



## Géolinguistique

15 | 2015

La géographie linguistique au Brésil

---

### *L'Atlas linguistique sonore des langues indigènes du Brésil : un projet en cours*

*The Sound Linguistic Atlas of Indigenous Languages in Brazil: An Ongoing Project*

Ana Suelly Arruda Câmara Cabral, Abdelhak Razky, Ariel Pheula do Couto e Silva, Jorge Domingues Lopes et Tabita Fernandes da Silva

---



#### Édition électronique

URL : <http://journals.openedition.org/geolinguistique/670>

DOI : 10.4000/geolinguistique.670

ISSN : 2650-8176

#### Éditeur

UGA Éditions/Université Grenoble Alpes

#### Édition imprimée

Date de publication : 1 décembre 2015

Pagination : 215-232

ISBN : 978-2-84310-317-9

ISSN : 0761-9081

#### Référence électronique

Ana Suelly Arruda Câmara Cabral, Abdelhak Razky, Ariel Pheula do Couto e Silva, Jorge Domingues Lopes et Tabita Fernandes da Silva, « *L'Atlas linguistique sonore des langues indigènes du Brésil : un projet en cours* », *Géolinguistique* [En ligne], 15 | 2015, mis en ligne le 15 février 2019, consulté le 13 septembre 2019. URL : <http://journals.openedition.org/geolinguistique/670> ; DOI : 10.4000/geolinguistique.670

---

# ***L'Atlas linguistique sonore des langues indigènes du Brésil : un projet en cours***

---

Ana Suelly Arruda Câmara Cabral

*Universidade de Brasília – CNPq (Brésil)*

Abdelhak Razky

*Universidade Federal do Pará – CNPq (Brésil)*

Ariel Pheula do Couto e Silva

*Universidade de Brasília – CAPES (Brésil)*

Jorge Domingues Lopes, Tabita Fernandes da Silva

*Universidade Federal do Pará (Brésil)*

## **Résumé**

Cet article offre un panorama des langues indigènes du Brésil classées d'après des critères génétiques en familles linguistiques distribuées selon des critères géographiques. Il s'agit de fournir une idée de l'état de vitalité de ces langues, tout en mettant l'accent sur l'importance des études dialectologiques, historiques, comparatives et sociolinguistiques des langues indigènes au Brésil. On présente par la suite le projet en cours *Atlas linguistique sonore des langues indigènes du Brésil* (ALSLiB) en soulignant certains de ses impacts, comme l'accès aux bases de données linguistiques des langues indigènes du Brésil qui privilégient une vision claire des variations linguistiques synchroniques et diachroniques, des frontières entre langues et dialectes, des emprunts entre langues génétiquement apparentées ou non, des emprunts entre ces langues d'une part et entre elles et le portugais et l'espagnol de l'autre, sans oublier les emprunts dans le sens inverse.

## Mots-clés

Langues indigènes du Brésil, changement linguistique, langues et dialectes, diversité linguistique.

## Abstract

*This article provides an overview of indigenous languages of Brazil classified according to genetic criteria into linguistic families distributed on a geographical basis. The paper provides an idea of the state of vitality of languages, while emphasizing the importance of dialectological studies, historical, comparative and sociolinguistic indigenous languages in Brazil. The current project Digital Atlas of Brazil's indigenous languages (ALSLiB) is presented here to highlight some of its impacts, such as access to language databases of indigenous languages of Brazil which offer a clear vision of synchronic and diachronic linguistic variations, languages and dialects borders, borrowing between languages genetically related or not, loans between these languages one the one hand and between them and Portuguese and Spanish of the other, not to mention the loans in the reverse side.*

## Keywords

*Indigenous languages of Brazil, linguistic change, languages and dialects, linguistic diversity.*

*«A extinção das línguas amazônicas acarreta, portanto, a perda de um considerável patrimônio humano em termos de experiência cognitiva. Preservar as condições de uso das línguas indígenas a par das condições de exercício das atividades culturais tradicionais dos povos que as falam deve ser a diretriz de uma política sábia de aproveitamento inteligente, não predatório, não destruidor da Amazônia. As culturas e as línguas indígenas, em sua grande diversidade, têm sido vistas como um precioso banco de conhecimentos humanos sobre a interação do homem com o meio ambiente<sup>1</sup>.»*

