gérondifs transitifs résistent au changement). La section présente va donc laisser de côté ces résidus d'un système d'accord différent sur les verbes dépendants, qui seront traités en détail au chapitre 15, II.

- les formes locales : les mots descriptifs et les formes dépendantes étant exceptés, le reste du système de marques de personne sur le verbe en émérillon est assez typique de la famille tupi-guarani. L'émérillon connaît néanmoins quelques spécificités au niveau des configurations locales. Une configuration de participants à un événement est dite locale quand les deux participants sont des personnes de l'interlocution : $1 \rightarrow 2^{74}$ ou $2 \rightarrow 1$. Nous insisterons sur les formes locales de l'émérillon, ces données remettant en question la hiérarchie des personnes traditionnellement acceptée $1 > 2 > 3$ (voir par exemple Jensen, 1990).

Le choix des séries sur le verbe transitif émérillon est maintenant expliqué.

III- 2. Les séries sur le verbe transitif

Tout verbe transitif prend deux arguments, A et P. Un fait de première importance en émérillon est que seulement un des arguments est marqué sur le verbe transitif. Que ce soit A qui soit marqué, par la série I, ou P, par la série II, dépend de leur position relative sur une hiérarchie des personnes.⁷⁵ La prise en

⁷³ Par "verbes dépendants", il faut comprendre "formes dépendantes du verbe" et non une classe particulière de verbe. Les propositions dépendantes du proto-tupi-guarani comptent parmi elles les subordonnées temporelles/conditionnelles, les gérondifs, les constructions à oblique topicalisé, et enfin les nominalisations (cf. Chapitre 15, I).

⁷⁴ $1 \rightarrow 2$ doit être lu "un sujet de première personne agit sur un objet de 2ème personne". $1/2 \leftrightarrow 3$ doit être lu "un sujet de première ou deuxième personne agit sur une troisième personne, ou vice-versa".

⁷⁵ Le terme "hiérarchie des personnes" que nous utilisons correspond grosso modo aux désignations rencontrées dans la littérature telles que "animacy hierarchy" (hiérarchie selon le caractère plus ou moins animé), "hiérarchie d'agentivité", "hiérarchie référentielle" ou "hiérarchie de topicalité inhérente". Notre choix n'est pas le résultat d'une prise de position théorique, mais découle du fait que, pour décrire l'émérillon, la référence aux personnes grammaticales est suffisante.

compte de la position des arguments sur une hiérarchie des personnes pour déterminer la marque d'accord à utiliser sur un verbe transitif a été largement répertoriée dans la littérature de linguistique typologique et de linguistique amérindienne, et toujours décrite comme un système mineur. La plupart de la littérature utilise le terme de système "direct-inverse" ou "inverse", et se base sur l'exemple prototypique des langues algonquines, où une marque sur le verbe indique si A est plus haut que P dans la hiérarchie ("direct"), ou le contraire ("inverse") : Payne (1997, p. 209-216), Givón (1994a et b), Klaiman (1991), Mithun (1999, p. 222-228). Une autre partie de la littérature ne restreint pas les systèmes d'accord prenant en compte une hiérarchie des personnes aux seuls systèmes dits "inverses", mais considère que ceux-ci constituent un cas particulier de systèmes "hiérarchiques" : DeLancey (non daté, 1981), Nichols (1992, p. 66-69), Heath (1998). Nous laissons pour la fin de cette partie (III-4) la discussion sur le caractère approprié des termes "hiérarchique" ou "inverse" pour ce genre de systèmes.

Pour savoir quel argument (A ou P) doit être marqué sur le verbe transitif, il faut distinguer trois configurations, selon les personnes des arguments :

- quand l'un des participants est une des personnes de l'interlocution⁷⁶, l'autre est une troisième personne
- quand les deux participants sont des troisièmes personnes
- les configurations locales (1 → 2 et 2 → 1).

III- 2. 1. Un des participants est une des personnes de l'interlocution, l'autre est une troisième personne

Cette première configuration obéit à une hiérarchie des personnes très claire : 1 > 3 et 2 > 3. Le participant placé le plus haut sur la hiérarchie est celui qui sera

⁷⁶ Nous entendons par "personnes de l'interlocution" la première et la deuxième personnes.

marqué sur le verbe, que ce soit A ou P. Par conséquent, quand $1 \rightarrow 3$ ou $2 \rightarrow 3$, seul A est marqué, par une marque de série I :

- (135) a-nupā. 1 → 3
 1SG.I-frapper
 Je l'ai frappé.
- (136) ele-ɕjika ! 2 → 3
 2SG.I-tuer
 Tu l'as tué !

Par contre, quand $3 \rightarrow 1$ ou $3 \rightarrow 2$, seulement P est marqué, par une marque de série II :

- (137) e-zo-zopoɕ. 3 → 1
 1SG.II-RED-nourrir
 On me nourrit tout le temps.
- (138) zawal de-suʔu. 3 → 2
 chien 2SG.II-mordre
 Un chien t'a mordu.

Donc, quand une personne de l'interlocution ($1^{\text{ère}}$ ou $2^{\text{ème}}$) interagit avec une $3^{\text{ème}}$ personne, c'est toujours la personne de l'interlocution qui est marquée, avec la série I pour A, avec la série II pour P. La hiérarchie des personnes est $1/2 > 3$.

III- 2. 2. Les deux participants sont des troisièmes personnes

Quand les deux arguments sont tous deux de $3^{\text{ème}}$ personne, le préfixe *o-* de $3^{\text{ème}}$ personne de série I est systématiquement utilisé, quels que soient les référents des arguments (humain/non humain, animé/inanimé, singulier/pluriel...)⁷⁷, sauf dans le cas particulier où l'objet est de $3^{\text{ème}}$ personne générique humaine sur lequel nous reviendrons plus tard en III-3. La présence de *o-* implique l'absence de toute personne de l'interlocution dans la situation décrite par l'énoncé.

⁷⁷ Dans ces exemples, le caractère singulier/pluriel, masculin/féminin, animé/inanimé... des participants est parfois donné par le contexte (textuel).

- (139) patu-pope **o-ijnuŋ.**
 marmite-dans 3.I-mettre
 Elle les met dans la marmite (les patates douces).
- (140) alakapusa-uhu **o-mõdul-oŋ** bal.
 fusil-grand 3.I-envoyer balle
 Les fusils tiraient des balles.
- (141) **o-mõ-gaʔu.**
 3.I-CAUS-boire.du.cachiri
 Ils la font boire.
- (142) **o-pelo-pelog** eiba Ø-owa.
 3.I-RED-lécher animal 3.II-visage
 Le chien lui lèche le visage.
- (143) eiba **o-itu-ĩtun** wila-k^wal.
 animal 3.I-RED-sentir arbre-trou
 Le chien renifle le trou de l'arbre.

Cette configuration 3 → 3 sort du cadre de la hiérarchie des personnes 1.2 → 3, mais elle révèle une autre hiérarchie : A > P (idée déjà exposée dans Couchili, Maurel et Queixalós, 2002). En effet, le préfixe *o-* qui appartient à la série I (qui marque les S et les A) peut être décrit ici comme une marque de A. On peut alors dire que dans la configuration 3 → 3, c'est A qui est marqué. Indirectement, la présence de *o-* indique également que P est de 3^{ème} personne, car si P était une personne de l'interlocution, le jeu de la hiérarchie des personnes conduirait à ce qu'il soit codé à la place de A. Une des conséquences de ces deux hiérarchies est que le préfixe *i-* de 3^{ème} personne de série II n'apparaît jamais sur les verbes indépendants. Il n'est utilisé que dans des groupes nominaux (comme possessif) ou postpositionnels, ainsi que sur les gérondifs.

L'émérillon connaît donc une hiérarchie claire 1/2 > 3 dans la configuration où une personne de l'interlocution est impliquée, et une hiérarchie des rôles sémantiques A > P quand A et P sont tous deux de 3^{ème} personne. Nous allons dans

la section suivante nous attacher au problème des hiérarchies dans les cas de configurations "locales".

III- 2. 3. Configurations locales

Heath (1998) consacre un article à l'indexation des combinaisons de 1^{ère} et 2^{ème} personne ($1 \leftrightarrow 2$) dans les langues amérindiennes. Il note que ces combinaisons sont normalement opaques et irrégulières : la seule généralisation qu'il peut faire est que les langues évitent les combinaisons transparentes $1 \leftrightarrow 2$. Il formule même précisément douze stratégies que les langues utilisent pour s'éloigner de la transparence maximale. Les langues tupi-guarani (émérillon inclus) ne font que refléter le caractère "brouillé" des marques $1 \leftrightarrow 2$ dans les langues amérindiennes (cf. par exemple Montserrat et Soares, 1983). Ainsi, en émérillon, alors que le respect de la hiérarchie est clair et systématique quand une seule des personnes de l'interlocution est présente, l'indexation se révèle fonctionner différemment quand les deux personnes de l'interlocution sont impliquées (configuration dite "locale"), ce qui rend ardue la hiérarchisation des deux premières personnes. L'article de Heath nous confirme que cette différence de traitement est fréquente dans les langues où une hiérarchie des personnes opère. Un autre exemple est donné par Gildea (1994) dans le système inverse du Caribe du Surinam. DeLancey (non daté) expose la nature déictique des systèmes hiérarchique et inverse et explique ainsi le fait qu'ils marquent le statut spécial des personnes de l'interlocution vis-à-vis des autres participants.

Un autre point intéressant est que, malgré la différence de traitement des configurations locales, la plupart des langues tupi-guarani sont décrites comme possédant une hiérarchie des personnes $1 > 2 > 3$ (voir par exemple Jensen 1990). Toutefois, les données émérillon montrent un fonctionnement différent des autres langues tupi-guarani, et n'indiquent clairement que $1/2 > 3$.

A partir de ces constats, les configurations "locales" en émérillon sont maintenant examinées, tout d'abord dans le sens 2 → 1, puis dans le sens 1 → 2, à la fois en relation avec ce qui se passe dans le reste de la famille et en relation avec les stratégies de "brouillage" des combinaisons 1 ↔ 2 formulées par Heath. La section III- 3 conclut sur la hiérarchisation qui guide le système d'indexation des personnes en émérillon.

III- 2. 3. 1. Les formes locales pour 2 → 1

La configuration 2 → 1 est marquée en émérillon de manière opaque, et différemment des autres langues tupi-guarani. Nous tentons cependant d'expliquer les formes synchroniques (notamment par des hypothèses diachroniques).

Le tableau 2 donne les quatre combinaisons possibles où une 2^{ème} personne singulier ou pluriel agit sur une 1^{ère} personne singulier ou exclusive.

A	P		
2 ^{ème} SG	1 ^{ère} SG	<i>ele-...</i> 2SG.I-...	<i>elep</i> 2SG
2 ^{ème} SG	1 ^{ère} EXCL	<i>ele-...</i> 2SG.I-....	<i>olone-kom ~ olep ~ olelep</i> PRO1EXCL-PL
2 ^{ème} PL	1 ^{ère} SG	<i>pe-...</i> 2PL.I-...	<i>pep (~elep)</i> 2PL
2 ^{ème} PL	1 ^{ère} EXCL	<i>pe-...</i> 2PL.I-...	<i>olone-kom ~ olelep</i> PRO1EXCL-PL

Tableau 13 : Combinaisons 2 → 1 en émérillon

(144) **ele-laho** **elep.** 2SG → 1SG
2SG.I-emmener 2SG
Tu m'as emmené.

(145) **pe-kual** **olelep.** 2PL → 1EXCL, exemple élicité
2PL.I-trouver 1EXCL
Vous nous avez trouvé.

Quand on a 2 → 1, A est marqué par un préfixe de série I, et une autre marque de personne est postposée au verbe. La logique laisserait penser que dans ce marquage apparemment "analytique", la deuxième marque va référer à P, ce qui

n'est en réalité pas toujours le cas. En effet, *olonekom* est le pronom de 1^{ère} personne exclusive suivi de la marque de pluriel. Quant aux pronoms *eleŋ*, *oleŋ/oleleŋ* et *peŋ*, ils peuvent avoir comme origine les pronoms indépendants du proto-tupi-guarani † *eré* (2SG), *oré* (1EXCL) et *pe...ẽ* (2PL)⁷⁸. Reste à savoir d'où vient le /ŋ/ final, peut-être de la nasalisation finale de *pe...ẽ*⁷⁹ ensuite étendu par analogie aux autres formes. Donc si *olonekom*, *oleŋ* et *oleleŋ* réfèrent clairement à P 1^{ère} personne exclusive, les formes *eleŋ* et *peŋ* ne réfèrent pas directement à P : ces marques de 2^{ème} personne singulier et pluriel indiquent indirectement que P est une 1^{ère} personne singulier. On a donc un double marquage de 2^{ème} personne qui signifie 2 → 1. La présence de ces formes postposées est cependant nécessaire pour distinguer 2 → 1SG de 2 → 3, comme en témoignent les exemples suivants avec *eleŋ*.

- (146) *ele-nupã*. 2 → 3
 2SG.I-frapper
 Tu le frappes.
- (147) **ele-nupã** **eleŋ**. 2 → 1SG
 2SG.I-frapper PRO2SG
 Tu m'a frappé.

Nous pouvons avancer une hypothèse pour cet usage apparemment bizarre, qui prend en compte un état de langue plus ancien (nous ferons référence ici au proto-tupi-guarani). Les reconstructions proposent que la hiérarchie soit 1 > 2 > 3 en proto-tupi-guarani. Dans la configuration 2 → 1, la 1^{ère} personne (plus haute sur la hiérarchie) est marquée sur le verbe (avec la série II car c'est un P). A est alors exprimé par un pronom spécial après le verbe.

⁷⁸ Ces reconstructions sont tirées de Jensen (1998a, p. 498).

⁷⁹ A moins que cette nasale soit le résultat de la grammaticalisation d'un des clitiques progressifs dont la forme est *-ŋ* et qui aurait pu se fixer sur ces marques de personne. Mais il semble envisageable que cette nasale soit liée aux marques de 2^{ème} personne de manière plus ancienne. Les données de la langue tupi Sateré-Mawé (Franceschini 2002) montrent l'existence d'une forme de pronom libre *en*, utilisée pour marquer un A de 2^{ème} personne singulier, quand P est de première personne.

III- 2. 3. 2. Les formes locales pour 1 → 2

Dans cette configuration, une autre stratégie est utilisée : on a seulement un indice sur le verbe. Les deux formes possibles sont *olo-* (je/nous → tu) et *apolo-* (je/nous → vous).

olo- indique qu'une 1^{ère} personne agit sur une 2^{ème} personne singulier.

- (150) "e-weɕu t-**olo**-ʔu".
 2SG.IMP-descendre BUT-A1/P2SG-manger
 "Descends que je te mange !"

apolo- indique qu'une 1^{ère} personne agit sur une 2^{ème} personne pluriel.

- (151) **apolo**-ekal. exemple élicité
 A1/P2PL-chercher
 Nous vous avons cherché.

Les formes 1 → 2 sont rares dans notre corpus naturel, et les résultats des élicitations ne sont pas aussi clairs que la présentation ci-dessus le laisserait penser. Pour 1SG → 2SG, on trouve toujours *olo-*, et pour 1SG/EXCL → 2PL, on trouve le plus souvent *apolo-*. Mais pour 1EXCL → 2SG, *olo-* et *apolo-* sont tous deux attestés par les informateurs. Il est difficile de savoir si cette variation est réelle ou due aux difficultés d'élicitation de telles formes.

