

L'objectif que nous nous étions fixé pour ce travail était d'offrir une description la plus complète possible de la grammaire de l'émérillon. Cette thèse constitue ainsi la première description d'envergure de cette langue, qui nous l'espérons pourra servir de référence pour de futures recherches. Quoique le cœur de ce travail ait trait à la morphosyntaxe, des aspects aussi éloignés que la phonétique ou l'organisation du discours sont aussi abordés.

L'analyse de la langue a été réalisée dans le cadre de la linguistique fonctionnelle et typologique. La problématique qui a été privilégiée consistait, à partir des descriptions produites, à situer l'émérillon dans le contexte plus large des langues du monde représentées par les études typologiques et dans celui plus restreint des langues tupi-guarani. Un troisième axe a ponctuellement été exploité : celui du caractère "amazonien" de l'émérillon.

Nous concluons maintenant sur la caractérisation de l'émérillon en termes typologiques, sur sa spécificité parmi les langues tupi-guarani, et sur sa conformité aux propriétés communes des langues amazoniennes.

## **I- L'émérillon et la typologie des langues du monde**

Cette partie veut caractériser typologiquement l'émérillon et souligner ses particularismes qui peuvent venir enrichir les études typologiques déjà existantes.

La typologie caractérise souvent les langues par le degré selon lequel les mots sont segmentables en morphèmes clairement individués (cf. Comrie 1981, p.39-50). Cette typologie morphologique suit deux paramètres : la synthèse et la fusion.

- D'une part, l'indice de synthèse indique combien de morphèmes forment un mot. Les langues sont placées sur un continuum entre langues isolantes (où

chaque mot est un seul morphème) et langues polysynthétiques (où les mots consistent en plusieurs morphèmes lexicaux ou grammaticaux).

Sur ce continuum, l'émérillon tend vers le côté polysynthétique : le prédicat est au minimum constitué de la racine et d'un indice de personne, mais peut aussi incorporer un nom, prendre de nombreux préfixes et suffixes auxquels s'ajoutent des clitiques. Comme les arguments sous forme de syntagmes nominaux ne sont pas obligatoires, le prédicat ainsi formé peut constituer à lui seul une phrase.

- D'autre part, l'indice de fusion indique le degré avec lequel les morphèmes à l'intérieur du mot sont segmentables et invariables. Les langues sont placées sur un continuum entre langues fusionnelles (ou flexionnelles, dans lesquelles un morphème non segmentable exprime plusieurs catégories grammaticales) et langues agglutinantes (où les mots sont constitués de plusieurs morphèmes faciles à segmenter, peu variables, et exprimant une seule catégorie).

Sur ce continuum-là, l'émérillon est une langue agglutinante. Les morphèmes de l'émérillon sont facilement segmentables et ne fusionnent pas plusieurs catégories. Cette caractérisation de l'émérillon comme langue à tendance agglutinante et polysynthétique est peu surprenante car il s'agit d'une situation typique pour une langue d'Amazonie.

Deux autres caractéristiques typologiques de l'émérillon sont également répandues parmi les langues d'Amazonie.

La première est que l'émérillon est à la fois "head-marking" et ce que Jelinek (1988) appelle une langue "à argument pronominal". En effet, l'émérillon ne nécessite pas de syntagmes nominaux pleins pour exprimer les arguments d'un prédicat. Les relations entre arguments et prédicat sont marquées sur le prédicat par un indice de personne, et non sur les arguments.

L'autre caractéristique typologique et amazonienne est que l'ordre des mots est flexible et dépend à la fois du rôle pragmatique des arguments (topique, focus) et

de la présence de particules de seconde position. Dans une phrase pragmatiquement neutre, il semble que l'ordre préféré soit SOV, l'ordre SV étant strict. Les groupes postpositionnels sont en position périphérique, soit finale, soit initiale. Les autres paramètres (postsposition, place de la subordonnée avant la principale, place des modificateurs avant le nom) renforcent l'analyse de l'émérillon comme une langue à verbe final.