(RODRIGUES, 1992, p. 4)

- 
1. «L'extinction des langues amazoniennes entraîne donc la perte d'un patrimoine humain considérable en termes d'expérience cognitive. Préserver les conditions d'utilisation des langues autochtones aux côtés des conditions d'activités culturelles traditionnelles des peuples qui les parlent devrait être la ligne directrice d'une politique

## 1. Introduction

Rodrigues (1993a, 1993b)<sup>2</sup> présente la première estimation existante sur la diversité des langues indigènes au Brésil depuis 500 ans. Selon cette estimation, il y aurait environ 1 200 langues différentes parlées par les peuples autochtones dans ce qui correspond au territoire brésilien actuel. Pour arriver à cette estimation, Rodrigues a considéré comme point de départ une liste de 76 peuples autochtones qui se trouvaient dans une zone étroite parallèle à la côte, de la rivière São Francisco au nord, jusqu'au Rio de Janeiro au sud, faite par le père jésuite Ferdinand Cardim au XVI<sup>e</sup> siècle (Cardim, 1978 [manuscrit de 1584]). Selon Rodrigues, Cardim fait une référence explicite à l'identité des 65 langues parlées dans cette extension géographique. Il s'agit de langues différentes les une des autres, mais aussi de la langue des indiens tupinambá de la côte (connus aussi sous le nom de tupiniquins, caetés, potiguaras, tamoios, etc.). En se basant sur le fait que la liste de Cardim a été écrite dans la langue des Tupinambás, Rodrigues remarque que c'était la même langue «[...] que hoje também é chamada de *tupi antigo* e que no século XVII foi denominada língua brasílica<sup>3</sup>» (2005, p. 1), et conclut que les sources d'information étaient les Tupinambas et que cette énumération représente donc leurs connaissances sur leurs voisins les plus immédiats. En outre, l'estimation de Rodrigues est fondée sur d'autres critères tels que la population actuelle, les limites géographiques de certaines régions du monde, comme en Afrique et en Inde, et sur le croisement d'informations des peuples indigènes du Brésil sur différentes périodes et endroits. Ceci corrobore l'estimation de Rodrigues.

D'après Rodrigues (2005), 180 langues indigènes sont parlées au Brésil. Ce chiffre est approximatif en raison de l'imprécision, dans certains cas, de la distinction entre langues et dialectes (des variétés d'une langue si peu différentes, qui ne rendent pas la communication difficile entre leurs locuteurs respectifs). Rodrigues était bien conscient que le chiffre de 180 langues pouvait baisser à cause, par exemple, de la mort de quelques langues considérées en extrême danger d'extinction. Rodrigues, le plus grand spécialiste de linguistique historique des langues indigènes du Brésil de nos jours,

---

sage, d'une utilisation intelligente, qui ne soit ni prédatrice ni destructive de l'Amazonie. Les cultures et langues indigènes dans leur diversité doivent être considérées comme une source précieuse de la connaissance humaine sur l'interaction entre l'homme et l'environnement.»

2. Le texte original a été présenté en 1992 à la réunion annuelle de la Sociedade brasileira para o progresso da ciência (SBPC).
3. «[...] appelée aujourd'hui Tupi ancien et au XVII<sup>e</sup> siècle *langue brasílica*».

a toujours contribué à la classification des langues, dans une perspective génétique, en soulignant que la méthode comparative permettait d'inclure, dans les mêmes groupes de langues, celles pour lesquelles il existe des preuves de descendance d'une même langue ancestrale. À partir des résultats de ses propres études comparatives ainsi que sur d'autres études qu'il croyait scientifiquement bien fondées, Rodrigues (2005) a postulé l'existence de 43 familles et de certaines langues qui, apparemment, ne sont pas apparentées à d'autres langues, conçues ainsi comme des *langues isolées*. Rodrigues explique aussi l'adoption du terme *tronc* linguistique, proposé par Morris Swadesh, pour distinguer des familles dont la diversité interne ne caractérise pas les langues descendantes comme très éloignées en termes temporels, c'est le cas de la famille Romane, par exemple, de celles dont les propriétés communes ne peuvent s'expliquer que par une origine commune plus éloignée, c'est-à-dire, par des relations génétiques caractérisées par une plus grande profondeur temporelle. Rodrigues (1999) croyait que le tronc Tupi en tant qu'organisme bien établi, comprenait dix familles, et était en contraste avec un autre tronc linguistique qu'il considérait comme encore très hypothétique, le tronc Macro-Jê, intégrant 12 familles linguistiques.