Jensen (1998a, p. 498) propose de reconstruire ces deux formes † *oro-* A1/P2SG, † *opo-* en proto-tupi-guarani. En linguistique tupi-guarani, on les appelle traditionnellement des morphèmes "porte-manteaux", du fait qu'elles réfèrent à la fois à A et à P en un seul morphème, une analyse que nous allons remettre en question ici, ainsi que la reconstruction † *opo-*. Il est intéressant de noter que ces formes connaissent une variation importante dans les langues tupi-guarani. D'ailleurs, quand Montserrat et Soares (1983) présentent cinq groupes de langues tupi-guarani pour lesquelles la hiérarchie des personnes fonctionne seulement de

manière partielle, dans les cinq groupes, la rupture de la hiérarchie concerne précisément le cas 1 → 2.

Revenons donc à notre objet d'analyse : les formes *olo-* et *apolo-* de l'émérillon. Le morphème *olo-* de l'émérillon est de manière assez évidente un réflexe du † *oro* proto-tupi-guarani. Par contre, son analyse en tant que morphème porte-manteau ne paraît se baser que sur un seul fait : que cette marque révèle les personnes à la fois de A et de P, fait qui se vérifie en réalité pour toutes les formes que nous avons exposées jusque là ! Par conséquent, une analyse en tant que marque de A ou P nous paraît plus pertinente. Or, *olo-* est homonyme avec la marque de série I de première personne exclusive (c).

(152) **olo-nupã**

- a) Je te frappe.
- b) Nous te frappons.
- c) Nous (le) frappons.

Cette homonymie nous fait tendre vers l'hypothèse d'une interprétation d'*olo-* comme marque de A et non de P. C'est ensuite le contexte qui permet de résoudre l'ambiguïté entre les interprétations a, b et c.

Si nous suivons cette hypothèse, il nous faut expliquer qu'en (152), quand la configuration est 1 → 2SG, cette marque explicitement plurielle rend aussi bien un A singulier que pluriel. Il existe un fait général lié à la communication entre deux personnes tel que la situation 1 → 2 crée une confrontation entre le locuteur et son allocataire qui est en position basse. Les langues utilisent alors souvent des stratégies telles que la pluralisation ou la substitution d'une personne pour une autre pour atténuer cette confrontation, comme par exemple en français quand le locuteur s'adresse à un allocataire unique en utilisant une forme de 2^{ème} personne plurielle ("vous de politesse") (Kerbrat-Orecchioni, 1992, dans la lignée de Brown et Levinson, 1987). Une analyse possible de la forme *olo-* dans la configuration locale

est que son caractère pluriel soit une manière de rabaisser ou "délayer" la 1^{ère} personne A, et par conséquent d'établir une relation moins "menaçante" pour l'allocutaire-P⁸¹. C'est exactement le point de vue qu'expose Heath (1998) de manière générale dans son article sur les combinaisons 1 ↔ 2 dans les langues amérindiennes : il compare le fait que les langues évitent les combinaisons 1 ↔ 2 transparentes aux restrictions d'utilisation des 1^{ère} et 2^{ème} personnes singulier dans le discours poli. Ainsi, parmi les stratégies d'évitement de la transparence, il donne les suivantes qui s'appliquent aux cas de *olo-* marque de A 1^{ère} personne exclusive utilisée pour exprimer 1SG/PL → 2SG.

- stratégie 10 : le sujet et l'objet sont en compétition pour une seule position
- stratégie 4 : neutralisation du nombre, incluant parfois l'usage de "pluriel" pour un "singulier" sémantique.

Ainsi *olo-* (et ses contreparties dans les autres langues tupi-guarani) illustre le manque de transparence dans les formes locales dû à des restrictions d'ordre socio-pragmatique.

En ce qui concerne l'autre forme *apolo*, nous réfutons non seulement son analyse comme morphème porte-manteau, mais aussi la reconstruction proto-tupi-guarani † *opo* proposée par Jensen (1998a, p. 498). Nous proposons plutôt de segmenter cette forme en deux indices successifs : *a-* (1^{ère} personne singulier de série I, référant à A) et *polo* (indice d'objet générique humain de série II). Il s'agirait de la seule combinaison possible de deux préfixes de personne sur le même verbe en émérillon.

La présence de l'objet générique humain *polo-* dans la configuration 1 → 2PL peut aussi être expliquée en terme de confrontation entre les deux personnes de

⁸¹ Cette analyse a été proposée pour des faits similaires en Quechua par Queixalós et García Rivera (à paraître), avec référence aux faits tupi-guarani.

l'interlocution : la substitution d'une marque de personne indéterminée pour une deuxième personne permet de créer une distanciation, qui adoucit la confrontation (cf. Brown & Levinson, 1987). Cette "substitution" marque une deuxième étape dans la grammaticalisation de *polo* : d'indice de personne indéterminée et générale de série II, *polo* devient indice d'objet de 2^{ème} personne pluriel dans la configuration 1 → 2PL.

La même hypothèse de segmentation a en fait été récemment développée par Cabral (2001a), à partir de quelques langues tupi-guarani⁸². Cet auteur propose en conséquence que la proto-forme de 1SG → 2PL soit † *a-poro* (et non † *opo*). † *a-poro* aurait ensuite été réduit phonologiquement dans certaines langues, ce qui explique la reconstruction † *opo-* de Jensen. Notons qu'une fois de plus, l'apparent manque de transparence dans les combinaisons 1 ↔ 2 peut être expliqué par des raisons d'ordre socio-pragmatique qui poussent à éviter la combinaison de deux indices reprenant les personnes de l'interlocution.

En conclusion, l'étude des séries sur les verbes transitifs a permis de montrer que l'émérillon suit une hiérarchie des personnes seulement entre les personnes de l'interlocution et une 3^{ème} personne : il s'agit de 1/2 > 3. Cette hiérarchie et le fait qu'elle ne s'applique pas aux configurations locales sont tous deux fréquents en typologie. L'étude plus spécifique des configurations locales nous permet d'affirmer que l'expression de 2 → 1 a évolué entre le proto-tupi-guarani et l'émérillon, et que l'analyse des formes de 1 → 2 comme formes porte-manteaux doit être rejetée à condition d'accepter des glissements fonctionnels de marques de personne selon des modèles socio-pragmatiques, et une nouvelle reconstruction de la proto-forme d'*apolo-*.

⁸² tembé, asuriní du Xingu, émérillon, jo'é et wayampi.

Maintenant que nous avons vu en détail le système d'accord sur le verbe, nous pouvons conclure sur les hiérarchies en jeu dans l'indexation des personnes en émérillon.

III- 3. La hiérarchisation des personnes et des rôles sémantiques

Après avoir récapitulé brièvement en quoi l'émérillon remettait en cause la hiérarchie $1 > 2$, nous proposons d'expliquer les formes locales à l'aide d'une autre hiérarchie, celle des rôles sémantiques. Ensuite, les deux hiérarchies seront ordonnées pour former un système cohérent.

Nous avons vu que la description des formes locales (celles qui impliquent les deux personnes de l'interlocution) est problématique en émérillon. Nous avons également observé que les formes pour $2 \rightarrow 1$ de l'émérillon divergent de celles des autres langues de la famille, et ne respectent pas la hiérarchie censée être $1 > 2$. Il nous faut donc remettre en question la hiérarchie $1 > 2 > 3$ attribuée à la famille tupi-guarani, du moins en ce qui concerne l'émérillon. Comme il est typologiquement commun que, dans une langue présentant une hiérarchie des personnes, le classement des personnes de l'interlocution vis-à-vis de la 3^{ème} personne soit net, mais que le classement des deux personnes de l'interlocution entre elles soit plus flou, cela peut expliquer que l'émérillon ait réorganisé la hiérarchie du proto-tupi-guarani en ce qui concerne $1 > 2$.

Récapitulons les analyses précédentes :

- Pour $2 \rightarrow 1$, A est marqué en priorité par la série I. L'autre marque, un pronom placé après le verbe, réfère à P de manière soit directe, soit indirecte quand elle est un vestige d'une marque de A du proto-tupi-guarani.
- Pour $1 \rightarrow 2SG$, nous avons proposé que *olo-* soit analysé comme une marque de A. Rien ne marque alors explicitement P.

- Pour $1 \rightarrow 2\text{PL}$, A est marqué par la série I, alors que P est marqué par un objet générique. C'est donc le seul cas en émérillon où l'on trouve deux positions d'indices de personne sur le verbe.

Notre conclusion (qui était déjà celle de Couchili, Maurel et Queixalós, 2002) est qu'aucune hiérarchie entre les deux personnes de l'interlocution ne ressort de manière évidente, mais que la hiérarchie opérante pour toutes ces situations "locales" est une hiérarchie des rôles sémantiques $A > P$. En effet, dans tous les cas, c'est A qui a la priorité sur P : A occupe toujours la position pré-verbale réservée aux marques de personne, alors que P est rejeté après le verbe ($2 \rightarrow 1$), incorporé entre le préfixe de personne de série I et le radical du verbe (*polo*), ou simplement déduit du contexte (*olo*). Un point intéressant est qu'il a déjà été fait appel à cette hiérarchie des rôles sémantiques pour la configuration impliquant deux 3^{ème} personnes. La hiérarchie des rôles sémantiques fonctionne dans les cas où la hiérarchie des personnes ne s'applique pas, à une exception près.

Il est en effet important de noter que la hiérarchie $A > P$ est inactive dans un cas particulier où pourtant une 3^{ème} personne agit sur une autre 3^{ème} personne. En effet, quand l'objet d'un verbe transitif est de 3^{ème} personne générique, c'est lui qui est marqué par *polo-*, comme en (153) et (154).

(153) **polo-**ʔu-pa boɟ.
 INDET.II-manger-COMPL serpent
 Le serpent a mangé tout le monde.

(154) ...i-moda-mãʔẽ-kom **polo-pihig** zo-kalakuli-kom o-ikiɟ...
 3.II-vol-REL-PL INDET.II-attraper INDET.II-argent-PL 3.I-prendre
 les voleurs attrapent les gens et prennent leur argent

Cette règle va à l'encontre de la hiérarchie de rôles sémantiques, car A devrait être marqué prioritairement dans cette configuration où $3 \rightarrow 3$. Ceci est d'autant plus surprenant que même sur les hiérarchies plus détaillées que l'on trouve dans la littérature, les 3^{èmes} personnes génériques sont encore plus basses que les autres

(cf. par exemple Payne 1997). Cette exception à la hiérarchie des rôles sémantiques doit probablement s'expliquer par le caractère "récent" de *polo-* parmi les indices de personne. Dans la section I-3.1., il a été vu que *polo-* est le résultat d'une grammaticalisation d'un nom générique incorporé en indice de personne. L'origine particulière de la structure impliquant *polo-* comme seul indice de personne semble expliquer la non conformité de cette structure aux hiérarchies gouvernant le système d'indexation de l'émérillon.

Au final, l'indexation des personnes en émérillon connaît deux hiérarchies différentes : une hiérarchie des personnes $1/2 > 3$, et une hiérarchie de rôles sémantiques $A > P$. Quand doit-on appliquer l'une ou l'autre ? Couchili, Maurel et Queixalós (2002) proposent la "hiérarchie des hiérarchies" personnes $>$ rôles, pour les raisons suivantes :

"Il y aura donc un seul préfixe, celui du participant le plus haut placé dans la hiérarchie personnelle, *indépendamment de son rôle sémantique*. Cette dernière précision fonde la hiérarchie personne $>$ rôle. Quant à la hiérarchie des rôles, elle se manifeste quand la question de la préséance de personne ne se pose plus, à savoir avec deux participants de 3°." (Couchili, Maurel et Queixalós, 2002, p.178)

Nous adhérons à cette analyse, mais l'étendons également à la configuration locale : notre hiérarchie des personnes se limitant à $1/2 > 3$, les cas impliquant les deux personnes de l'interlocution sortent du cadre de la hiérarchie des personnes et obéissent alors à la hiérarchie des rôles sémantiques. On peut donc résumer le système d'accord hiérarchique de l'émérillon comme suit : l'indexation des personnes suit la hiérarchie des personnes $1/2 > 3$. Quand celle-ci ne s'applique pas ($3 \rightarrow 3$, $1 \rightarrow 2$, $2 \rightarrow 1$), la hiérarchie des rôles sémantiques $A > P$ rentre en jeu.

Le fait qu'une hiérarchie des personnes régit le système d'accord des langues tupi-guarani a conduit certains auteurs à interpréter ce système comme un système inverse. Nous débattons maintenant de cette hypothèse.

III- 4. La hiérarchie des personnes en émérillon et en tupi-guarani : un système inverse ?

Dans cette section, nous discutons de la possibilité d'analyser le système d'accord de l'émérillon comme un système "inverse". Nous commençons par définir la notion d'inverse en typologie. Ensuite, nous introduisons l'hypothèse de D. Payne (1994) qui voit dans les langues tupi-guarani des systèmes inverses. Enfin, nous donnons notre avis motivé par nos données émérillon, et rejettons cette analyse, pour préférer le terme de "système hiérarchique".

La définition de l'inverse se base sur l'idée d'une "hiérarchie de topicalité inhérente" $1 > 2 > 3$, la topicalité d'un participant étant le résultat d'un croisement de paramètres tels que son agentivité, son caractère animé et individué, son importance locale ou globale et son caractère prédictible (Givón, 1994b). Les constructions "inverses" marquent le fait que l'agent est "moins topique" que le patient (Payne T. 1997). Ainsi, si dans une proposition transitive, la 1^{ère} personne agit sur la 2^{nde} ou la 3^{ème}, ou si la 2^{ème} agit sur la 3^{ème}, l'action se déroule dans la direction naturelle (A est plus topique que P) : la situation est appelée "directe". Si le cours de l'action est inversé (P est plus topique que A), on dira alors que la situation est "inverse". La notion d'"inverse" a été mise à jour pour les langues algonquines, qui constituent maintenant le prototype de l'inverse. Voici des exemples d'une langue algonquine, l'ojibwa, repris de Rhodes (1976, p.85-86, cité dans Mithun 1999, p. 223) :

(155) n-bi:n-a > nbina:
 1-amener-DIR
 Je l'amène.

(156) n-bi:n-igw > nbi:nig
 1-amener-INV
 Il/elle m'amène.

Une langue à inverse grammatical est une langue qui exprime le caractère inverse d'une situation par une construction morphosyntaxique qui reste transitive, contrairement à la construction passive.

Mais Givón (1994b) ou Klaiman (1991)⁸³ favorisent une définition fonctionnelle de l'inverse. Il y a inverse dès que P est plus topique que A mais que A est encore topique, et ce, quelle que soit la construction utilisée. Ainsi, Thomas Payne (1997), dans sa typologie des inverses, propose un sous-type appelé "accord verbal spécifique aux situations inverses" ("special verb agreement for inverse situations") et l'illustre par le wayampi, une langue très proche de l'émérillon. D'un autre côté, Heath (1998) s'élève contre l'extension récente du terme "inverse" à des formes qui ne connaissent pas de morphème inverse, extension qui détruit l'utilité d'un tel terme linguistique. En effet, de manière plus traditionnelle,

"Inverse marking is a feature of languages that have a rich agreement system, typically languages where all the arguments are marked in the verbal inflection." (Jelinek 1990, p. 227)

Ainsi, derrière la notion d'inverse, on retrouve la discussion théorique qui consiste à débattre si l'on doit partir de la forme (approche syntaxique) ou de la fonction (approche fonctionnelle).