Mais la caractéristique de l'émérillon la plus intéressante typologiquement est le système d'indexation des personnes sur le prédicat. En effet, l'émérillon connaît un système d'accord mineur dans les langues du monde : un système hiérarchique. L'indice sélectionné par le prédicat indique soit le sujet (indice de série I) soit l'objet (indice de série II), selon la position relative de ces deux arguments sur une hiérarchie des personnes  $1/2 > 3$  et sur une hiérarchie des rôles sémantiques  $A > P$ . Ce système n'est ni répandu dans les langues du monde, ni caractéristique des langues amazoniennes où l'ergativité est importante. Il s'agit plutôt d'une spécificité tupi-guarani.

L'émérillon partage avec les autres langues tupi-guarani une autre caractéristique typologique singulière : certains noms prédisent avec un sens possessif en utilisant exactement la même morphologie que les verbes en fonction de prédicat.

Enfin, l'émérillon connaît quelques caractéristiques intéressantes qui lui sont particulières, notamment l'origine de la sérialisation verbale dans une construction dépendante (le gérondif), et une structure possessive unique des noms référant à des animaux domestiques qui peut être analysée comme un cas de classificateur nominal unique dans la langue.

## II- L'émérillon comme langue tupi-guarani

L'émérillon est classé parmi le 8<sup>ème</sup> sous-groupe des langues tupi-guarani (Rodrigues 1984-1985, Rodrigues et Cabral 2002). Les langues de cette famille étant caractérisées par une grande similarité aux niveaux lexical et morphologique, leur comparaison est aisée, et leurs divergences frappantes.

Dans l'ensemble, l'émérillon possède les principales caractéristiques morphosyntaxiques de la famille tupi-guarani. Il conserve les deux séries majeures d'indices de personne et le système d'accord hiérarchique (certains détails étant modifiés), le suffixe référential *-a* dans certains contextes seulement et le préfixe relationnel typiques de cette famille. Ses principaux démonstratifs, postpositions, numéraux et morphèmes de voix sont des cognats des morphèmes similaires du reste de la famille. Sur le plan syntaxique, l'émérillon est comme les autres langues tupi-guarani une langue où le prédicat porte obligatoirement un indice de personne mais peut facilement se passer d'arguments exprimés par des noms pleins.

Dans le détail de ces systèmes conservés par l'émérillon, des modifications sont cependant observées par rapport aux autres langues tupi-guarani : la création d'indices de personne indéterminée, le marquage complexe de la configuration où une seconde personne agit sur une première personne, l'absence du *-a* sur les arguments sujet et objet, la perte du morphème de réciproque et sa prise en charge par le morphème de réfléchi (et conséquemment, la simplification du marquage de la coréférentialité).

Enfin, certains systèmes tupi-guarani ont été largement modifiés par l'émérillon. Ainsi, seul l'émérillon scinde les mots dits "descriptifs" en deux classes, l'une comparable aux noms : celle des nominoïdes ; l'autre assimilable ni à des noms, ni à des verbes, ni à des adjectifs : celle des attributifs. De même, seul l'émérillon connaît la réduplication des syllabes initiales d'un verbe. La réduplication dans les autres langues tupi-guarani est décrite comme affectant les syllabes finales. Par

ailleurs, une évolution que l'émérillon (et quelques autres langues de la famille) a connue est la perte des systèmes d'indexation de personne absolutif et coréférentiel dans les formes verbales dépendantes, et de la plupart des marques de dépendance, comme par exemple la marque d'oblique topicalisé sur le prédicat. Enfin, l'émérillon a développé de nouvelles subordonnées et modifié certaines constructions gérondives en constructions sérielles.