La réduction de 1 200 à 180 langues autochtones tout au long des 500 dernières années a été pour Rodrigues l'effet d'un processus de colonisation extrêmement violent et continu, qui existe toujours et qui n'a pas été interrompu même depuis l'indépendance du pays au début du XIX<sup>e</sup> siècle, ou avec l'établissement du régime républicain à la fin de ce siècle, ou même encore avec la promulgation de la *Constitution citoyenne* de 1988.

En dépit de son avis favorable aux perspectives ouvertes par une première Constitution qui reconnaît les droits fondamentaux des peuples autochtones, y compris les droits linguistiques, Rodrigues souligne que la relation entre la société majoritaire et les groupes minoritaires autochtones n'a jamais changée. D'un autre côté, il croyait qu'avec la Constitution de 1988 plusieurs développements importants étaient en cours pour un grand nombre de ces minorités, sur plusieurs plans, y compris l'accès aux projets d'éducation plus spécifiques et en tenant compte les langues maternelles des peuples indigènes du Brésil.

Nous présentons ici les langues indigènes du Brésil classées par Rodrigues selon des critères génétiques propres à la méthode comparative (voir Rodrigues, 1985, 1986, 1999); Rodrigues & Cabral, 2012), mais avec des adaptations faites par ses disciples<sup>4</sup>, considérant aussi les résultats des en-

---

4. Ana Suelly Arruda Câmara Cabral (voir Cabral & Solano, 2006; Rodrigues & Cabral, 2012; Cabral *et al.*, 2011, 2014; entre autres), Beatriz Carretta Corrêa da

quêtes sociolinguistiques et des nouvelles études comparatives faites dernièrement par des chercheurs du Laboratoire des langues et littératures indigènes de l'université de Brasília (LALLI/UnB).

### **Langues indigènes du Brésil classées d'après des critères génétiques**

Tronc Tupí = 10 familles

- Tupí-Guaraní (Bolivie, Brésil, Colombie, Guyane française, Pérou, Venezuela)
- Awetí (Tupí, Brésil)
- Mawé (Tupí, Brésil)
- Mundurukú (Tupí, Brésil)
- Jurúna (Tupí Brésil)
- Arikém (Tupí, Brésil)
- Tuparí (Tupí, Brésil)
- Mondé (Tupí, Brésil)
- Puroburá (Tupí, Brésil)
- Ramaráma (Tupí, Brésil)

Tronc Macro-Jê = 12 familles

- Jê (Brésil)
- Kamakã † (Brésil)
- Maxakalí (Brésil)
- Krenak (Brésil)
- Purí † (Brésil)
- Karirí † (Brésil)
- Yatê (Brésil)
- Karajá (Brésil)
- Ofayé (Brésil)
- Boróro (Brésil)
- Guató (Brésil)
- Rikbaktsá (Brésil)

---

Silva (2011), Tabita Fernandes da Silva (2010; Silva & Cabral, 2006), Anderbio Marcio Silva Martins (2007, 2011), Sanderson Castro Soares de Oliveira (2014), Jorge Domingues Lopes (2014), Lidiane Szerwinsk Camargos (2013), Ariel Pheula do Couto e Silva (2014), Eliete de Jesus Bararuá Solano (2004, 2009) et Wary Kamaiurá (2012).