Passons à l'hypothèse de Doris Payne (1994), qui analyse les langues tupi-guarani comme des langues à système d'accord inverse. Cet auteur considère que l'usage des séries I ou II constitue un système d'inverse dans ces langues. Quand $1 \rightarrow 2/3$ ou $2 \rightarrow 3$, la situation est directe et la série I est utilisée pour marquer A. A l'opposé, quand $3 \rightarrow 1/2$ ou $2 \rightarrow 1$, on a affaire à une situation

inverse et la série II marque P. Elle propose ensuite qu'un préfixe, le préfixe *r-*, soit analysé comme la marque d'inverse des langues tupi-guarani. Ce préfixe, souvent appelé préfixe relationnel⁸⁴, s'insère entre certains racines verbales, nominales ou postpositionnelles et leur complément. Mais il nous paraît être une mauvaise marque d'inverse, vu qu'il n'apparaît qu'avec une classe de verbes (lexicalement définie), mais par contre apparaît sur d'autres parties du discours.

Si nous continuons cependant à suivre l'hypothèse de l'inverse, nous aboutissons à appeler "inverse" un système très éloigné du prototype que constituent les langues algonquines : en émérillon, aucun morphème ne peut être interprété comme une marque d'inverse, et le système est limité aux cas où une et seulement une des personnes de l'interlocution est impliquée. En effet, pour pouvoir interpréter le système d'accord émérillon comme un système inverse, il faudrait accepter de décrire la simple distribution des marques de personne en deux séries distinctes comme un système inverse (sans oublier le fait que certaines marques n'appartiennent pas de manière claire à l'une ou l'autre des séries). Cette supposition est possible dans le cadre d'une définition fonctionnelle de l'inverse, telle que celle de Givón (1994b).

Tout en admettant que l'on peut identifier une "fonction" inverse dans le système d'accord émérillon, cela ne nous paraît pas suffisant pour parler d'un système inverse en émérillon. Nous n'avons pas réellement affaire à un "système", tout d'abord parce que deux configurations sur trois y échappent (formes locales et 3 → 3), puis parce qu'il ne possède aucun marqueur spécifique. Selon nous, l'analyse du système d'accord émérillon comme un système inverse n'est pas nécessaire ni enrichissante pour comprendre nos données : parler d'un système hiérarchique suffit. Nous n'avons par exemple pas de marques de 3^{ème} personne "obviative" ou "proximale" (une distinction interne à la 3^{ème} personne, présente dans les langues

⁸³ cf. par exemple l'analyse que ce dernier auteur fait de l'Arizona Tewa.

⁸⁴ Le préfixe relationnel de l'émérillon fera l'objet de la partie II du chapitre 4.

algonquines) qu'un système inverse pourrait prendre sous sa coupe, étant donné que quand deux 3^{ème} personnes sont impliquées, aucune alternance comparable au direct/inverse n'est possible. Après avoir donné nos raisons pour écarter l'hypothèse de l'inverse, nous concluons brièvement sur le type de système d'accord de l'émérillon.

Nous avons écarté dès le début de cette partie III deux types de systèmes d'accord : le système absolutif reconstruit par Jensen (1990) pour les verbes dépendantes, car il n'a plus lieu en émérillon dans les propositions dépendantes, et le système actif/statif car l'analyse des mots "descriptifs" comme des noms et non des verbes en émérillon détruit la notion d'intransitivité scindée, qui est le fondement d'un système actif/statif (III-1). Nous avons également exprimé notre réticence à l'égard de l'appellation "système inverse" pour une langue comme l'émérillon (III-5). Le système d'accord de l'émérillon se rapproche à certains égards des systèmes nominatif-accusatif. En effet, S et A sont tous deux marqués par la série I, que l'on pourrait interpréter comme une marque nominative, et P seul est marqué par la série II, qui serait alors une marque d'accusatif. Mais un élément manque à cette description : seulement un argument est en réalité marqué sur les verbes transitifs, le choix se faisant selon une hiérarchie des personnes et une hiérarchie de rôles sémantiques. Cette notion de hiérarchie nous semble être un fait primordial et organisateur du système d'accord émérillon, que nous classerons donc parmi les systèmes d'accord hiérarchique. Dans son exposé des différents types d'"alignement", Nichols (1992) reconnaît les systèmes d'accord hiérarchiques comme constituant un des grands types d'"alignement" dans les langues du monde. Ils incluent les systèmes purement inverses, et leur particularité est définie ainsi :

“Access to inflectional slots for subject and/or object is based on person, number, and/or animacy rather than (or no less than) on syntactic relations” (Nichols, 1992, p.66).

C'est donc sur la conclusion que l'émérillon connaît un système d'accord hiérarchique que nous terminons ce chapitre.

Ce chapitre était consacré aux indices de personne de l'émérillon, qui sont des morphèmes fréquents, indispensables et poly-fonctionnels de la langue. Nous avons vu en I les formes des indices de personne (I-1), avec leur sens (I-2) et leur origine (I-3). Nous avons ensuite donné la distribution détaillée de chaque série d'indices (II). Enfin, nous avons discuté du système d'accord en personnes sur le verbe en émérillon, qui s'organise de manière hiérarchique (III et IV).

Le chapitre suivant est consacré à deux autres morphèmes typiquement tupi-guarani : le *-a* et le relationnel.

Chapitre 4 : Le suffixe référentiant *-a* et le préfixe relationnel *l-*

Dans ce chapitre, nous présentons deux morphèmes : le suffixe référentiant *-a* et le préfixe relationnel *l-*. Ils sont tous deux non seulement typiques et répandus dans la famille tupi-guarani, mais ils sont aussi largement utilisés en émérillon, dans différents types de constituants. En raccourci, le suffixe *-a* indique qu'un constituant ne joue pas de rôle prédicatif, donc qu'il est dépendant d'un prédicat ou d'une autre tête de syntagme. Quant au relationnel, il apparaît dans différents contextes pour lier une tête de syntagme et un complément que celle-ci régit fortement.

I- Le suffixe *-a*

Le suffixe *-a* est présent dans un grand nombre de langues tupi-guarani, avec une distribution variable quant aux contextes phonologiques et syntaxiques (Cabral 2001b). Il est souvent considéré comme une marque de cas, qui prend différents noms selon les auteurs, sans que cela n'implique des analyses vraiment différentes : cas nucléaire (Seki 2000), cas nominal (Rodrigues 1981 ; Jensen 1989), cas argumentatif (Rodrigues 2001). Dans les exemples émérillon, nous l'avons tout simplement glosé *-a*.

Dans cette partie, nous exposons d'abord en I-1 les contextes morphosyntaxiques et phonologiques dans lesquels on trouve le *-a* en émérillon. Nous reprenons ensuite en I-2 la proposition de Queixalós (2001c) de l'analyser comme un suffixe référentiant, proposition que nous adoptons dans cette thèse. Enfin, nous présentons en I-3. le fait que *-a* ait une variante *aʔe*, fait qui nous semble confirmer et enrichir l'analyse de Queixalós de manière intéressante.

I- 1. Distribution du suffixe *-a*

Dans cette partie, nous allons d'abord présenter la distribution du suffixe *-a* en émérillon, en termes morphosyntaxiques et phonologiques.

Ce suffixe apparaît dans de nombreux contextes morphosyntaxiques, notamment à l'intérieur du syntagme génitival entre deux éléments nominaux (157), entre un groupe nominal et une postposition (158), entre une proposition (entre crochets) et un subordonnant (159) et entre un constituant et certains clitiques (comme la particule contrastive *-nẽ* en (160)).

(157) apam-**a**-baʔek^wəl-**a**-kom aipo za-pulu.
 étranger-a-coutume-a-PL maintenant INDET.I-utiliser
 Maintenant on utilise les coutumes des étrangers.

(158) wane pita(ŋ)-kom-**a**-pe.
 bon enfant-PL-a-pour
 C'est bon pour les enfants.

(159) [pitaŋ o-ʔal]-**a**-nawe, eap̄ o-kakuwa ikiʔi.
 enfant 3.I-tomber-quand vite 3.I-grandir désormais
 Une fois que l'enfant est né, il grandit alors vite.

(160) apam-**a**-nẽ nōde-apisi-tanẽ.
 étranger-a-CONTR 2.II-massacrer-DESID
 Les étrangers voulaient nous massacrer.

Dans l'exemple (157), on voit que le suffixe *-a* s'intercale aussi entre un nominal et le suffixe de pluriel. Nous verrons dans le prochain chapitre en III-2. que Queixalós rend compte de ce phénomène en considérant que le pluriel *-kom* est en fait un nom "ensemble, groupe", qui constituerait donc un syntagme génitival avec le nominal qu'il "pluralise".

La non réalisation phonologique du suffixe *-a* dans certains contextes (en (162) par opposition à (161)) n'est que le résultat d'une règle morphophonologique qui efface la première voyelle d'un suffixe après un morphème à finale vocalique, règle d'ailleurs valable pour d'autres morphèmes, tel le morphème de pluriel *-(o)ŋ* en (163) et (164).

- (161) lekol-**a**-zal
 école-a-maître
 le maître d'école, l'instituteur
- (162) teko-**Ø**-awu
 Emérillon-a-langue
 la langue des Emérillons, l'émérillon
- (163) o-paʔam-**oŋ**.
 3.I-être.debout-PL.S
 Ils sont debout.
- (164) wilo t-**apiɕ** o-baʔe-**ŋ**.
 feuille.de.palmier NSP-maison 3.I-faire-PL.S
 Ils ont fait une maison de feuille de palmier.

Nous aborderons au chapitre 9, III-1, la question de sa possible appartenance à un paradigme de cas. Après avoir vu la distribution de ce morphème, nous passons à son analyse fonctionnelle.

I- 2. Le suffixe *-a* : un morphème translationnel ou référentiant

Queixalós (2001c) a proposé une hypothèse qui rend compte des différents contextes d'apparition du morphème *-a* non seulement en émérillon, mais aussi dans l'ensemble des langues tupi-guarani. En effet, dans d'autres langues de la famille, le suffixe *-a* connaît une distribution plus large. En plus des contextes donnés ci-dessus pour l'émérillon, le suffixe *-a* peut se trouver également sur les sujets et les objets, sur les prédicats équatifs, ainsi que sur les verbes pour les "nominaliser"⁸⁵.

Pour résumer, cet auteur suppose un état initial des langues tupi-guarani où toutes les racines lexicales ont vocation prédicative, et où le suffixe *-a* permet de construire de la référence à partir d'une racine qui ne réfère pas. Ce *-a* marque indirectement la translation d'une racine capable de prédiquer à une fonction subordonnée (telle que sujet, objet, circonstant, complément du nom). Mais comme il se trouve aussi sur un prédicat dans une relation équative, il n'a pas le seul rôle

⁸⁵ Pour une description détaillée des nominalisations d'action, voir le chapitre 6, V-2.1.

de translatif, mais donne plutôt au constituant sur lequel il s'affixe la capacité de référer. A partir de cet état initial, les diverses langues tupi-guarani ont connu des types d'érosion divers du morphème *-a*. En émérillon, la récession du *-a* a touché sur le plan syntaxique les fonctions sujet et objet (sauf quand ils sont suivis d'une particule), et sur le plan phonologique le contexte post-vocalique. Ce *-a* reste utilisé dans la construction génitive et sur les compléments du verbe autres que sujet et objet (objets de postposition, subordonnées), dans une position post-consonantique.

Comme la distribution du suffixe *-a* émérillon n'inclut pas l'affixation sur un prédicat de type équatif (cf. I-1), l'analyse du *-a* comme "translationnel" est suffisante pour cette langue. Elle nous semble rendre compte de manière adéquate de l'utilisation synchronique de ce morphème en émérillon, comme de sa diachronie et de la variété de sa présence dans les diverses langues tupi-guarani.

I- 3. La variante *a?e*

Ayant adopté comme analyse du *-a* l'hypothèse proposée par Queixalós (2001c), nous ajoutons une information intéressante qui soutient cette analyse : le suffixe *-a* connaît en émérillon une variante *a?e* (qui est par ailleurs un démonstratif). Nous donnons des exemples illustratifs de cette variante que nous glosons également *-a*. Nous discutons ensuite de sa nature probablement démonstrative, et en tirons des hypothèses d'ordre diachronique.

I- 3. 1. Distribution de *a?e*

Dans les constructions génitives, on trouve souvent le morphème *a?e* à la place du *-a*.

- (165) apam-**a?e**-sipala-te nōde-l-apidɔ-am.
 étranger-a-fer-FOC 2.II-RELN-maison-TRANSL
 C'est le fer des étrangers (qu'on utilise) pour nos maisons.

Il semblerait que *a?e* et *-a* soient des variantes interchangeables⁸⁶. Ainsi, l'exemple (166) est parallèle à (167), mais avec *a?e* au lieu de *-a* :

- (166) *zai-a?e-ba?ek^wəl*
 lune-a-histoire
 l'histoire de la lune
- (167) *apam-a-ba?ek^wəl*
 étranger-a-coutume
 la coutume des étrangers

Le morphème *a?e* n'est pas seulement une variante de *-a* dans la construction génitive, mais dans d'autres contextes aussi, comme devant une postposition.

- (168) *10heures-a?e-nawe o-kel.*
 10h-a-à 3.I-dormir
 A 10h, ils sont allés dormir.
- (169) *apam-a?e-wi wane-ma?ẽ za-ikiɕ-tal.* phrase élicitée
 étranger-a-ABL être.bon-REL INDET.I-prendre-FUT
 On va prendre ce qui est bien aux étrangers.

Nous n'avons pas d'avis définitif sur la grammaticalité de *a?e* devant un subordonnant. De plus, des premiers résultats d'enquête tendraient à montrer que *a?e* n'est pas une variante acceptable de *-a* avant les clitiques tels que *-nẽ* "contrastif", *-we* "aussi".

I- 3. 2. Hypothèses diachroniques

Une hypothèse que nous faisons est que le morphème *-a*, si répandu dans la famille tupi-guarani, ait un lien avec le démonstratif *a?e*, qui semble également répandu dans la famille tupi-guarani⁸⁷. En émerillon, *a?e* est aussi une forme de

⁸⁶ Notre enquête sur la question a consisté à prendre tous les exemples de construction génitive avec *-a* ou *a?e* dont nous disposions à l'époque dans les textes et à demander si la substitution par l'autre variante était acceptable. Dans tous les cas, elle l'était.

⁸⁷ Jensen (1998a, p. 551) propose la reconstruction **a?e* "lui, celui-là, là-bas" (visible ou invisible) pour le proto-tupi-guarani.

démonstratif : c'est un déictique discursif, utilisé comme anaphorique⁸⁸. On le trouve comme pronom et comme modifieur :

- (170) fu āhā nōde-peḏḡu, **aʔe**-a-te nōde-baʔe.
 IDEO seulement 1INCL.II-souffler DEM-a-FOC 1INCL.II-faire
 Il nous a soufflés et c'est ça qui nous a créé.
- (171) **aʔe** kalug-a-pope-we o-nupā-ŋ o-nupā-ŋ.
 DEM soirée-a-dans-aussi 3.I-frapper-PL.S 3.I-frapper-PL.S
 A cette soirée-là, ils l'ont frappée (la liane).

Ce qui est intéressant, c'est que l'hypothèse d'avoir le démonstratif *aʔe* comme marqueur référentiel est tout à fait explicable fonctionnellement : un argument accompagné d'un déictique est forcément référentiel. Or, parmi les langues citées par Queixalós (2001c) comme utilisant un morphème dit "translationnel", le *nootka* et le *nahuatl* utilisent une marque de "définitude" comme translationnel. Si la présence d'un article ou d'un démonstratif auprès d'un constituant le fait référer, alors indirectement cela indique sa "translation" de constituant potentiellement prédicatif à un constituant ne prédiquant pas.