Ces remarques faites, il est clair que l'émérillon est un membre de la famille tupi-guarani non marginal. Il possède cependant des particularités qui en font un élément intéressant et sans doute innovant. Cela dit, il ne s'écarte pas de manière marquée des autres langues du 8<sup>ème</sup> sous-groupe. Ainsi, la plupart des innovations connues par l'émérillon se retrouve également en wayampi. Il sera intéressant de comparer l'émérillon au jo'é sur le plan morphosyntaxique quand la description de cette langue le permettra. Enfin, précisons que nous n'avons pas repéré de phénomènes d'ordre syntaxique qui joueraient un rôle dans la diversification dialectale de l'émérillon<sup>342</sup>. Nous espérons que la description de l'émérillon et les remarques comparatives proposées dans ce travail participeront à l'amélioration toujours en cours de la classification interne de la famille tupi-guarani.

### III- L'émérillon comme langue amazonienne

La linguistique amazonienne a connu un essor considérable dans les années 1980-1990.<sup>343</sup> Elle a offert aux typologues une source d'informations importante sur les langues amazoniennes alors peu connues. Il existe donc une littérature "aréale" ou "régionale", dont les ouvrages les plus connus sont sans doute le *Handbook of*

---

<sup>342</sup> Seules des variantes phonétiques de certains phonèmes selon la zone d'habitation (est ou ouest) ont été notées, ainsi qu'une variante *kōde-* du morphème *nōde-* "1INCL.II" qui est spécifique à la zone ouest.

<sup>343</sup> Cette dynamique continue, comme l'attestent des ouvrages récents tels que Queixalós et Renault-Lescure (eds), *As línguas amazônicas hoje* (2000) ou des actes de colloques organisés au Brésil tels que Cabral et Rodrigues (eds), *Línguas Indígenas Brasileiras. Fonologia, Gramática e História. Atas do I Encontro Internacional do GTLI da ANPOLL* (2002).

*Amazonian Languages* édité par Derbyshire et Pullum (1986a et 1988), *Amazonian Linguistics, Studies in Lowland South American Languages* édité par Doris Payne (1990) et *The Amazonian Languages* édité par Dixon et Aikhenvald (1999a). Ces ouvrages démontrent qu'il n'est pas anodin de parler de "langues amazoniennes". Ils soulignent en effet qu'il existe des similarités dans la structure grammaticale et phonologique de ces langues dépassant les groupes génétiques actuellement postulés (par exemple Derbyshire et Pullum, 1986b, p.1).

Ainsi, par exemple, David Payne (1990) donne cinq formes largement répandues en Amérique du Sud : une particule négative /ma/, un affixe causatif /mV/, un préfixe verbal causatif /V-/, un suffixe verbal directionnel ou cas locatif /-pV/ et un auxiliaire /ka/ qui coexiste dans de nombreuses langues avec un verbe lexical "dire, travailler", et/ou des affixes de changement de la valence phonologiquement proches. Ces formes "pan-américaines" sont expliquées par la diffusion plutôt que par l'héritage génétique commun. L'émérillon possède deux morphèmes qui peuvent leur être liés : le causatif *bo-* (dont la réalisation nasale est *mõ-*) doit certainement être rapproché de l'affixe causatif /mV/, et de manière plus hypothétique, le causatif émérillon *okal* pourrait être lié au morphème /ka/ avec cependant la seule fonction de changement de la valence.

Cette dernière partie de notre conclusion va permettre de confirmer le caractère typiquement amazonien de l'émérillon. Nous reprenons pour cela les caractéristiques générales des langues amazoniennes mises en avant par Derbyshire et Pullum (1986b), Doris Payne (1990), et Dixon et Aikhenvald (1999b). De manière intéressante, ces trois articles, tirés des ouvrages sur les langues amazoniennes précédemment cités, se recoupent rarement : l'article de Derbyshire et Pullum liste 10 propriétés essentiellement syntaxiques, celui de Doris Payne en liste 8 qui dressent la typologie morphologique de ces langues, et celui de Dixon et Aikhenvald propose 15 propriétés, dont certaines répètent celles des articles précédents et ne seront donc pas reprises ici, les autres se focalisant sur des points

plus précis de la morphosyntaxe des langues d'Amazonie. Au total, nous présentons donc 26 propriétés des langues d'Amazonie qui nous permettent de situer l'émérillon dans ce contexte "régional".