## 2 familles étendues

- Aruák ou Arawák (Bolivie, Brésil, Guyane, Guyane française, Pérou, Suriname)
- Karíb (Brésil, Colombie, Guyane, Guyane française, Suriname, Venezuela)

## 3 familles de grandeur moyenne

- Páno-Takana (Bolivie, Brésil, Pérou)
- Tukáno (Brésil, Colombie, Équateur, Pérou)
- Makú (-Puináve) (Brésil, Colombie, Venezuela)

## 10 petites familles

- Arawá (Brésil et Pérou)
- Katukína (Brésil)
- Múra (Brésil)
- Nambiquára (Brésil)
- Txapakúra (Bolivie, Brésil)
- Witóto (Brésil, Colombie, Pérou)
- Yanomámi (Brésil, Venezuela)
- Txapakúra (Brésil)
- Irantxe (Brésil) (deux langues)
- Jabotí (Brésil) (deux langues)

## 6 langues isolées

- Tikúna (Brésil, Pérou et Colombie)
- Trumai (Brésil)
- Kwazá (Brésil)
- Aikanã (Brésil)
- Kanoé (Brésil)
- Chiquitano (Brésil, Bolivie)

En dépit de cette riche diversité linguistique, la grande majorité des langues est considérée comme sérieusement en danger, soit parce que leur utilisation est en déclin progressif soit parce qu'elles ne sont plus transmises aux nouvelles générations. Cela advient depuis plusieurs années de politiques d'assimilation promues par le gouvernement fédéral brésilien, dont le but a été l'intégration des communautés autochtones dites minoritaires dans la société brésilienne dominante, tout en décourageant et en supprimant des milliers d'années de diversité linguistique et de connaissances millénaires.

Le degré de différenciation entre les langues appartenant aux deux troncs linguistiques ci-dessus représente une différenciation génétique considérable des langues à travers le temps, soit de nature lexicale et morphosyntaxique, comme c'est le cas des différences entre les familles Arikém et Juruna (tronc Tupi), soit entre eux et les langues de la famille Tupi-Guarani, la famille la plus nombreuse du même tronc. Malheureusement, la majorité de ces familles ainsi que les langues isolées, à l'exception du Tikuna, sont menacées d'extinction, malgré leur importance tant pour les communautés parlantes de ces langues, que pour la connaissance de la diversification humaine en Amérique du Sud.

L'état de santé général des langues autochtones du Brésil est difficile. Si l'on prend comme exemple les 13 langues de la famille Tupi-Guarani parlées dans l'État du Pará — zo'é, apiaká, kayabí, mbyá, asuriní du Xingu, araweté, anambé, asuriní do Tocantins, suruí aikewára, tembé et guajajára —, à l'exception du zo'é, qui totalise 260 individus, utilisant tous la langue zo'é comme langue de communication; du guajajára, qui totalise à peu près 8 000 locuteurs (avec un degré de connaissance allant du locuteur natif, en passant par celui qui pratique la langue native dans des situations très limitées, jusqu'à celui qui comprend, mais qui ne parle pas la langue native); et de la langue araweté, toujours parlée comme langue première, malgré les effets du contact essentiellement récent, à la suite de la construction de la centrale hydroélectrique Belo Monte; les langues restantes sont fortement menacées d'extinction pour les prochaines générations si elles ne continuent pas à être transmises aux enfants, comme c'est le cas observé chez la majorité des locuteurs des langues asuriní du Xingu, asuriní du Tocantins, suruí aikewára et kayabí de certains villages. Les autres langues, l'apiaká et l'anambé, ne sont plus transmises aux nouvelles générations et n'ont que très peu de gardiens d'une connaissance mémorisée de la langue native.

La famille Puruborá, constituée actuellement d'une unique langue éponyme, n'existe que partiellement dans les mémoires d'une douzaine de personnes au maximum. Parmi les langues de la famille Juruna, constituée des langues juruna et xipaya, seul le juruna est entièrement parlé par tous les membres de la communauté soit une population d'environ 260 individus, tandis que la langue xipaya n'est que partiellement retenue dans la mémoire d'une seule femme. En retournant à d'autres langues de la famille Tupi-Guarani, les langues dites guaraní — le mbya, le nhandeva et le kaiowá — sont pleinement parlées par une population stable, ainsi que le tapirapé, le kamaiurá, le wayampí, l'amondáwa, le tembé, le uru-eu-au-au, et le tenharim. Certaines ont moins d'une dizaine de locuteurs, c'est le cas du karipuna, du piripkura (trois locuteurs), du juma (cinq locuteurs), du

xeta (deux locuteurs), du jiahuju (trois locuteurs) et d'awre awra (un seul locuteur)<sup>5</sup>.