On peut donc imaginer le scénario d'émergence des constructions utilisant *-a/aʔe* en émérillon, comme par exemple la construction génitive. Le schéma suivant résume l'hypothèse de scénario :

étape 1	[apam _i], [aʔe _i baʔel ^w əl]	les étrangers, l'histoire de ceux-ci
étape 2	[apam aʔe baʔel ^w əl]	l'histoire de ceux (qui sont des) étrangers
étape 3	[apam-a-baʔek ^w əl]	l'histoire des étrangers

Tableau 14 : Scénario d'émergence des constructions génitatives en *aʔe* et *-a*

On part d'un état de la langue 1 où les modifieurs précèdent généralement la tête, et où l'on peut disloquer des éléments topicalisés en position initiale.

⁸⁸ Pour une description de *aʔe* comme pronom démonstratif, cf. Chapitre 6, III-2.3 et comme adjectif démonstratif, cf. Chapitre 7, III-3.

Imaginons une situation initiale où une suite telle que *apam aʔe baʔekʷəl* est interprétée comme "les étrangers, l'histoire de ceux-ci"⁸⁹. Dans ce morceau d'énoncé, *aʔe* est analysé comme le modifieur de *baʔekʷəl*, et renvoie à l'antécédent *apam* (ce qui est symbolisé dans notre schéma par leurs indices identiques). L'ordre canonique des modifieurs élimine l'hypothèse d'une lecture où *aʔe* modifierait *apam*. Dans une deuxième étape, on peut facilement imaginer une réanalyse en un seul syntagme de *apam aʔe baʔekʷəl* comme "l'histoire des étrangers", où *aʔe* n'a plus de sens déictique propre ("*l'histoire de ces étrangers"). Enfin, l'érosion phonologique de *aʔe* le réduit le plus souvent à la seule voyelle *-a*. Notons que cette réduction phonologique en /a/ est attestée quand *aʔe* a sa valeur de déictique discursif.

- (172) **a-l-ehe-te** ...
 DEM-RELN-à.cause.de-FOC
 C'est pour ça que ...

A partir de ce scénario, deux hypothèses sont possibles :

- L'hypothèse forte consisterait à considérer que le démonstratif tupi-guarani *aʔe* constituerait l'unique source diachronique du fameux morphème tupi-guarani *-a*.

⁸⁹ On trouve en fait ce genre de constructions dans les langues actuelles de la famille. L'exemple 971 de Seki (2000) correspond à cette structure, avec en plus, des occurrences supplémentaires du *-a*. Seki décrit cette structure comme une parataxe par juxtaposition d'un nom en fonction de modifieur génitival à gauche d'une construction dans laquelle il est repris par le démonstratif *a'e* (Seki 2000, p. 260).

(1) **kuikuro a'e-a retma-a** amoete utsu
 K. ceux-ci-N village-N loin Intens (N = cas nucléaire)
 Les kuikuro, le village de ceux-ci est très loin.

On doit simplement remarquer la similarité de fonction entre $-a_1$ et les deux réflexes de † *aʔe*, qui en fait des variantes du même morphème référentiant. La convergence de la variante a_2 de *aʔe* et du morphème $-a_1$ en une seule forme $-a$ indifférenciée s'explique par leur similarité fonctionnelle et leur homonymie. Selon ce scénario, une co-occurrence de $-a$ et de *aʔe* doit être envisagée, co-occurrence qui se serait réduit phonologiquement en *aʔe*. Deux phrases successives du corpus semblent authentifier cette analyse :

- (173) o-koal **aʔe** kito.
 3.I-trouver DEM grenouille.
 Il a trouvé la grenouille de celui-ci.
- (174) o-zal-**a** **aʔe**-le-iba o-koal.
 3.COREF-maître-a DEM-RELN-animal 3.I-trouver
 Il a trouvé la grenouille de son maître.

Dans la première phrase, *aʔe* doit être vu comme le modifieur génitival de *kito* (renvoyant à l'antécédent "il/garçon", c'est-à-dire au maître du chien). Dans la deuxième phrase, l'antécédent est repris explicitement, avec un $-a$. On a donc une co-occurrence de $-a$ et *aʔe*.

Seules des informations comparables sur d'autres langues de la famille pourraient venir consolider l'hypothèse que le morphème $-a$, si répandu dans la famille tupi-guarani, ait un lien avec le démonstratif *aʔe*, et préciser les scénarios d'évolution concevables.

Cette partie consacrée au suffixe $-a$, suffixe tupi-guarani typique et omniprésent dans ces langues, a présenté d'abord sa distribution en émerillon, puis une analyse récente comme suffixe "translationnel" et "référentiant" proposée par Queixalós et enfin, à partir d'une possible variante *aʔe* de $-a$, quelques idées nouvelles sur l'origine de ce suffixe.

II- Le préfixe relationnel *l-*

Le morphème que nous traitons maintenant apparaît dans toutes les structures émérillon où peut apparaître la série II d'indices de personne (cf. Chapitre 3, II). Ce morphème relationnel "relie" donc une tête de syntagme (verbe, postposition ou nom) et son complément, que celui-ci soit justement réalisé par un indice de série II ou par un nominal explicite. Quand le complément est un nominal et non un indice, le relationnel apparaît en plus du suffixe *-a* décrit ci-dessus.

Ce morphème a été reconstruit comme † *r-* pour le proto-tupi-guarani⁹⁰ (Cabral 2000b) et son rôle de "lien" lui a valu l'appellation de "relationnel" (Rodrigues 1981). Il se retrouve en effet dans les mêmes contextes dans toute la famille tupi-guarani, mais aussi dans d'autres familles de la région comme les familles caribes et macro-jê (Cabral 2000b). L'analyse de sa fonction et de son rapport aux indices de personne prête à la controverse.

Dans cette partie, nous commençons par décrire le comportement du morphème relationnel en émérillon, comportement tout à fait représentatif de celui qu'il a dans l'ensemble des langues tupi-guarani. En II-1, nous précisons la distribution de ce morphème en émérillon, en II-2 les différentes formes qu'il prend, et en II-3 quelles racines (lexicalement déterminées) demandent le relationnel parmi les noms, les postpositions et les verbes. Ensuite, nous abordons une discussion valable pour toute la famille tupi-guarani : nous traitons de la fonction de ce morphème (II-4) et de sa position par rapport au système d'indexation des personnes (II-5).

II- 1. La distribution du relationnel

Le relationnel se trouve dans les contextes où la série II peut se trouver : il peut relier un nom et son possesseur, une postposition et son objet, un verbe et son objet. Dans les trois cas, il s'insère entre la tête de syntagme et son complément le

⁹⁰ Ce *r-* correspond à notre *l-* en émérillon.

plus fortement régi (qui précède la tête), que celui-ci soit réalisé par un indice de série II de 1^{ère} ou 2^{ème} personne ou par un nominal entier. Il n'apparaît pas pour autant dans toutes les occurrences de ces structures, car il n'est nécessaire qu'avant certains radicaux (cf. II-3).

Un syntagme génitif contient souvent un morphème relationnel *l-*, que le génitif soit marqué par *-a* (175) ou non (175), ou par *aʔe* (176).

(175) *pulelu-l-aʔil zawal-a-l-aʔil o-ekal.*
 crapaud-RELN-fils chien-a-RELN-fils 3.I-chercher
 Le petit crapaud cherche le petit chien.

(176) *wilaʔa-uhu-aʔe-l-o-wəl-a-l-ehe o-ze-pihig-o kupao.*
 comou-grand-DEM-RELN-feuille-vidé-a-RELN-POSTP 3.I-REFL-prendre-CONT PL.S
 Ils s'accrochent aux feuilles du comou.

On trouve aussi le relationnel précédant les mêmes lexèmes nominaux quand ils prennent un indice de série II de première ou deuxième personne.

(177) *nōde-l-aʔil-a-kom*
 1INCL.II-RELN-fils-a-PL
 nos enfants

Le relationnel se trouve aussi entre certaines postpositions et leur objet.

(178) *o-sisig-a-l-ehe o-zebaladz pia pia.*
 3.COREF-sœur-a-RELN-POSTP 3.I-jouer nuit nuit
 Il couche avec sa sœur toutes les nuits.

Comme nous le verrons au chapitre 15, III-2, les postpositions se sont développées en subordinants. Dans ce cas-là, le relationnel s'insère entre la proposition et le subordinant.

(179) [*awak^{wəl} o-kidʒe*]-*l-ehe,* *ka-wi o-wag.*
 homme 3.I-avoir.peur-RELN-SUB guêpe-ABL 3.I-aller
 Parce que l'homme a peur, il s'éloigne des guêpes.

II- 2. Les formes du relationnel

Le morphème relationnel connaît trois différentes formes phonétiques *l-* ~ *le-* ~ *n-*. Ces trois formes peuvent facilement être décrites comme trois allomorphes du même phonème. *le-* est utilisé devant une consonne (*lek^{wal}91*) et devant un mot commençant par la voyelle /i/ (*iba, ipil, itʃig, ikiɟ*). *n-* est utilisé devant les racines qui portent le trait de nasalité (*anō, amiŋ*). Enfin, c'est *l-* qui est utilisé dans le reste des contextes (c'est-à-dire devant les racines commençant par une voyelle autre que /i/ et ne portant pas le trait de nasalité). *l-* est clairement la variante la plus fréquente.

De plus, une racine qui prend normalement le relationnel *l-* prendra souvent *n-* (parfois /d/ réalisé [nd]) après l'indice de 2^{ème} personne du pluriel de série II *pe-*, devant un verbe ou une postposition.⁹²

- (183) *pe-n-edu.* phrase élicitée
 2PL.II-RELN-entendre
 Il vous entend.
- (184) "wai *pe-d-ehe* *za-maʔẽ-tane.*"
 NEG 2PL.II-RELN-POSTP INDET.I-voir-DESID
 "On ne veut pas vous voir".

Mais pour la forme du relationnel entre un nom et un indice *pe-*, la situation est plus compliquée et plus floue. Nos données sont toutes élicitées, car les textes ne contiennent qu'un seul exemple de possesseur de 2^{ème} personne du pluriel (*pe-tunaʔiŋ* "votre cœur", racine qui ne prend pas le relationnel). Voici des

⁹¹ Les racines qui prennent le relationnel sont toutes censées commencer par une voyelle. Rodrigues et Cabral (2002) donnent l'étymologie suivante pour "épouse" : **-er-eko-át* (où *er* est le morphème de causatif-comitatif et *eko* signifie "vivre"). Si cette étymologie est bonne, en émérillon, le premier /e/ a été réanalysé comme faisant part du relationnel, et la racine est maintenant *lek^{wal}* : c'est ainsi qu'elle apparaît, sans le relationnel, après *o-* et comme nom incorporé.

⁹² Cette forme nasalisée du relationnel après un indice de série II de 2^{ème} personne du pluriel est reconstruite par Jensen (1998a, p. 498) pour le proto-tupi-guarani. Jensen pense que cette nasalité est le résultat d'une propagation de la nasalité du pronom libre de 2^{ème} personne du pluriel qui est reconstruit comme *pe ...ẽ* et qui aurait donné le préfixe *pe-*.

exemples élicités, où *apicɔ̃* est un nom qui demande le relationnel. Dans l'exemple (185), le relationnel semble prendre la forme *d-* après le 2^{ème} personne du pluriel. Le même sens de "votre maison" peut être exprimé comme en (186), où la forme de relationnel est *l-* mais l'indice de personne est *pene*. On trouve aussi cette forme *pene* avec *ial*, un nom qui ne prend pas le relationnel (187).

- | | | |
|-------|--|-----------------|
| (185) | <i>pe-d-apicɔ̃</i>
2PL.II-RELN-maison
"votre maison" | exemple élicité |
| (186) | <i>pene-l-apicɔ̃</i>
?-RELN-maison
"votre maison" | exemple élicité |
| (187) | <i>pene-ial</i>
2PL.II-pirogue
"votre pirogue" | exemple élicité |

Les deux derniers exemples sont problématiques, notamment car ils sont élicités. Hors contexte, il est difficile de savoir s'il faut interpréter *pene* comme un indice de série II (variante du *pe-* de l'exemple d'avant) ou comme le pronom libre de 2^{ème} personne du pluriel *pene(kom)*⁹³, qui serait ici en position de complément génitival (litt. "la maison/pirogue de vous"). Mais étant donné que la forme *pene* peut aussi être utilisée devant des postpositions (188), alors que les postpositions n'acceptent pas les pronoms libres comme objet, *pene-* doit forcément être analysé comme une variante du *pe-* de série II dans l'exemple suivant. Si *pene-* existe comme variante de *pe-* devant les postpositions, il semble possible de l'analyser de même devant les noms comme en (186) ou (187).

- | | | |
|-------|---|-----------------|
| (188) | <i>pene-poli</i>
2PL.II-à.côté.de
à côté de vous | exemple élicité |
|-------|---|-----------------|

Il est tout à fait probable que la forme *pene-* soit une réanalyse de la séquence de l'indice de série II *pe-* suivi du relationnel *l- ~n-* en un indice de personne de série II *pene-*, qui constitue alors une variante de *pe-*. Cette réanalyse effectuée, *pene*

⁹³ Les pronoms libres seront présentés au chapitre 6, II.

peut logiquement être combiné au relationnel comme en (186). Jensen (1998a, p.502) décrit un phénomène similaire en guarani mbya : *pe* a apparemment été remplacé systématiquement par *pene*, qui peut être utilisé en conjonction avec le relationnel *r-*. Cependant, l'acceptabilité des formes en *pene-* et *pede-* devant un nom en émérillon n'a pas l'air claire et consensuelle pour nos informateurs. On peut juste formuler quelques remarques : *pene-* et *pede-* sont très bien acceptés avec un nom qui prend le relationnel, mais très mal avec les emprunts et les nominoïdes. La variante en *pede-* est pour certains locuteurs réservé aux noms à relationnel.

De même, le préfixe de 1^{ère} personne exclusive de série II, *ole-* (reconstruit pour le proto-tupi-guarani † *oré-* par Jensen 1998a) connaît une variante dans laquelle on peut reconnaître un relationnel : *olone-* et une autre dans laquelle la nasalité du relationnel se serait propagée : *onone-*.

S'il est possible mais non certifié que *pe-* et *ole-* de série II aient des variantes *pene-*, *olone-* et *onone-*, il est clair que les pronoms libres leur correspondant (2^{ème} personne du pluriel et 1^{ère} exclusive) sont *pene* et *olone* en émérillon, alors que leur reconstruction par Jensen (1998a) pour le proto-tupi-guarani est † *pe...ẽ* et *oré*. Il est fort probable qu'il faille voir là un figement du relationnel sur le pronom.

II- 3. Les différentes racines qui prennent le relationnel

Le relationnel apparaît sur une classe de racines déterminée lexicalement, généralement appelée classe II (d'après Rodrigues 1953). Cette classe de racine ne correspond pas à une partie du discours, mais elle englobe des noms, des postpositions et des verbes. Elle a pour caractéristique phonologique de ne comprendre que des mots à initiale vocalique⁹⁴. Mais cette propriété n'est pas définitoire, étant donné que d'autres racines à initiale vocalique ne prennent pas le relationnel. Par exemple, la racine *i* "mère" est toujours possédée, mais elle ne

⁹⁴ A l'exception de la racine *lek^wal* "épouse", cf. note 91.