La première série de 10 généralisations reprises ici et commentées pour l'émérillon est issue de Derbyshire et Pullum (1986b, p.19).

- (1) "Verb agreement with both subject and object (cf. e.g. the Bantu languages) plus null realization of subject and object pronouns, making it extremely common for a sentence to have no full NP subject or object."

Cette caractérisation est tout à fait juste pour l'émérillon. Rappelons que la série I marque le sujet sur le verbe, et la série II l'objet, mais qu'un seul indice est présent sur le verbe. Le choix de l'indice en question se fait selon les hiérarchies des personnes  $1/2 > 3$  et des rôles sémantiques  $A > P$ . De plus, les syntagmes nominaux pleins en fonction d'argument ne sont pas obligatoires et les pronoms personnels ne sont utilisés que pour marquer une emphase particulière.

- (2) "Predictability of whether subject and object will be realized as full NPs or merely indicated by verbal affixes, depending on whether they represent given or new information."

Nous avons en effet vu que les arguments d'un verbe sont très rarement exprimés sous forme de syntagmes nominaux pleins. S'ils le sont, c'est effectivement pour introduire leur référent dans le discours.

- (3) "Use of nominalizations to substitute for finite relative clauses and other types of subordinate clauses..."

Ceci définit parfaitement la situation du proto-tupi-guarani. Cependant, en émérillon, les propositions "dépendantes" ont subi un changement dont le résultat consiste en des subordonnées finies, relatives ou circonstancielles.

- (4) "A combination of phrase constituent orders which, while not unique, is unusual, and seems to be found regardless of whether main clauses have VO or OV constituent order: N-Adj, Gen-N, and NP-P."

En effet, l'émérillon a bien un génitif préposé au nom modifié et des postpositions. Pour autant, l'ordre N-Adj n'est pas applicable à l'émérillon, qui ne connaît pas d'adjectif en tant que tel. Par contre, un attributif, une fois intégré dans une relative, peut servir à modifier un nom. Mais l'attributif précède alors le nom, qui est suivi du relativisateur.

- (5) "No construction identifiable as an agentive passive."

Cette propriété s'applique telle quelle à l'émérillon.

- (6) "No indirect statement form, hence heavy reliance on direct rather than indirect speech constructions."

L'émérillon ne connaît en effet pas le discours indirect.

- (7) "No coordinating conjunctions, hence reliance on juxtaposition to express logical coordination."

Nous avons vu qu'il existe une seule conjonction en émérillon : *o?olam*. La juxtaposition est donc importante, même si d'autres outils servent à souligner la relation de coordination (connecteur de discours *kol*, particules, pluriel *-kom*, démonstratif comme objet de postposition).

- (8) "Extensive use of right-dislocated paratactic constructions..."

Notre corpus montre quelques exemples de telles dislocations, sans que l'on puisse considérer leur usage comme répandu.

- (9) "Extensive use of particles that are phrasal subconstituents syntactically and phonologically but are sentence operators or modifiers semantically."

Les particules de l'émérillon ont largement été discutées et présentées. Ce sont des clitics de constituant (souvent du premier constituant de la proposition). Leur

fonction couvre des thèmes très variés comme le mode, la quantification, les rôles pragmatiques... Leur quantité et leur usage est remarquablement important.

(10) "Ergative subject-marking, or relics of such marking."

En tupi-guarani, les formes verbales dépendantes sont caractérisées par un marquage absolutif. Il ne s'agit pas à proprement parler d'un système ergatif, le sujet d'un verbe transitif n'étant jamais indexé sur un verbe dépendant, mais d'un système absolutif : sur un verbe dépendant, c'est le sujet qui est marqué si le verbe est intransitif, et l'objet si le verbe est transitif, les deux avec des indices de série II. Ce système dit "absolutif" a quasiment disparu en émérillon, où seules les quelques constructions gérondives et les rares nominalisations le maintiennent.