Une lueur d'espoir dans l'avenir de ces langues est l'éveil constaté chez les peuples autochtones pour la reconnaissance de l'importance des langues maternelles. Les enseignants, en particulier, ont joint des initiatives du gouvernement dans la formation d'enseignants indigènes du secondaire, de licence en études interculturelles indigènes et aussi dans des programmes d'études supérieures en linguistique, au niveau du master et du doctorat, dont les efforts se concentrent sur la documentation et sur l'étude scientifique des langues natives, en envisageant une éducation différenciée et spécifique, tout en respectant les différentes réalités des peuples autochtones du Brésil.

## 2. Les études scientifiques des langues indigènes du Brésil

L'étude scientifique des langues indigènes du Brésil, bien qu'avancée pour certaines d'entre elles, est encore loin de servir aux locuteurs de ces langues. Plusieurs thèses de doctorat et de mémoires de maîtrise, ainsi que de nombreux articles sur ces langues sont écrits soit en anglais, soit en français, ce qui rend leur accès impossible aux communautés indigènes qui ne parlent que le portugais ou l'espagnol, en plus de leurs langues maternelles. De la même manière, les bases de données de grands projets de documentation n'ont eu aucune utilité pour les peuples autochtones, ne leur étant pas disponible.

Néanmoins, la diversité linguistique que représente l'ensemble des langues indigènes brésiliennes s'ouvre vers une multitude de nouvelles études qui peuvent en fait contribuer à la connaissance linguistique de ces langues, pour le développement des sciences du langage, ainsi que pour des programmes éducatifs (parmi lesquels des programmes d'éducation formelle) qui envisagent le renforcement culturel et linguistique des ces peuples.

---

5. Le dernier recensement effectué au Brésil (2010) a constaté que les personnes qui s'identifient comme autochtones ont mentionnés les noms de 270 langues. Néanmoins, seules 200 de ces langues, comme nous l'avons déjà dit, sont des langues qui demeurent, soit parlées soit gardées dans la mémoire des derniers individus qui ont eu de la chance d'être élèves en écoutant leurs parents communiquer dans la langue de leurs ancêtres. Parmi les 70 langues restantes du recensement, il y a des langues natives des immigrants qui vivent actuellement au Brésil, des langues déjà éteintes et connues d'après des listes de mots ou d'après des documents historiques. Mais il y a aussi des noms de langues qui correspondent en réalité à des noms de peuples dont la langue originelle a disparu sans laisser de traces et sur lesquelles il n'y a pas d'information.

Parmi ces nouvelles études, on notera les études linguistiques appliquées, comme l'enseignement des langues maternelles, et les études sur la dialectologie et la géographie linguistique, qui sont à leur tour étroitement liées à l'histoire des langues.

Dans ce contexte, le projet *Atlas linguistique sonore des langues indigènes du Brésil* (ALSLiB) présenté au Conseil national de développement scientifique et technologique (CNPq), en 2013, représente une contribution fondamentale pour l'étude linguistique de nos langues maternelles. Ce projet met en évidence les langues indigènes du Brésil distribuées selon des critères géographiques et génétiques et les traits partagés par ses variétés ou d'autres langues génétiquement proches, et ceux acquis par le contact avec d'autres langues apparentées ou non.

Le souci rencontré par l'affaiblissement rapide de l'usage des langues qui sont encore parlées au Brésil (makurap, jabuti, kanoé, akuntsu, parintintin, jiahó, juma, asurini du Xingu) et beaucoup d'autres langues fortement menacées, et par le progrès des études scientifiques portant sur plusieurs d'entre elles souligne la nécessité de la construction d'un atlas linguistique des langues indigènes du Brésil, sur la base de la production réussie de l'*Atlas linguistique du Brésil* (projet *Atlas linguistique du Brésil* – ALiB), à partir de la publication d'un ensemble d'atlas régionaux, afin d'atteindre l'objectif final sous la forme d'un atlas national.