Ces noms ont deux autres spécificités. Premièrement, quand leur possesseur est de 3^{ème} personne non coréférente au sujet, ce n'est pas *i-* qui marque ce possesseur, mais un de ses allomorphes spécifiques aux noms à relationnel (*t-* ou zéro). Deuxièmement, même si ces noms sont obligatoirement possédés, il arrive que l'on veuille en parler sans préciser leur possesseur. Dans ce cas-là, on utilise un préfixe de personne non-spécifiée, ce qui est attesté dans beaucoup d'autres langues amazoniennes (Aikhenvald, 1996).⁹⁵ Les exemples suivants montrent le nom *apitɕ* "maison" avec un possesseur de troisième personne coréférente au sujet (192), avec un possesseur de troisième personne non coréférente au sujet (193), avec un possesseur non-spécifié (194), et avec un possesseur qui est un nominal (195) ou une deuxième personne (196).

- (192) pitəŋ o-nan o-ho **o-apitɕ**-a-koti. phrase élicitée
 enfant 3.I-courir 3.I-aller 3.COREF-maison-a-vers
 L'enfant part en courant chez lui.
- (193) pitəŋ o-nan o-ho **Ø-apitɕ**-a-koti. phrase élicitée
 enfant 3.I-courir 3.I-aller 3.II-maison-a-vers
 L'enfant_i part en courant vers sa_j maison.
- (194) kob **t-apitɕ** ɕuliba-we-ʔe.
 COP NSP-maison escalier-aussi-INTENS
 Il y a une maison avec un escalier.
- (195) kol əŋ **t-apitɕ** teko-**l-apitɕ**-a-te.
 puis DEM NSP-maison Emérillon-RELN-maison-a-FOC
 Et cette maison, c'est la maison des Emérillons.
- (196) apam-aʔe-sipala-te nōde-**l-apitɕ**-am.
 étranger-DEM-fer-FOC 2.II-RELN-maison-TRANSL
 C'est le fer des étrangers (qu'on utilise) pour nos maisons.

⁹⁵ Ce morphème de possesseur non-spécifié de l'émérillon peut être considéré comme de la dérivation : il rend autonome le nom obligatoirement possédé par rapport aux possessifs, c'est-à-dire qu'il modifie la valence du nom. Nous remercions D. Creissels pour cette proposition d'analyse. On trouve un morphème de même fonction sur une certaine classe de noms dans les langues mayas (Grinevald, c.p) : ainsi, la racine *q'ab'a* "main" est soit possédée et donc préfixée d'un indice possessif (*n-q'ab'a* "ma main"), soit non possédée et donc suffixée (*q'ab'a-dj* "une main").

Le tableau suivant propose des exemples de racines nominales fréquentes prenant le relationnel.⁹⁶ Elles sont classées selon les formes que prennent leur relationnel, leur possesseur de 3^{ème} personne non coréférente avec le sujet et leur possesseur non-spécifié. On donne enfin leur traduction.

Racine (avec <i>zo-</i> , <i>o-</i> , incorporée)	Avec possesseur = 1 ^{ère} ou 2 ^{ème} pers, ou GN	Avec possesseur = 3 ^{ème} pers	Avec possesseur non spécifié	Traduction	
1	<i>apɨɟ</i>	<i>l-</i>	\emptyset -	<i>t-</i>	maison, carbet
	<i>ɪlu</i>	<i>l-</i>	\emptyset -	<i>t-</i>	vêtements
	<i>owa</i>	<i>l-</i>	\emptyset -	<i>t-</i>	visage
2	<i>ɪwɨk^wəl</i>	<i>l-</i>	<i>t-</i>	<i>t-</i>	corps, sang
	<i>(?)u</i>	<i>l-</i>	<i>t-</i>	<i>t-</i>	père
3	<i>amɨŋ</i>	<i>n-</i>	<i>t-</i>	?	grand-père
4	<i>lek^wal</i>	<i>le-</i>	<i>i-</i>	?	épouse
5	<i>el (hel ?)</i>	<i>l-</i>	<i>h-</i>	?	nom

Tableau 16 : Racines nominales qui prennent le relationnel

Ces racines se répartissent en cinq groupes distincts, uniquement définis par la nature des allomorphes de leurs préfixes. Ce regroupement ne semble pas correspondre à une classification sémantique. Nous avons ordonné ces groupes selon leur importance quantitative : les noms des deux premiers groupes (qui prennent la forme *l-* du relationnel) sont beaucoup plus nombreux que les autres. Dans la limite de nos connaissances actuelles, les deux derniers groupes sont restreints à une unité chacun.

⁹⁶ Nous n'avons pas toutes les informations présentées dans ce tableau pour une liste de noms plus importante.

Pour les comparer avec les noms qui ne prennent pas le relationnel, ces derniers n'ont ni de préfixe relationnel, ni de préfixe de possesseur non spécifié, et leur indice possessif de 3^{ème} personne non coréférente est toujours *i-*.

II- 3.2. Le relationnel sur les postpositions/subordonnants

Parmi les postpositions/subordonnants les plus courantes, seules deux racines prennent le relationnel : *ehe* ("avec, destination, source, parce que") et *upi* ("le long de, en parallèle à, au fur et à mesure que")⁹⁷. Vu qu'elles prennent un zéro comme possesseur de 3^{ème} personne, et le relationnel *l-*, elles correspondent aux racines nominales du type 1 dans le tableau précédant.

Racine (avec <i>zo-</i> , <i>o-</i>)	Avec objet = 1 ^{ère} ou 2 ^{ème} pers, GN, ou proposition	Avec objet = 3 ^{ème} pers	Traduction
<i>ehe</i>	<i>l-</i>	∅-	avec, destination, source, parce que
<i>upi</i>	<i>l-</i>	∅-	le long de, en parallèle à, au fur et à mesure

Tableau 17 : Postpositions qui prennent le relationnel

La deuxième colonne indique que ces deux postpositions prennent le relationnel quand elles sont précédées d'un objet de 1^{ère} ou 2^{ème} personne (197), d'un objet nominal (198). La forme du relationnel est *l-*⁹⁸ avec les deux postpositions.

⁹⁷ On trouve aussi le relationnel *n-* sur *oãtfã* "devant" et *l-* sur *akapilel* "derrière" que nous analysons comme des postpositions, mais qui sont des innovations émérillon dont nous ignorons l'origine exacte (cf. Chapitre 9, II-1.2.2.)

⁹⁸ Sauf avec la 2^{ème} personne du pluriel, cf. II-2.

(197) "d-a-maʔě-tal-ai-dʒi de-**l-ehe**."
 NEG-1SG.I-voir-FUT-beaucoup-NEG 2SG.II-RELN-POSTP
 "Je ne veux plus te voir."

(198) o-sisig-a-**l-ehe** o-zebaladz pia pia.
 3.COREF-sœur-a-RELN-POSTP 3.I-jouer nuit nuit
 Il couche avec sa sœur toutes les nuits.

Comme l'indique la première colonne du tableau, les postpositions n'utilisent pas le relationnel avec un objet de 3^{ème} personne coréférentielle au sujet *o-* (199) ou un objet dit "indéterminé" *zo-* (200), comme les noms.

(199) ka o-mõ-waŋ zawal **o-ehe**.
 guêpe 3.I-CAUS-se.disperser chien 3.COREF-POSTP
 Le chien disperse les guêpes sur lui-même.

(200) kol a-potal zãdam-am t-o-maʔě **zo-ehe**.
 donc 1SG.I-vouloir gendarme-TRANSL BUT-3.I-regarder INDET.II-POSTP
 Donc je voudrais des gendarmes pour surveiller les gens.

Enfin, la troisième colonne montre que quand l'objet de ces postpositions est de troisième personne non coréférentielle avec le sujet, il est marqué par un zéro et non par un *i-*.

(201) di-si-maʔě-ni **Ø-ehe**.
 NEG-INCL.I-voir-NEG 3.II-POSTP
 On ne le voit pas.

Ces deux postpositions pouvant être analysées, en synchronie, comme des subordinants, le relationnel apparaît aussi entre les postpositions/subordonnants et la proposition qui les précède.

(202) [awak^wəl o-kidʒe]-**l-ehe**, ka-wi o-wag.
 homme 3.I-avoir.peur-RELN-SUB guêpe-ABL 3.I-aller
 Parce que l'homme a peur, il s'éloigne des guêpes.

(203) o-wil o-ʔa, [bokal-a-pe o-akaŋ o-mõde]-**l-ehe**.
 3.I-se.détacher 3.I-tomber bocal-a-dans 3.COREF-tête 3.I-mettre-REL-SUB
 Il est tombé, parce qu'il a mis la tête dans le bocal.

Cette distribution étant très spécifique, et ces contextes peu courants dans nos textes, notre corpus contient des illustrations de seulement trois verbes avec le relationnel : *aihi* "aimer, apprécier" (204) et (205), *ekal* "chercher" (206) et *esag* "voir"⁹⁹ (207).

Voici donc la liste (probablement non exhaustive) des verbes émérillon qui prennent le relationnel, classés selon la forme du relationnel qu'ils prennent (*l-* ~ *le-* ~ *n-*). Encore une fois, ceux qui prennent le relationnel *l-* sont les plus nombreux.

Verbe	Relationnel	Traduction
<i>aihi</i>	<i>l-</i>	aimer, apprécier
<i>ekal</i>	<i>l-</i>	chercher
<i>elaho</i>	<i>l-</i>	porter, amener
<i>esag</i>	<i>l-</i>	voir
<i>ezal</i>	<i>l-</i>	laisser
<i>edu</i>	<i>l-</i>	entendre, écouter
<i>ipil</i>	<i>le-</i>	soulever
<i>itʃig</i>	<i>le-</i>	faire tomber par terre
<i>ikiɖʒ ~ ikiɖʒ</i>	<i>le-</i>	prendre, acheter
<i>anõ</i>	<i>n-</i>	attendre

Tableau 18 : Racines verbales qui prennent le relationnel

Rappelons que les trois formes du relationnel peuvent être expliquées par des règles morphophonologiques : *le-* est utilisé devant consonne et devant un mot commençant par la voyelle /i/, *n-* est utilisé devant les racines qui portent le trait de nasalité et *l-* dans les autres contextes.

⁹⁹ Ce verbe n'est utilisé que comme gérondif, et pas comme verbe indépendant.

Nous avons vu, pour l'émérillon, la distribution générale du relationnel, puis ses différentes formes. Nous venons d'exposer les types de racines qui prennent le relationnel, et pour chaque type, d'illustrer les contextes dans lesquels on le trouve, par opposition aux contextes où l'on trouve d'autres morphèmes qui indiquent généralement les personnes. Les faits décrits sont équivalents à ce qu'on trouve dans d'autres langues tupi-guarani, mis à part quelques variations phonologiques (Cabral 2000b). Dans les deux sections suivantes, nous allons partir de ces faits pour discuter de manière plus globale de la fonction du relationnel ainsi que de sa position par rapport au système d'indexation des personnes. Cette discussion se base sur les données émérillon mais veut être valable pour l'ensemble des langues tupi-guarani.

II- 4. Discussion sur la fonction du relationnel

Le préfixe relationnel *r-* est présent dans toutes les langues de la famille tupi-guarani. Il s'agit d'un phénomène typique des langues tupi-guarani et même régional. Des explications de sa fonction ont été proposées pour chaque langue particulière, et des discussions ont été menées pour l'ensemble de la famille (Cabral 2000b, Payne 1994 et surtout Jensen 1998a, p.557-562).

"The prefix *r-* is an inescapable subject of discussion by Tupí-Guaraní linguists. Nevertheless, there has been a lack of uniformity in the way it is described. It has been referred to as a linking prefix, a relational, an attributive, or simply as R-. [...]" (Jensen 1998a, p.502)

Dans cette partie, nous allons voir les différentes analyses du préfixe *r-* qui ont été proposées. Ce qui en ressort finalement, est une définition du morphème basée sur sa distribution. Nous donnerons ensuite quelques propositions diachroniques qui ont été lancées au sujet de ce même morphème.

II- 4. 1. Fonction

Diverses fonctions ont été considérées par les différents auteurs :

- Jensen (1998a, p. 557-559) a éliminé les analyses du relationnel comme épenthèse et comme morphème conditionné phonologiquement
- L'idée que les noms prenant le relationnel seraient "inaliénables" et ceux sans seraient "aliénables" paraît raisonnable à première vue. Mais on ne voit pas ce qui, au niveau sémantique, pourrait justifier que "père" et "aimer/apprécier" prennent le relationnel, et non "mère" et "aimer/vouloir" ? Ainsi, Praça (1999, sur le tapirapé), note que les deux classes de racines (celles qui prennent le relationnel et celles qui ne le prennent pas) ne se distinguent pas sémantiquement. Nous rejetons donc l'explication de la fonction du relationnel en termes d'inaliénabilité (vue comme une propriété sémantique), d'autant plus qu'elle serait difficile à appliquer aux verbes et aux postpositions.

Trois autres propositions posent le même problème : elles expliquent la présence du relationnel sur certains verbes, et non son absence sur les autres. D'autre part, ces explications ne sont pas extensibles aux propositions et aux noms :

- Seki propose que les préfixes relationnels spécifient que la racine est reliée à un participant inactif (Seki 1990, pour le kamaiura).
- Une analyse proche est celle d'Harrison (Harrison 1986, pour le guajajara), qui affirme que le relationnel marque sur les verbes intransitifs le non-contrôle de l'action par le référent ainsi marqué, et sur les verbes transitifs une relation anormale dans la hiérarchie de S et O.
- Cette dernière analyse nous rappelle l'hypothèse de Payne (1994), pour qui le morphème *r-* serait un marqueur d'inverse. Nous avons rejeté cette hypothèse au chapitre 3, III-4.

La seule fonction que l'on peut donner au relationnel est syntaxique¹⁰⁰ : signaler une relation de dépendance, de détermination entre une tête de syntagme (nom, verbe ou postposition) et son complément obligatoire (possesseur ou objet). Cabral parle de cette marque de dépendance en terme de "stratégie pour valider, en syntaxe, ce qui dans le lexique est relatif, c'est-à-dire, ce qui est prédestiné à être mis en relation avec un déterminant" (Cabral 2000b, p. 240).¹⁰¹ Pour Rodrigues (1981), comme pour beaucoup d'autres auteurs à sa suite, le relationnel (*r-* ~ zéro) indiquerait que le déterminant est le constituant contigu ou immédiatement précédent. Ceci implique par opposition que la présence d'un *i-* ~ *t-* ~ zéro (considéré par Rodrigues comme un autre type de relationnel) signale la non contiguïté du déterminant. Nous reviendrons sur la question du rapport (syntagmatique ou paradigmatic) entre le relationnel et les marques de personne dans la section suivante.

Mais si cette fonction de "lien de dépendance" paraît claire, on peut, comme dans le cas de la distribution de la série II (Chapitre 3, II), se demander s'il est nécessaire de chercher une fonction commune dans les divers usages du relationnel, ou si ceux-ci s'expliquent par une grammaticalisation d'un même morphème dans différents contextes à une même époque. Quelques hypothèses diachroniques intéressantes ont justement été proposées par divers auteurs.

II- 4. 2. Hypothèses diachroniques

Dans cette section, nous regroupons trois hypothèses diachroniques qui ont été proposées autour du problème du relationnel. Elles touchent chacune à des aspects différents de la question, et ne se contredisent donc pas.

¹⁰⁰ A ceci près qu'il dépend aussi d'un conditionnement lexical non expliqué.

¹⁰¹ Notre traduction de "uma estratégia para licenciar, na sintaxe, o que no léxico é relativo, ou seja, o que está fadado a ser relacionado a um determinante".