Nous donnons maintenant quelques autres propriétés des langues amazoniennes listées par Doris Payne (1990) et discutons leur application à l'émérillon.

(11) "Indigenous SA languages are dominantly polysynthetic."

L'émérillon, sur l'échelle de synthèse entre langue isolante et langue polysynthétique, penche en effet plus du côté polysynthétique. La plupart des catégories grammaticales sont marquées sur le prédicat par des affixes qui peuvent facilement se cumuler. Chaque affixe a un sens propre, non fusionné, et peut être facilement segmenté.

(12) "Suffixation predominates over prefixation in the highly polysynthetic lowland SA languages. [...] Tupi-Guarani languages tend toward having a greater number of prefixes than do most Western Amazonian families. [...] Verb agreement is frequently prefixal".

En émérillon, les préfixes incluent le préfixe de négation, les indices de personne et trois morphèmes de voix. Les suffixes incluent les affixes de TAM, le morphème de voix causative indirecte, et le suffixe de négation. Notons que

l'intégralité des clitiques (affixes de constituants ou particules) s'ajoute à l'extrême droite des constituants concernés.

- (13) "Non-additive morphological processes are found throughout lowland SA, but none is particularly characteristic of the region."

Aucun procédé de ce genre n'a été repéré en émérillon.

- (14) "Reduplication is not uncommon; nevertheless it is not as striking a characteristic of this region as it is, say, of Oceanic languages."

La reduplication est fortement utilisée en émérillon, selon les deux patrons monosyllabique et dissyllabique, et avec comme fonction l'expression de la répétition.

- (15) "Compounding and incorporation are common throughout the region".

La composition ne s'est pas révélée très importante dans notre corpus. Il est vrai que celui-ci contient peu de vocabulaire spécifique comme les noms de plantes ou de végétaux, sûrement plus susceptibles d'être composés. L'incorporation d'un objet possédé dans le verbe est par contre bien attestée.

- (16) "Nominal morphology is much less complex than verbal morphology in lowland SA languages."

Les noms émérillons portent quelques suffixes et servent d'objets de postposition. Leur morphologie est beaucoup moins riche que celle des verbes.

- (17) "Many languages of this region lack, or else have an extremely small set of morphologically simple adjectives".

L'émérillon ne possède pas de classe d'adjectifs à proprement parler. Deux catégories de mots sont sémantiquement équivalentes aux adjectifs d'autres langues : les nominoïdes et les attributifs.

- (18) "Common in the western Amazonian region are morphemes which indicate the type of movement accompanying the action of the verb, the location of

the action, or the directional orientation of the action or agent carrying out the action."

Même si les langues tupi-guarani ne sont pas concernées par cette généralisation, on trouve cependant en émérillon des indications systématiques de direction données par un verbe de mouvement dans une série verbale.

Enfin, voici d'autres remarques générales sur les langues d'Amazonie présentées dans Dixon et Aikhenvald (1999b, p.8-9)<sup>344</sup>.

(19) "Many languages have extensive classifier and/or gender systems."

Ceci ne s'applique pas à l'émérillon, qui ne connaît pas la distinction de genre et ne possède qu'un seul morphème susceptible d'être considéré comme un classificateur : le mot *iba* "animal domestique", obligatoire pour marquer le possesseur d'un animal.

(20) "There are very few oblique cases – often just a locative and an instrumental/comitative."

Nous n'avons pas analysé les marques d'oblique en termes de "cas", mais de "postpositions". Les postpositions de l'émérillon sont relativement nombreuses.

(21) "The bound pronominal forms marking a possessor within an NP are typically the same as one of the bound pronominal paradigms for marking core arguments of a clause."

En effet, en émérillon, ce sont les indices de série II qui servent à la fois à marquer un possesseur dans un syntagme nominal et l'objet sur le verbe.