Bien sûr, comme chaque langue réunit un ensemble riche de faits linguistiques qui reflètent les différentes expériences quotidiennes et historiques de leurs communautés dans l'environnement où elles ont vécues, aucun linguiste ne peut, à lui tout seul, épuiser les possibilités d'études d'une langue naturelle, quelles soient de nature descriptive et/ou théorique sur la grammaire, le lexique, les situations de contact, l'histoire individuelle d'une langue ou d'un groupe de langues, ou encore des études de nature dialectologique et sociolinguistique.

L'ALSLiB a été conçu pour être un instrument clé d'information linguistique, sociolinguistique et culturelle des langues indigènes brésiliennes, vues de façon indépendante et aussi à travers les connexions génétiques qui les unissent à d'autres langues (familles, sous-groupes des familles, etc.).

Le projet ALSLiB s'inscrit dans le cadre d'un partenariat entre chercheurs représentant deux institutions d'enseignement supérieur, l'université de Brasilia (UnB) et l'université fédérale du Pará (UFPA). Cette coopération s'étend progressivement et intègre d'autres chercheurs de ces deux institutions, d'autres universités et centres de recherche, ainsi que des associations indigènes. L'objectif à moyen et long terme est de produire un atlas interactif qui devrait refléter la diversité linguistique des langues indigènes du Brésil. L'atlas permettra d'illustrer les différentes langues et

leurs respectives variétés, en soulignant les principales zones linguistiques, avec leurs traits linguistiques et le chemin pris par certains mots, lesquels reflètent la migration et les contacts entre les locuteurs de langues apparentées et entre locuteurs de langues génétiquement distinctes. L'atlas sera également une source de diffusion de connaissances sur la préhistoire des langues et des peuples indigènes du Brésil, dans la mesure où il montrera des mots apparentés à travers les langues, des changements sonores, sémantiques et grammaticaux qu'elles ont subis, entre autres.

Le projet *Atlas linguistique sonore des langues indigènes du Brésil* (ALSLiB) devra englober, d'une part, une vision diachronique — la diversification ou la dialectisation des langues ancestrales dans des familles et/ou troncs linguistiques — et, d'autre part, les variations synchroniques des langues autochtones et de leurs frontières dialectales respectives. Ainsi conçu, il devra présenter la dialectologie à la lumière de la recherche linguistique et sociolinguistique associée à la géographie, de sorte à enrichir les connaissances linguistiques sur les langues en plaçant l'attention sur certains aspects de leurs utilisations, et sur les facteurs de conditionnement socioculturels, tels que la classe d'âge, le sexe, les relations de parenté, les fonctions sociales et d'autres paramètres qui apparaîtront dans le cadre de la recherche sociolinguistique. L'ALSLiB intègre les dimensions diatopique, diagénérique, diagénérationnelle, diastratique, diaphasique et diaréférentielle aux côtés des paramètres génétiques. Ces derniers sont cruciaux pour la compréhension de la diversité linguistique du Brésil et de sa préhistoire linguistique, mais il est également important pour l'avancement des connaissances sur les modèles génétiques de la diversification, ainsi que sur les causes internes et externes qui ont induit des changements des langues à travers le temps.

L'ALSLiB mettra en évidence le Brésil, comme le pays d'Amérique latine qui possède le plus grand nombre de langues et la plus grande diversité linguistique, avec 43 familles linguistiques, la plupart desquelles ont une préhistoire qui peut être reconstruite (voir Rodrigues, 2001, 2007; Rodrigues & Cabral, 2012).

### 3. Un atlas numérique pour les langues autochtones du Brésil

Le souci de stockage de données linguistiques des langues autochtones du Brésil a grandi dès les années 1980. Actuellement, plusieurs projets et programmes nationaux et internationaux cherchent à documenter le nombre de langues en voie de disparition, soit au Brésil soit dans d'autres parties du monde. Selon Razky (2003), la dialectologie moderne et la sociolinguistique ont toujours eu un intérêt pour le « stockage des données, la créa-

tion de bases textuelles diatopiques, diaphasiques et diastratiques» et aussi pour le « développement de cartes linguistiques phonétiques, morpho-syntaxiques, les données écrites ou sonores, lexico-sémantiques et pragmatiques recoltées sur le terrain ». Razky (2003) note également que le développement de l'industrie informatique a conduit certains précurseurs américains et européens dans le domaine de la dialectologie à utiliser les données sonores dans la représentation cartographique. Au Brésil, la première manifestation de cette possibilité a eu lieu dans l'État du Pará avec la première phase de l'AliPA – *Atlas Geossociolingüístico do Pará* (voir Razky, 1998, 2003, 2005).