La première hypothèse part du fait que seulement une classe de radicaux prend le relationnel, et que cette classe ne peut pas être distinguée sémantiquement. Ribeiro (c.p.) en conclut que la distinction entre les deux classes de racines est purement morphologique, certainement avec une motivation phonologique en diachronie. Jensen (1998a, p. 562) formule explicitement cette hypothèse : "The *r-* morpheme may originally have occurred with all vowel-initial stems ; those Class I stems which in Proto-Tupí-Guaraní are vowel initial may have lost an initial consonant."

Une seconde hypothèse veut expliquer le fait que le relationnel n'est nécessaire qu'avec les indices de 1^{ère} et 2^{ème} personne (et ni avec les 3^{ème} personnes, ni avec la personne indéterminée, ni avec le possesseur non-spécifié pour les noms). On ne peut pas expliquer la présence du relationnel par un lien particulier avec les personnes de l'interlocution, vu qu'il est aussi utilisé avec un nom plein le précédant. Ribeiro (c.p.) propose une explication d'ordre diachronique : si des morphèmes pronominaux s'affixent directement à la racine sans requérir le préfixe relationnel, cela signifie probablement que ces morphèmes étaient déjà des préfixes à l'époque où les relationnels sont apparus. Cela veut dire que *i-*, *t-*, et *o-* devaient être des préfixes avant que les autres indices de personne soient devenus morphologiquement dépendants (ce qui est cohérent avec l'évolution des marques de personne sur les verbes indépendants proposées par Gildea 2002). Nous pouvons à cet effet rappeler que les indices de série II (sauf le *i-*) dérivent probablement de pronoms indépendants, et qu'ils sont encore considérés dans certaines langues de la famille (par exemple, le tupinambá, Rodrigues 1981 et le kamaiurá, Seki 2000) comme des mots indépendants, alors que le *i-* est un préfixe non-accentué.

Enfin, une troisième hypothèse très intéressante, proposée par Payne (1994), fait du relationnel une ancienne marque casuelle. Elle évoque la possibilité que le

relationnel ait été originellement une marque de génitif en pré-proto-tupi-guarani¹⁰². Les verbes dépendants auraient pu à une époque être traités comme des noms (ce qui expliquerait la structure O-RELN-V)¹⁰³. Ensuite, la présence du *r-* dans les propositions principales aurait pu apparaître à la suite d'une réinterprétation des constructions subordonnées ou nominalisées/statives en propositions principales. De même, il est envisageable que les postpositions qui prennent le relationnel soient dérivées de noms.

Ces trois hypothèses concernant le relationnel *r-* sont intéressantes. Les deux premières nous paraissent fort probables. Quant à la troisième, elle sous-entend beaucoup d'autres changements diachroniques dont nous ne sommes pas à même de juger la validité.

II- 5. Discussion sur la position du relationnel par rapport au système d'indexation des personnes

Un autre point est problématique dans la présentation du préfixe *r-*. Il existe plusieurs façons de le présenter dans son rapport aux indices de personne (marquant les possesseurs des noms, les objets des postpositions et des verbes) et au morphème de possesseur non spécifié. Ainsi, certains linguistes considèrent les indices de troisième personne et les préfixes de possesseur non spécifié comme des allomorphes du relationnel (II-5.1.), alors que d'autres considèrent que le relationnel est un morphème unique, extérieur aux paradigmes des indices de personne (II- 5.2). C'est cette seconde attitude que nous avons adoptée.

¹⁰² Cette hypothèse est compatible avec la précédente : cette marque casuelle aurait marqué les pronoms de 1^{ère} et 2^{ème} personne, alors que la 3^{ème} personne était déjà exprimée par un préfixe et ne pouvait donc recevoir de cas.

¹⁰³ Ou alors certains verbes régissaient un complément au cas *-r*.

Le problème majeur est qu'aucune interprétation n'est entièrement satisfaisante, car le système est de toute façon asymétrique et il faut forcément envisager un paradigme incomplet quel qu'il soit. Nous allons maintenant expliquer pourquoi nous avons choisi la présentation faite jusqu'ici (qui correspond à celle de Jensen), en reconnaissant qu'il s'agit d'un choix théorique et que les autres possibilités ne sont pas foncièrement fausses pour autant.

Reprenons des données de base de l'émérillon avec un paradigme nominal d'une racine qui prend le relationnel (*owa*), et un autre d'une racine qui ne prend pas le relationnel (*kija*). Nous donnons les différentes formes possibles en limitant les différentes 1^{ère} et 2^{ème} personnes à celles du singulier pour plus de simplicité.

<i>e-l-owa</i>	mon visage	<i>e-kija</i>	mon hamac
<i>de-l-owa</i>	ton visage	<i>de-kija</i>	ton hamac
<i>X-(a)-l-owa</i>	le visage de X	<i>X-(a)-kija</i>	le hamac de X
\emptyset - <i>owa</i>	son visage	<i>i-kija</i>	son hamac
<i>o-owa</i> ¹⁰⁴	son propre visage	<i>o-kija</i>	son propre hamac
<i>t-owa</i>	le visage	<i>kija</i>	le hamac

Tableau 19 : Paradigmes nominaux avec et sans relationnel

Tout le monde s'accorderait sur le fait que *e-* et *de-* soient des indices de personne (marquant ici le possesseur du nom), et que *l-* soit un relationnel. Les morphèmes qui posent problème sont ceux qui apparaissent à la 3^{ème} personne (\emptyset - et *i-*, *o-* coréférentiel) et la marque de possesseur non spécifié (ici, *t-*).

¹⁰⁴ Cette forme est en réalité réalisée phonétiquement [owa].

Indices de personne	Relationnels
e-, de- (indices 1 ^{ère} et 2 ^{ème} personnes)	l-
<p style="text-align: center;">i- ~ Ø- o- t-</p>	
	existence d'un relationnel Ø ?

Tableau 20 : Problèmes de classement de préfixes en indices de personne ou relationnels

A partir de ce constat, deux façons de présenter le système sont possibles. La première (II-5.1.) consiste à considérer les indices de 3^{ème} personne et la marque de possesseur non spécifié comme des relationnels. Nous intitulons cette approche "analyse en paradigme de relationnels". La deuxième (II-5.2.), que nous avons suivi jusqu'à maintenant, ne considère comme relationnel que la marque qui apparaît dans les contextes illustrés dans le Tableau 15, p. 120, c'est-à-dire entre certains noms, postpositions et verbes et leur objet (le possesseur pour les noms), quand celui-ci est un nom ou une marque de série II de 1^{ère} ou 2^{ème} personne. Nous dénommons cette analyse "l'analyse du relationnel unique". Ces deux analyses sont maintenant présentées.

II- 5. 1. L'analyse en paradigme de relationnels

La première analyse représente approximativement les opinions (plus détaillées, et plus variées) de Dietrich (1990), Seki (2000), Rodrigues (1981), Harrison (1986), et de Cabral (2000b). Elle limite le paradigme de série II aux préfixes qui sont communs à toutes les racines (concrètement, dans le tableau, aux deux paradigmes-types *owa* et *kija*). Par conséquent, elle exclut de la série II les formes

de troisième personne (ici *i-* et zéro). Ces formes sont considérées comme des "relationnels" avec les marques *l-*, *o-* (coréférentiel) et *t-* (non spécifié) car elles sont en rapport paradigmatiques. De plus, un relationnel "zéro" est introduit dans le paradigme de type *kija*, par souci de parallélisme avec celui de *owa* où *l-* est présent. Selon cette analyse, toutes les formes en gras du tableau ci-dessous sont des relationnels (et donc exclues des séries d'indices de personne, qui, eux, se retrouvent sans 3^{ème} personne).

type <i>owa</i>	type <i>kija</i>	
e-l- N	e- Ø- N	1 ^{ère} SG
de-l- N	de- Ø- N	2 ^{ème} SG
X-(a)-l- N	X-(a)- Ø- N	synt.génitival
Ø- N	i- N	3 ^{ème}
o- N	o- N	3 ^{ème} coref
t- N	N	sans possesseur

Tableau 21 : Paradigmes de noms selon l'analyse en paradigme de relationnels

Les différentes formes du "relationnel" sont expliquées par leur distribution : *i-* et zéro avec un déterminant non contigu, *l-* et zéro avec un déterminant contigu, *t-* de possesseur non spécifié, et *o-* de coréférence avec le sujet.

Le traitement de différents préfixes comme "relationnels" est basé sur l'observation qu'ils occupent la même position paradigmatique. En utilisant ce même argument, le *i-* et le *o-* pourraient de manière aussi légitime être vus comme appartenant au paradigme des indices de personne en émérillon. Mais dans les langues où les indices de personne sont phonologiquement traités comme des clitiques ou des mots indépendants (sauf *i-* qui est un préfixe), cela constitue un argument pour présenter le *i-* et le *o-* à part des indices de série II.

Ce qui nous pose problème dans cette analyse, c'est l'hétérogénéité des valeurs et de la distribution des relationnels, et cette volonté de "décomposer" les paradigmes d'indices de personne pour pouvoir créer une classe de préfixes relationnels, quant au fond, seulement *l-* a clairement un rôle de "lien". Un inconvénient majeur réside dans la multiplication des relationnels "zéro". Ribeiro (c.p.) fait remarquer qu'en suivant cette analyse, on peut trouver des relationnels dans toutes les langues qui font une distinction entre noms obligatoirement possédés et noms aliénables et qui utilisent la marque de 3^{ème} personne par défaut. Il précise que l'utilisation de *i-* et *t-* comme "marqueurs de non contiguïté" découle simplement de la nécessité pour certains lexèmes d'apparaître avec un complément.

II- 5. 2. L'analyse du relationnel unique

L'autre analyse possible est celle suivie par Jensen (1998a), Payne (1994), Da Silva 2002, Ribeiro (c.p.), Gildea (c.p.), et celle que nous adoptons dans cette thèse. Elle inclut le *i-*, le zéro et le *o-* dans les paradigmes d'indices de personne. L'argument est ici aussi le fait que ces indices ont la même distribution que les autres marques de série II (*e-* et *de-*), et la même relation avec la tête (ils spécifient la personne du complément). En suivant cette présentation, on a donc un seul relationnel : le *l-* qui est en relation syntagmatique avec les marques de personne (en gras dans le tableau suivant). Aucun relationnel zéro n'est postulé.

e- l -owa	e-kija
de- l -owa	de-kija
X-(a)- l -owa	X-(a)-kija
∅-owa	i-kija
o-owa	o-kija
t-owa	kija

Tableau 22 : Paradigmes de noms selon l'analyse du relationnel unique

Cette analyse tient d'autant bien en émérillon que les indices de série II sont tous des préfixes (phonologiquement liés), et que l'on trouve quelques exemples de combinaison d'un relationnel avec une marque de 3^{ème} personne. Il est possible que par analogie, cette marque s'étende à la 3^{ème} personne. De tels exemples sont peu fréquents, produits par des locuteurs plutôt jeunes, et ne se retrouvent pas dans d'autres langues tupi-guarani : ils sont apparemment "innovateurs". Il est évident que la combinaison d'un relationnel avec une marque de 3^{ème} personne rend difficile l'analyse de *l-* et des indices de troisième personne comme une classe unique de préfixes mutuellement exclusifs (analyse en paradigme de relationnels).

- | | |
|---|---------|
| (208) o-l-ilu
3.COREF-RELN-vêtements
ses propres vêtements | élicité |
| (209) i-n-amij
3.II-RELN-grand-père
son grand-père | élicité |
| (210) i-l-upi
3.II-RELN-le.long.de
le long de lui | élicité |

Le principal inconvénient de cette présentation est de masquer la particularité de la 3^{ème} personne dans le paradigme de série II, personne qui ne se combine pas avec le relationnel (probablement dû à une grammaticalisation précoce). Mais après tout, la particularité de la 3^{ème} personne dans un paradigme est plus que fréquente dans la typologie des langues du monde. C'est donc l'analyse du relationnel unique que nous adoptons dans cette thèse. Ce choix n'a pas de conséquences directes sur l'analyse d'autres composantes de la langue, les descriptions suivantes sont ainsi toutes compatibles avec l'autre hypothèse.

Nous venons d'exposer dans les deux derniers chapitres les indices de personne, le suffixe *-a* et le relationnel, tous des morphèmes essentiels, fréquents et à large

distribution en émérillon et qui se retrouvent dans les autres langues de la famille tupi-guarani. Dans le 3^{ème} chapitre de cette partie consacrée à la morphologie de base de l'émérillon, nous allons présenter différents morphèmes de nombre, qui connaissent aussi une large distribution mais ne constituent pas une caractéristique de la famille tupi-guarani.

Chapitre 5 : Le nombre

Contrairement aux catégories présentées dans les deux chapitres précédents, le nombre ne connaît pas de marques répandues dans les langues tupi-guarani. Il semble que chaque langue tupi-guarani ait innové en la matière. Il est intéressant de remarquer qu'aucun des morphèmes de pluriel de l'émérillon n'a été reconstruit en proto-tupi-guarani. Il semblerait en fait qu'aucun morphème de nombre ne soit commun aux langues tupi-guarani. Aucune reconstruction d'un morphème de nombre n'a donc été proposée pour le proto-tupi-guarani. Une explication à cela peut être que, si l'on suit Aikhenvald, le nombre n'est pas une caractéristique fondamentale des langues amazoniennes :

"Number is usually obligatory with pronouns/cross-referencing markers, but optional with nouns and often confined to human or animate referents. [...] Number is usually marked once in an NP. [...] Usual opposition is singular vs non-singular (plural, collective)."
(Aikhenvald, 1996, Unit 3 p.5)

En émérillon, il existe trois morphèmes de nombre, avec pour chacun une histoire singulière : le pluriel de sujet *-(o)ŋ* est une innovation émérillon apparemment native, l'autre pluriel de sujet *kupa* est peut-être du à une interférence avec le wayampi et le clitique *kom* qui pluralise les noms, leur possesseur et aussi l'objet des verbes et des postpositions, est un emprunt aux langues caribes. Le système de marquage du nombre est assez important en émérillon, et contredit sur plusieurs points la généralisation d'Aikhenvald donnée ci-dessus : dans cette langue, le nombre n'est pas obligatoire sur le verbe avec les indices de 3^{ème} personne¹⁰⁵, il n'est pas restreint aux animés, et il n'est pas limité au groupe nominal. Au niveau sémantique, il est difficile de dire si ces morphèmes de "non

singulier" expriment un pluriel ou un collectif. En conclusion, le nombre constitue une partie importante de la morphologie de base de l'émérillon, qui, comme la morphologie des deux chapitres précédents, peut être commune aux noms et aux verbes, mais qui peut aussi départager arguments et prédicats (cf. Chapitre 1, II et III).

Dans ce chapitre, nous exposerons trois morphèmes : en I, le pluriel du sujet en $-(o)\eta$, en II le pluriel du sujet en *kupa*, et en III le clitique $-kom$, qui pluralise les noms, leur possesseur et aussi l'objet des verbes et des postpositions. Pour chaque morphème de pluriel, nous verrons ses formes et ses fonctions, son statut et son origine.

I- Le pluriel du sujet en $-(o)\eta$

Le morphème $-(o)\eta$ est une marque facultative de pluriel d'un sujet de 3^{ème} personne. Il se cliticise sur un élément du groupe verbal, le plus souvent sur le prédicat, et semble être particulier à l'émérillon. Nous verrons dans un premier temps sa forme et sa fonction (I- 1) , puis dans un second temps son statut (I- 2), et nous finirons en proposant une hypothèse sur son origine (I-3).

I- 1. Forme et fonction de $-(o)\eta$

Le morphème $-(o)\eta$ connaît deux allomorphes : $-o\eta$ est utilisé après une finale consonantique, et $-\eta$ après une finale vocalique.