---

<sup>344</sup> On ne reprend ici que les remarques qui diffèrent de celles déjà données ci-dessus.

- (22) "If there are several prefix positions, the bound pronominal prefix(es) will typically appear further from the root than prefixes that mark valency-changing derivations."

C'est bien le cas en émérillon, où les préfixes de changement de valence apparaissent entre les indices de personne et la racine. Cette propriété semble en fait plus relever de la typologie générale des langues agglutinantes que d'une spécificité amazonienne.

- (23) "Most verbal categories (e.g. tense, aspect, modality, direction) are expressed through optional suffixes."

Les suffixes de TAM sont bien optionnels en émérillon. La direction est exprimée par des constructions sérielles.

- (24) "If there is noun incorporation, typically only those nouns which are obligatorily possessed can be incorporated, and they typically precede the verb root."

L'incorporation nominale en émérillon a été décrite exactement ainsi.

- (25) "In many languages adverbs and adpositions may be incorporated into the verb, typically following the verb root."

Ce phénomène n'est pas attesté en émérillon.

- (26) "There is generally only one small class of lexical numbers."

La numération de l'émérillon est en effet réduite. Les chiffres jusqu'à 4 ou 5 sont utilisés. Au delà, l'usage d'emprunts du français est quasiment automatique.

Cette comparaison systématique entre les propriétés reconnues comme typiques des langues amazoniennes et les propriétés de l'émérillon montre à quel point les propriétés "régionales" sont effectivement partagées : l'émérillon confirme la plupart de ces généralisations. Quand il ne le fait pas, c'est souvent parce qu'il a connu une évolution particulière au sein de la famille tupi-guarani.

Avant de clore cette conclusion, reprenons les points les plus intéressants de la langue émérillon. Les premiers sont les thèmes où l'émérillon contraste avec les études typologiques actuelles, par exemple le système hiérarchique d'indexation des personnes, encore peu discuté typologiquement, les prédications nominales possessives avec la même morphologie sur les verbes et les noms qui prédisent, ou l'origine des constructions sérielles dans les constructions gérondives. D'autres thèmes sont apparus comme particulièrement remarquables sur le plan comparatif, soulignant des divergences entre l'émérillon et les autres langues de la famille tupi-guarani, par exemple l'existence de deux catégories de mots descriptifs en émérillon, le marquage des personnes quand une deuxième personne agit sur une première personne ou la reduplication des syllabes initiales et non des syllabes finales. Sur le plan "aréal", de nombreuses analyses ont montré de manière saillante que la grammaire de l'émérillon concordait tout à fait avec les caractéristiques attribuées aux langues amazoniennes, ce qui confirme la validité de l'existence d'une linguistique "aréale" d'Amazonie. Finalement, quelques autres thèmes se sont révélés plus riches que ce que les descriptions disponibles des langues tupi-guarani ou des langues d'Amazonie auraient laissé supposer à priori, par exemple les idéophones, la reduplication ou la morphologie de pluriel.

Ces conclusions tirées, nous espérons que le présent travail constitue une base pour de futures recherches visant à consolider et élargir cette description de la langue émérillon. Dans l'avenir, la constitution d'un corpus plus important, incluant notamment des conversations, pourra permettre de confirmer et de préciser les analyses proposées dans ce travail (par exemple pour la description des configurations où les personnes de l'interlocution agissent l'une sur l'autre ou celle des déictiques...). Sur le plan de la linguistique amérindienne, nous souhaitons que ce travail serve à enrichir les recherches sur la famille tupi-guarani et sur

l'ensemble des langues d'Amazonie. Sur un autre plan, ce travail peut constituer un outil utile pour une plus ample documentation de la langue et de la culture émérillon en facilitant la transcription de données. Enfin, de cette thèse pourra être extrait un ouvrage non spécialisé accessible aux acteurs de l'éducation bilingue et à la communauté émérillon.