L'importance accrue d'un atlas numérique (ou parlant), c'est qu'il permet un contact direct avec l'information contenue dans les cartes. Au lieu d'accéder à des données de transcription lexicales ou phonétiques, l'utilisateur d'un programme informatique a également la possibilité d'écouter les données numérisées (textes entiers dans des genres différents, des données linguistiques précédemment sélectionnées et classées, images, vidéos...). Comme le souligne Razky (2005), l'atlas sonore rend la représentation cartographique plus proche de la réalité linguistique décrite, une fois que la subjectivité inhérente à toute transcription phonétique est surmontée par la présentation de données réelles qui peuvent être entendues plus d'une fois et transcrites plusieurs fois s'il était nécessaire.

#### 4. Contribution de l'innovation technologique

L'*Atlas sonore des langues indigènes du Brésil*, tel que nous l'avons proposé, sera une source importante de renseignements sur la vie des langues indigènes du Brésil, étant une base de données importante qui favorise l'analyse de variables sociales qui n'ont pas été entièrement traitées par ailleurs, et qui mettra en évidence son importance en associant ses outils à d'autres initiatives et d'autres outils de stockage et de diffusion linguistique.

Le projet ALSLiB prévoit le développement de ses propres paquets de logiciels pour l'entrée des données spécifiques, ainsi que le développement de mécanismes pour accéder à ces données, ce qui implique nécessairement une nouvelle contribution dans le domaine de l'ingénierie linguistique.

#### 5. Références de base pour le projet ALSLiB

En proposant la formulation d'un projet de l'*Atlas des langues pour les langues autochtones du Brésil*, nous prenons comme référence principale le projet *Atlas linguistique du Brésil* (Cardoso, 1996; Comitê nacional do

projeto ALiB, 2003), et des atlas régionaux du Brésil, tel que l'*Atlas linguistique de Sergipe* (Rossi & Ferreira, 1994 ; Ferreira *et al.*, 1987), l'*Atlas linguistique de Paraíba* (Aragão & Menezes, 1984), l'*Atlas linguistique du Rio de Janeiro* (Votre, 1978), l'*Atlas linguistique du Pará* (Razky, 1998, 2002, 2003, 2005 ; Razky & Carvalho, 2002), l'*Atlas linguistique du Rio Grande do Norte* (Aragão & Pereira, 2005) et l'*Atlas linguistique de Paraná* (Aguilera, 1994). On n'oublie pas ici, bien entendu, l'*Atlas linguistique de la France* (Gilliéron, 1902-1910) en raison de son importance dans l'histoire des atlas géolinguistiques des langues du monde, et l'*Atlas linguistique guarani-roman* (Thun *et al.*, 2009), car il a été la première initiative d'un atlas linguistique impliquant des régions linguistiques autochtones adjacentes au Brésil et génétiquement liées à des langues autochtones de cette région. Nous avons utilisé aussi les études importantes sur les atlas sonores développées par Razky (2003, 2005), afin d'ouvrir de nouvelles voies pour la transparence, l'utilité et l'appréciation des données linguistiques d'un atlas linguistique.

## 6. Objectifs à atteindre

Les objectifs spécifiques du projet *Atlas linguistique sonore des langues indigènes du Brésil* comportent, dans une première phase, les étapes suivantes :

- enregistrer et décrire de nouvelles données, tout en mettant en évidence des phénomènes linguistiques qui montrent les différences, mais aussi des similitudes entre les langues autochtones d'un même groupe linguistique et celles entre ces variantes ;
- identifier et cartographier les variations phonétiques de chaque langue, en appliquant un questionnaire phonétique et phonologique à des groupes de locuteurs définis précédemment ;
- organiser la bibliothèque sonore en utilisant des programmes de documentation linguistique comme ELAN ;
- utiliser des outils acoustiques pour le nettoyage, la segmentation et la base de données de catégorisation ;
- identifier et cartographier les variations lexicales, grammaticales et sémantiques de chaque langue, en appliquant des questionnaires spécifiques à des groupes de locuteurs définis précédemment ;
- recueillir des textes sonores représentatifs des langues en usage, illustrant les variations en fonction de critères sociolinguistiques ;