(211) sikãĩ-nam baipuli o-pihig-**oŋ**.
 petit-quand tapir 3.I-attraper-PL.S
 Ils ont attrapé le tapir quand il était petit.

(212) aʔe-kom kuɕabulu-am o-ho-**ŋ**.
 DEM-PL sirène-TRANSL 3.I-aller-PL.S
 Celles-ci sont devenues des sirènes.

¹⁰⁵ Typologiquement, il est assez courant que le verbe ne marque la distinction singulier/pluriel qu'aux deux premières personnes.

Ce morphème a pour fonction de pluraliser un sujet de 3^{ème} personne : en effet, seuls les indices de 3^{ème} personne ne sont pas intrinsèquement marqués pour le nombre. Les indices de personne de l'émérillon distinguent 1^{ère} personne singulier, inclusive et exclusive, et 2^{ème} personne singulier et pluriel¹⁰⁶, et ne peuvent donc pas se combiner avec le morphème *-oŋ*.

- (213) olo-kusug ***-oŋ**
 1EXCL.I-laver -PL.S
 Nous l'avons lavé.

Par contre, le clitique *-(o)ŋ* indique le pluriel du sujet de 3^{ème} personne, même dans les situations où l'objet est marqué sur le verbe transitif et non le sujet de 3^{ème} personne, qui est inféré à partir du système hiérarchique d'indexation des personnes¹⁰⁷.

- (214) e-potal-**oŋ**
 1SG.II-aimer-PL.S
 Ils m'aiment.
- (215) wiŋ'i-ne nōde-mōdokal-**oŋ**.
 loin-CONTR 1INCL.II-envoyer-PL.S
 Ils nous ont envoyé loin.

Le morphème *-(o)ŋ* ne pluralise qu'un sujet de 3^{ème} personne, et pas un objet de 3^{ème} personne :

- (216) ***a-ikiŋ-oŋ**.
 1SG.I-prendre-PL.S
 Je les ai pris.

Nous avons vu en introduction que dans les langues amazoniennes, il était fréquent que les morphèmes de nombre ne s'appliquent qu'aux référents animés ou humains. Dans nos textes, *-oŋ* pluralise toujours un sujet animé (homme, animal, ou objet inanimé qui se métamorphose en objet animé). Néanmoins, un tel phénomène

¹⁰⁶ Pour une description des indices de personne, voir Chapitre 3, I.

¹⁰⁷ Pour une description du système hiérarchique de marquage des personnes, voir Chapitre 3, III.

résulte probablement du fait que, dans les textes du corpus, la plupart des sujets réfèrent à des êtres animés. Nous avons cependant élicité des exemples de sujets inanimés qui déclenchent la pluralisation en *-oŋ*.

- (217) *maŋ ebig-oŋ.* exemple élicité
 mangue sucré-PL.S
 Les mangues sont sucrées.

Enfin, on peut noter le caractère facultatif de ce morphème : il n'apparaît pas systématiquement chaque fois que le sujet est une 3^{ème} personne du pluriel. Dans le morceau de texte suivant, le sujet est toujours "les femmes". Dans la première phrase, le pluriel est seulement marqué par *-kom* sur le groupe nominal sujet ; dans les suivantes, le sujet n'est plus présent sous forme de groupe nominal, et *-oŋ* ne marque le pluriel du sujet que dans la seconde phrase (sur un argument du verbe), et pas dans la troisième, où le pluriel n'est plus du tout marqué.

- (218) *wãiwĩ-kom-enam o-wikilog o-wikilog.*
 femme-PL-chang.top 3.I-vider 3.I-vider
 Les femmes voient les poissons.
- (219) *o-ba?e pali-am-oŋ.*
 3.I-faire boucan-TRANSL-PL.S
 Elles font un boucan.
- (220) *o-ijunŋ t-o-ka?ẽ ike?i.*
 3.I-mettre BUT-3.I-boucaner alors
 Alors elles les mettent à boucaner.

Il ne serait cependant pas juste d'affirmer que la marque *-oŋ* n'est présente que pour résoudre une ambiguïté. Dans la phrase suivante, le pluriel est marqué à la fois sur le groupe nominal (qui ne peut être que le sujet des deux prédicats intransitifs), et sur le prédicat de la deuxième proposition.

- (221) aʔe-nawe sikāi-we wāiwī-kom, aʔuwe d-o-zemōdiʔal¹⁰⁸-i-ŋ.
 DEM-quand petit-aussi femme-PL pas.encore NEG-3.I-avoir.ses.règles-NEG-PL.S
 A cette époque, les filles étaient petites aussi, elles n'avaient pas encore leurs règles.

I- 2. Statut de *-(o)ŋ* : un clitique de groupe verbal

Dans cette section, nous allons démontrer que ce morphème est un clitique de fin de groupe verbal. En effet, il peut apparaître sur différents types de prédicats, ainsi que sur d'autres constituants du groupe verbal, mais seulement si ceux-ci suivent le prédicat.

Les exemples (211) à (215) et (221) possèdent une marque de pluriel *-(o)ŋ* sur un verbe. On peut aussi trouver *-(o)ŋ* sur d'autres types de prédicats : attributif en (217), nominal en (222) et (223) ou numéral en (224).

- (222) Ø-apidɕ-ŋ. phrase élicitée
 3.II-maison-PL.S
 Ils ont une maison.
- (223) i-awu-ŋ ikiʔi.
 3.II-parole-PL.S maintenant
 Maintenant ils parlent.
- (224) mokoŋ-ŋ. phrase élicitée
 deux-PL.S
 Ils sont deux.

Sur le prédicat, le morphème de pluriel *-(o)ŋ* se suffixe après les morphèmes de temps et de négation.

- (225) i-wilakoti-enam beɕu o-ɨnuŋ-tal-ŋ.
 3-sous-chang.top cassave 3.I-mettre-FUT-PL.S
 En dessous, ils vont faire des cassaves.
- (226) aʔuwe d-o-zemōdiʔal-i-ŋ.
 pas.encore NEG-3.I-avoir.ses.règles-NEG-PL.S
 Elles n'avaient pas encore leurs règles.

¹⁰⁸ Il est probable que ce mot puisse encore être segmenté.

Comme nous l'avons déjà remarqué, cette marque de pluriel peut se placer sur un autre mot que le prédicat, par exemple l'argument objet en (227) et (228). Elle continue alors de marquer le pluriel du sujet.

(227) o-baʔe pali-am-**oŋ**.
3.I-faire boucan-TRANSL-PL.S
Elles font un boucan.

(228) o-boʔi baipuli-l-ie-**ŋ**.
3.I-trancher tapir-RELN-ventre-PL.S
Elles ouvrent le ventre du tapir.

On peut aussi la trouver sur un groupe postpositionnel.

(229) o-meʔeŋ-gatu sipala-m i-ɕuʔe-**ŋ**.
3.I-donner-bien métal-TRANSL 3.II-pour-PL.S
Ils lui donnent du fer.

(230) o-ho i-koti-**ŋ**.
3.I-aller 3.II-chez-PL.S
Elles vont chez lui.

(231) owi-o-wil i-ʔal-**oŋ**.
RED-3.I-monter 3.II-sur-PL.S
Ils sont montés dessus (sur l'oiseau).

Un fait à noter est que le morphème $-(o)ŋ$ ne peut pas s'affixer à un élément précédent le verbe. Il est même nettement placé vers la fin de la phrase. Ainsi, en (232), cette marque s'affixe à un groupe postpositionnel qui suit le verbe. En (233), on voit qu'il est impossible que le pluriel s'affixe au même groupe postpositionnel si celui-ci précède le verbe, même si cette position est naturelle pour un groupe postpositionnel (234).

(232) lekolazal-a-kom o-ʔu(l)-tal e-koti-**ŋ**. exemple élicité
instituteur-a-PL 3.I-venir-FUT 1SG.II-chez-PL.S
Les instituteurs vont venir chez moi.

- (233) *e-koti-ŋ lekolazal-a-kom o-ʔu(l)-tal. exemple élicité
 1SG.II-chez-PL.S instituteur-a-PL 3.I-venir-FUT
 Les instituteurs vont venir chez moi.
- (234) e-koti lekolazal-a-kom o-ʔu(l)-tal-(oŋ). exemple élicité
 1SG.II-chez instituteur-a-PL 3.I-venir-FUT-PL.S
 Les instituteurs vont venir chez moi.

Ainsi, *-(o)ŋ* apparaît plutôt en fin de phrase. Les éléments qui peuvent le suivre sont : les compléments circonstanciels, un vocatif, ou la particule de fin de phrase *(i)keʔi ~ (i)kiʔi*.

- (235) zewe nani o-ho-ŋ i-kalug.
 tous.les.jours ainsi 3.I-aller-PL.S 3.II-crépuscule
 Tous les jours, elles y vont jusqu'au soir.
- (236) i-awu-ŋ ikiʔi.
 3.II-parole-PL.S maintenant
 Maintenant ils parlent.

Suite à l'étude de cette distribution, nous devons conclure que :

- *-(o)ŋ* est un clitique, car il n'apparaît pas toujours sur des mots de la même classe
- c'est un clitique de fin de groupe verbal, car il s'attache toujours au dernier élément de celui-ci.

Cette analyse est confirmée par le fait qu'avec une série de verbes comme en (237) et en (238) ou avec une suite verbe-gérondif à sujet pluriel comme en (239), le pluriel apparaît toujours sur le dernier verbe.

- (237) munuʔaŋ-a-nate o-hē-hem o-ho-ŋ.
 après-midi-a-POSTP 3.I-RED-sortir 3.I-aller-PL.S.
 L'après-midi, elles sortent.
- (238) imani mun-a-kom o-ho o-(w)eta beku o-ilul-oŋ.
 plusieurs personnes-a-PL 3.I-aller 3.I-couper liane.sp. 3.I-amener-PL.S
 Beaucoup de gens sont allées couper la liane et la ramener.

- (239) zewe o-ho baipuli-l-esak-**oŋ**.
 tous.les.jours 3.I-aller tapir-RELN-voir-PL.S
 Tous les jours, elles vont voir le tapir.

Le tableau suivant récapitule la distribution de $-(o)\eta$. Le groupe verbal auquel il s'affixe est mis entre crochets.

[prédicat verbal] $-(o)\eta$
[attributif] $-(o)\eta$
[prédicat nominal] $-(o)\eta$
[prédicat numéral] $-(o)\eta$
[V O] $-(o)\eta$
[V O GP] $-(o)\eta$
[V ₁ ...V _n] $-(o)\eta$

Tableau 23 : Distribution du clitique $-(o)\eta$ sur différents types de groupes verbaux

L'étude de $-(o)\eta$ constitue un argument pour l'existence d'un groupe verbal en émerillon et fait ressortir les différents types de groupes verbaux possibles. La position spécifique de ce morphème de pluriel nous fait en outre émettre une hypothèse quant à son origine.

I- 3. Origine de $-(o)\eta$

Rappelons d'abord que le morphème $-(o)\eta$ ne semble pas avoir d'équivalents dans les autres langues tupi-guarani, et qu'une analyse en termes d'emprunt ne s'impose pas non plus. Nous pensons qu'il s'agit plutôt d'une innovation émerillon.

On peut faire la supposition que ce clitique soit issu de la grammaticalisation d'un verbe en position de gérondif, ce qui expliquerait sa position à l'extrême droite

du groupe verbal¹⁰⁹. Cette idée est purement intuitive, et ne se base sur aucune piste quant à l'existence d'un verbe pouvant donner lieu à ce type d'évolution. Toutefois, ce genre de processus a déjà été décrit en tupi-guarani : Maurel (2000) rappelle la proposition de Jensen de la grammaticalisation d'un verbe "être assis" en un clitique "progressif" *-ijn* (continu, pour nous). En émérillon, l'autre morphème de pluriel du sujet a, lui, clairement connu une évolution comparable. Nous allons maintenant présenter ce morphème : *kupa*.

II- Le pluriel du sujet en *kupa*

Un autre morphème peut indiquer un sujet de 3^{ème} personne du pluriel, il s'agit de *kupa*. Dans notre corpus, son utilisation est beaucoup moins fréquente que celle de *-(o)ŋ*. De même, l'élicitation d'une forme plurielle engendre plus probablement une forme en *-(o)ŋ* qu'en *kupa*, et une phrase comportant *kupa* n'est pas toujours jugée acceptable. Dans les textes, on trouve 24 occurrences, produites par seulement 5 de nos locuteurs sur 13, c'est-à-dire dans 7 textes sur 22 (alors que *-(o)ŋ* est présent dans 16 textes sur 22). Ces locuteurs appartiennent à trois générations différentes et ont tous à un moment donné de leur vie habité au côté des Wayampis. Il est important d'ajouter à ceci que, pour les locuteurs émérillon, *kupa* "sonne" plus comme du wayampi que de l'émérillon. Pour toutes ces raisons, *kupa* a un caractère marginal dans la langue. Nous allons examiner sa forme et sa fonction (II-1) puis nous interroger sur son statut et son origine (II-2).

II- 1. Forme et fonction de *kupa*

Le morphème *kupa* partage la même fonction que *-oŋ*, avec peut-être une nuance sémantique : on a plutôt tendance à le traduire par "ensemble", donc à insister plus sur le caractère "collectif" du sujet, que sur son nombre (deux ou plus).

¹⁰⁹ Pour une description des gérondifs en proto-tupi-guarani, voir Chapitre 13, II.

- (240) ze-poli-katu o-apig-o **kupa-o.**
 RECIP-à.côté.de-bien 3.I-s'asseoir-CONT PL.S-CONT
 Elles s'assoient l'une à côté de l'autre.
- (241) oke-o-kel-o **kupa-n.**
 RED-3.I-dormir-CONT PL.S-CONT
 Ils dorment.
- (242) aʔe o-zewil-o **kupa-nam, ...**
 DEM 3.I-revenir-CONT PL.S-quand
 Quand ils revinrent, ...

Sa position dans la phrase est toujours en fin de proposition, ce qui sera expliqué en II-2.

Au niveau de sa forme, il est invariable, mais très souvent suivi de *-o* (14 occurrences sur 24). *-o* est normalement analysé comme un clitique de continu¹¹⁰. Sur *kupa*, dans certains cas comme en (240), *-o* peut tout à fait être analysé de cette manière. Dans d'autres, le sens continu n'est pas évident :

- (243) ..., eʔi baipuli-l-ie-pope-ne **kupa-o.**
 3.I.dire tapir-RELN-ventre-dans-CONTR PL.S-o
 ..., disent-elle ensemble dans le ventre du tapir.

Quatre autres occurrences de *kupa* prennent le clitique *-n* de continu, comme en (241), ce qui tendrait à nous faire analyser le *-o* également comme un continu. Enfin, quatre autres occurrences encore ne prennent pas de suffixe et précèdent immédiatement un subordonnant, comme en (242). La typologie des langues du monde nous montre qu'il est commun qu'un verbe ne prenne pas de marques temporo-aspectuelles quand il est dépendant, et que les suffixes d'aspect sur les verbes finis soient en relation paradigmatique avec les marques de subordination sur les verbes non-finis. Cette remarque typologique pourrait être un autre argument pour voir dans le *-o* une marque d'aspect continu, dont la présence alterne avec celle d'un subordonnant. En bref, que *-o* sur *kupa* soit analysé comme le clitique de continu ou non, il demeure que cette co-occurrence est fréquente et non-expliquée.

¹¹⁰ Pour une description des clitiques de continu *-o* et *-n*, voir Chapitre 12, II-1.

II- 2. Statut et origine de *kupa*

Une autre question reste en suspens, à savoir à quelle partie du discours appartient *kupa*. Ce morphème semble être accentué sur la deuxième syllabe, ce qui en ferait plutôt un mot indépendant qu'un clitique. C'est ici que les informations comparatives et diachroniques sont intéressantes.

Sur le plan diachronique, *kupa* est une forme résiduelle du verbe † *kúβ* "être ensemble"¹¹¹ suivi du suffixe de gérondif † *a* (Jensen 1989, p.104). Cette information explique indirectement pourquoi *kupa* est toujours placé en fin de proposition (c'est la position des gérondifs) et également le sémantisme plus collectif venant du sens de "être ensemble". Ces propriétés résultent de la grammaticalisation du gérondif du verbe † *kúβ* en morphème de nombre.

Sur le plan comparatif, le wayampi utilise le morphème *kupa* pour marquer le pluriel d'un sujet de 3^{ème} personne (Grenand 1980, p. 68). Ceci nous permet de spéculer que l'utilisation de *kupa* en émérillon résulte d'une interférence avec le wayampi. Cette interférence serait le résultat, chez certains Emérillons, de l'interaction quotidienne avec des locuteurs wayampi, interférence généralement rejetée et considérée comme non intégrée dans la langue par la communauté linguistique. Cette spéculation expliquerait pourquoi *kupa* est peu fréquent et peu apprécié en émérillon. Mais il est difficile de savoir quel est la nature exacte de cette interférence. Soit l'émérillon n'a jamais connu de forme *kupa*, et certains locuteurs l'utilisent en prenant comme modèle le wayampi. Soit l'émérillon avait lui aussi une forme *kupa*, dont l'utilisation a diminué quand les gérondifs ont été remplacés par la structure sérielle¹¹². Le fait que cette forme perdure en wayampi de manière apparemment fréquente serait en train de ralentir la disparition de *kupa* en émérillon.

¹¹¹ Ce verbe n'a pas de réflexe en émérillon.

¹¹² L'évolution des gérondifs en structure sérielle est discutée au Chapitre 15, II-4.

Nous espérons ainsi avoir donné quelques éclaircissements sur *kupa*, une marque de nombre marginale en émérillon. Nous portons maintenant attention au clitique *kom* qui est une marque de pluriel omniprésente et à l'histoire connue.

III- Le pluriel en *-kom*

Le morphème *-kom* est très courant car il connaît une très large distribution. Après avoir vu ses différentes fonctions et discuté son statut, nous donnerons son origine : c'est en effet un emprunt aux langues caribes.

III- 1. Fonctions de *-kom*

Les deux usages principaux de *kom* sont le pluriel de la tête de syntagme nominal (nom, pronom personnel ou démonstratif, ...) et le pluriel des compléments (possesseur du nom, objet de postposition, objet du verbe). Suffixé à un syntagme nominal, il peut aussi jouer un rôle de coordinateur. Notons que dans tous ces cas, il est facultatif.

La première grande fonction de *-kom* est de pluraliser différents types de syntagmes nominaux : un nom (244), un pronom personnel (245), un pronom démonstratif (246), une nominalisation (247) et une proposition relative sans tête (248). Le suffixe *-a*¹¹³ apparaît alors sur la tête de syntagme à finale consonantique, quelque soit la fonction de ce syntagme (comme devant les particules, cf. Chapitre 4, I).

(244) wãĩwĩ-**kom**-enam o-wikilog o-wikilog.
 femme-PL-chang.top 3.I-vider 3.I-vider
 Les femmes vident les poissons.

(245) k^wi-ena olone-**kom**-uwe pe-l-upi olo-ho-tal-uwe.
 un.jour-peut-être PRO1EXCL-PL-aussi chemin-RELN-sur 1EXCL.I-aller-FUT-aussi
 Un jour peut-être, nous aussi, nous irons aussi sur le chemin.

¹¹³ Le chapitre précédant est en partie dédié à l'analyse de ce *-a*.

- (246) wɨŋ-a-**kom**-āhā o-wɨl o-ho.
DEM-a-**PL**-seulement 3.I-monter 3.I-aller
Eux seulement montent.
- (247) la.mairie o-meʔeŋ kalakuli [ɨal-a-maʔē-hal]-a-**kom**-a-pe. phrase élicitée
mairie 3.I-donner argent pirogue-a-regarder-NOMN-a-PL-a-à
La mairie donne de l'argent à ceux qui gardent les pirogues.
- (248) [o-ʔi-wi-wi o-ʔul-a-maʔē]-**kom** amō
3.COREF-fleuve-ABL-RED 3.I-venir-a-REL-PL autre
les autres, qui viennent de leur fleuve...

Dans cette même position de fin de syntagme nominal, *kom* est souvent utilisé pour coordonner une liste de noms. La phrase suivante décrit divers types de poissons qui remontent à la surface après que l'on ait répandu le jus de la liane à nivrée dans un bassin. Il est évident qu'il ne s'agit pas d'un seul poisson de chaque espèce et de plusieurs représentants de la dernière (celle dont le nom porte le clitique *-kom*). Ici, *kom* n'indique donc pas le pluriel de *pakupitaŋ*, mais coordonne les trois noms de poissons.

- (249) uluwi, taleʔil, pakupitaŋ-a-**kom** ð-wem.
torche.argentée aïmara pacou.sp.-a-PL 3.I-sortir
Les torches argentées, les aïmaras et les pacous remontent.

Cette coordination fonctionne avec des noms pluriels ou singuliers. Dans la phrase suivante, deux noms au singulier sont coordonnés par l'affixation de *kom* à l'ensemble. Dans le cadre de l'Histoire de la Grenouille, cette phrase ne peut être comprise que comme en 1), avec *pulelu zawalakom* formant un syntagme complexe en position d'objet de la postposition *ehe*. Le sujet est alors "le garçon" repris dans la phrase par la marque de sujet sur le verbe *o-*. Hors contexte, la phrase pourrait être traduite comme en 2) si l'on interprète *kom* comme le pluriel de *zawal*, et donc *pulelu* comme le sujet de la phrase.

- (250) pulelu zawal-a-**kom**-a-l-ehe o-maʔē-maʔē kiʔi.
crapaud chien-a-PL-a-RELN-POSTP 3.I-RED-regarder alors
1) Il regarde alors le crapaud et le chien.
2) Le crapaud regarde alors les chiens.

La deuxième grande fonction de *-kom* est de pluraliser le complément de trois grands types de racine : nom, postposition et verbe. Il pluralise leur complément de 3^{ème} personne (généralement marqué par la série II¹¹⁴), à savoir respectivement le possesseur du nom, l'objet de la postposition et l'objet du verbe.

- (251) *zawal, pulelu, i-ɖʒal-a-kom o-k^wa-n.* possesseur du nom
 chien crapaud 3.I-maître-a-PL 3.I-partir-CONT
 Le chien, le crapaud et leur maître partent.
- (252) *o-wig Ø-ehe-kom.* objet de postposition
 3.I-arriver 3.II-à-PL
 Il arrive près d'elles.
- (253) *wɨŋ-a-te o-zika-kom.* pluriel de l'objet
 DEM-a-FOC 3.I-tuer-PL
 C'est lui qui les a tués. phrase élicitée

Les exemples de *-kom* sur le verbe, tel (253), sont rares. Il faut noter que contrairement aux occurrences avec les noms, la racine précédente (qu'elle ait une finale vocalique ou consonantique) ne prend pas le suffixe *-a* puisqu'il s'agit du prédicat. Le marquage du pluriel de l'objet sur le verbe est indépendant du marquage du sujet ou de l'objet sous forme d'indice préfixé au verbe.

Le tableau suivant récapitule l'élément que *-kom* peut pluraliser :

¹¹⁴ Le possesseur du nom et l'objet de postposition peuvent aussi être marqués par le *o-* coréférentiel réfléchi, et l'objet de postposition par le *ze-* coréférentiel réciproque (cf. Chapitre 17). Le verbe prend l'indice de sujet ou d'objet selon les hiérarchies des personnes et des rôles sémantiques (cf. Chapitre III).

SN-kom	(IP)-N-(a)-kom	pluralisation de la tête de syntagme nominal
	PRO-kom ¹¹⁵	
	DEM-(a)-kom	
	NOMN-(a)-kom	
	REL-(a)-kom	
	[N ₁ N ₂ ...N _n -(a)]-kom	coordination des différents noms du syntagme
IP-Racine-kom	IP-N-(a)-kom	pluralisation du complément de 3 ^{ème} personne
	IP-POSTP-kom	
	IP-V-kom	

Tableau 24 : Distribution du pluralisateur -kom

Ainsi, quand *-kom* s'ajoute à un nom précédé d'un indice possessif de troisième personne (dans les configurations en gras dans le tableau), il y a une ambiguïté possible. Hors contexte, il peut soit pluraliser soit le nom (tête de syntagme nominal), soit le possesseur (complément de la racine nominale), soit les deux :

- (254) *i-kiwil-a-kom*
 3.II-frère-a-PL
 a) ses frères
 b) leur frère
 c) leurs frères

En appliquant cette remarque à l'exemple (251), *i-dʒal-a-kom* pourrait hors contexte être traduit "ses maîtres". Le contexte suffit souvent à lever l'ambiguïté : dans l'Histoire de la Grenouille, il n'y a que deux animaux domestiques avec un seul et même maître.

Notons que, de manière complètement étrangère à la logique du système présenté ci-dessus, certains verbes intransitifs prennent le *-kom*. Maurel (1998)

donne, dans son tableau des marques personnelles sujet de l'émérillon, *si- ...-kom* pour la 1^{ère} personne inclusive. Notre corpus révèle de tels exemples, avec comme sujet *si-* (1^{ère} inclusive) et *za-* (personne indéterminée). Nous considérons cet usage comme une exception.

(255) aʔe-kom nōde-kom se¹¹⁶-iko-kom.
 DEM-PL PRO1INCL-PL 1INCL.I-être-PL
 Ces personnes-là (les survivants des guerres), c'est nous.

(256) za-ɕʒapiaka-kom napial-ehe.
 INDET.I-penser-PL chose-à
 On pense à ce genre de choses.

Nous avons vu la large et diverse distribution de *-kom* et expliqué sa fonction. Nous pouvons maintenant discuter de son statut.

III- 2. Statut de *-kom* : un clitique

Il est difficile de définir précisément le statut de *-kom*. Sur les noms et les postpositions, il est généralement plutôt produit comme une partie du mot (sauf quand il sert de coordination, où il est alors séparé). Etant donné qu'il s'accroche à diverses parties du discours, il doit être considéré comme un clitique. Toutefois, lorsqu'il porte sur le verbe, il est plutôt séparé du verbe et semble porter un accent. Il est donc difficile de préciser la partie du discours à laquelle appartient *kom*.

L'hypothèse d'analyser *kom* comme un nom signifiant "ensemble, groupe" a été proposée par Queixalós (2001c). Après un nom, *kom* serait donc la tête de syntagme nominal et le nom "pluralisé" serait son modifieur génitival : *wãĩwĩ-kom* "les femmes" devrait être compris littéralement comme "l'ensemble des femmes". L'auteur n'explique pas comment il faudrait alors analyser les cas où *kom* pluralise le possesseur du nom et l'objet d'une postposition. Selon cette analyse, il serait possible de voir dans nos exemples de coordination (249) et (250) un syntagme

¹¹⁵ On ne trouve jamais le *-a* entre un pronom et *-kom* car les pronoms ont tous une finale vocalique.

¹¹⁶ *si-* est réalisé /se-/ devant un /i/.

génitif dont la tête serait le nom *kom* "ensemble, groupe", et le déterminant l'ensemble de la liste de noms le précédant, coordonnés par simple juxtaposition. Ainsi, la phrase (249) serait à traduire littéralement "l'ensemble des torches, aimaras et pacous remontent à la surface". Après un verbe, *kom* constituerait un syntagme nominal objet indépendant, ce que semble corroborer la réalité prosodique. Il n'est cependant pas dans la position "préférée" de l'objet qui est préverbale.

En conclusion, l'analyse de *kom* comme clitique ou comme nom n'explique pas la variété de ses utilisations, variété qui est aussi importante dans sa famille de langues d'origine, comme nous allons le voir maintenant.

III- 3. Origine de *-kom* : un emprunt aux langues caribes

Jensen (1979, cité dans Dietrich 1990) et Couchili, Maurel et Queixalós (2002) reconnaissent dans *kom* un emprunt aux langues caribes. L'emprunt de ce morphème a aussi eu lieu en wayampi (*kũ*, sur les noms, Grenand 1980), et en jo'é (*kã*, marque collective sur les noms référentiels, Cabral, c.p.).

D'après les reconstructions du proto-caribe de Gildea (1998), † *komo* suivait un nom possédé pour marquer le nombre du possesseur (préfixe possessif ou groupe nominal précédant le nom déterminé)¹¹⁷. Sur les verbes nominalisés, il pluralisait le possesseur absolutif, et est devenu une marque post-verbale de nombre quand les formes nominalisées ont été réanalysées comme des verbes finis. Au niveau sémantique, cette particule exprime plus précisément le collectif, et oppose ainsi "tous" à "moins que tous". La description proposée par Lescure (c.p.) du *kon* en kali'na correspond à cette reconstruction. Elle précise que sur un nom possédé, *kon* réfère au possesseur, mais peut parfois indiquer le nombre du possédé.

¹¹⁷ Hoff (1968) indique les mêmes fonctions pour *koŋ*, mais le présente comme un des nombreux allomorphes de *ne*.

En émérillon, *kom* a donc gardé toutes ces fonctions : marquer le nombre du possesseur, du nom possédé, de l'objet du verbe (et du sujet apparemment seulement avec *si-* et *za-*), fonctions auxquelles a été ajoutée celle d'indiquer le nombre des objets de postposition. Enfin, il nous est difficile de spécifier son sens comme "pluralisateur" ou comme "collectivisateur" à cette étape de notre enquête.

Nous venons de présenter les trois marques de nombre de l'émérillon. Rappelons que le nombre n'est pas habituellement une catégorie très marquée dans les langues d'Amazonie. Et de fait, aucun des trois morphèmes de pluriel émérillon n'est un réflexe du proto-tupi-guarani, ni n'a la même origine diachronique. L'un d'entre eux, le clitique de fin de groupe verbal $-(o)\eta$, qui pluralise le sujet, est probablement une innovation émérillon. Un autre pluralisateur de sujet, *kupa*, est minoritaire dans la langue et peut-être une interférence du wayampi. Enfin, le clitique $-kom$, qui est utilisé très fréquemment aussi bien comme pluralisateur de tête de syntagme nominal que comme pluralisateur de complément, est un emprunt aux langues caribes.

Conclusion

Cette présentation des marques de pluriel clôt la partie II consacrée à la morphologie essentielle de l'émérillon, celle que l'on retrouve fréquemment et/ou avec une large distribution. Les marques de personne, le suffixe $-a$ et le préfixe relationnel $l-$ font de l'émérillon une langue tupi-guarani assez typique. Par contre, son système de marquage du pluriel est plus "original", étant donné que l'ensemble de la famille tupi-guarani ne partage aucune marque de pluriel. Maintenant qu'est posée la distinction entre, d'une part, le nom et le verbe émérillon, et d'autre part, entre l'argument et le prédicat, et que l'essentiel de la morphologie de base de l'émérillon a été présenté, nous allons nous intéresser à la structure des constituants de niveaux supérieurs qui remplissent les fonctions argumentales et prédicatives

dans cette langue. La partie III sera consacrée au syntagme nominal, et la partie IV à la structure de la phrase simple.