- identifier des isoglosses, en tenant compte de la délimitation des frontières linguistiques entre les langues du même groupe génétique ou entre des variétés de la même langue ;
- fournir un ensemble de données qui sont utilisées par des études lexicographiques et grammaticales afin de développer, dans le premier cas, des exemples de synonymie, d'antonymie, de polysémie, entre autres, et dans le second cas, les usages linguistiques qui, généralement, ne sont pas incorporés dans l'étude des grammaires, et qui sont également utiles pour le développement des manuels de soutien pédagogique des langues autochtones dans les écoles de villages. Les données devraient également servir aux ethnologues et universitaires d'autres domaines connexes à la linguistique.

## 7. La méthodologie

L'atlas identifie la variation dans l'espace et dans le temps en utilisant la méthodologie de la géolinguistique, tout en observant les individus en interaction et les *collections d'individus*. La covariation sociolinguistique est une des nos priorités et l'intégration de la dimension sociale que nous avons privilégiée assure la valeur des informations provenant d'un champ bien marqué et permet de comprendre les mécanismes internes impliqués dans la variation et dans les changements des langues.

La recherche envisage l'étude des variations phonétiques, lexicales, sémantiques et grammaticales en fonction de leurs motivations pragmatiques respectives. Initialement elle sera sociolinguistique, selon les paramètres de recherche de la géosociolinguistique. Les informateurs sont d'abord stratifiés selon des critères diagénériques et diagénérationnels. La deuxième étape concerne l'utilisation de questionnaires qui contemplent les variations linguistiques.

Le projet ALSLiB compte sur la participation effective des chercheurs autochtones, des étudiants diplômés en linguistique ou étudiants en interculturel qui se consacrent à l'étude de leur langue maternelle. Nous mettons l'accent, durant toutes les phases du projet, sur l'importance de la participation active des Indiens. C'est une occasion importante de contribuer également à la formation de chercheurs autochtones dans le domaine de la dialectologie et de la géographie linguistique.

## 8. Sélection des points de l'enquête

Le choix des langues a été fait en fonction de l'intérêt des communautés autochtones déjà contactées et connaissant les activités des chercheurs du projet. Les critères de sélection des points de l'enquête dépendent de critères tels que l'existence d'un seul village, de plusieurs villages ou de plus d'un territoire indigène.

Les dix premières langues choisies sont les langues asuriní du asurini, tembe, guajajara, kamaiurá, zo'é, ka'apor, surui-aikewára (fam. Tupi-Guarani, tronc Tupi), tupari (fam. Tuparí, tronc Tupi), suruí paiter (fam. Mondé, tronc Tupi), awetí (fam. Awetí, tronc Tupi), xikrin, kraho, xavante, xerente, laklanō (fam. Jê, tronc Macro-Jê), kalapalo (fam. Karib), mehinako (fam. Arawak) et kaxinawá (fam. Pano). La recherche compte sur la participation de chercheurs indigènes, non pas comme de simples informateurs, mais comme de véritables agents dans le processus de la recherche scientifique.

## 9. Techniques de recueil de données et stratification sociale des informateurs

Pour la préparation de l'atlas sonore, nous adoptons, pour les données en portugais, une adaptation du questionnaire phonético-phonologique préparé par le comité du projet *Atlas linguistique du Brésil* (ALiB), composé de 157 questions afin de vérifier les changements phonétiques dans l'espace social et dans le temps.

Pour répondre au questionnaire phonétique nous choisissons, pour chaque localité, quatre locuteurs. Les critères sont les mêmes que pour les données en portugais. En plus de ces questionnaires, nous appliquerons pour chaque langue le questionnaire de Brent et Kaufman, qui a été développé au sein du *Projet de documentation des langues autochtones de l'Amérique du Sud* (SAILDP), révisé à partir de Berlin *et al.* (1986).

Pour les données lexicales, nous utiliserons un questionnaire regroupant des champs sémantiques comme les parties du corps humain, les termes de parenté, les animaux et les plantes. Le choix des informateurs repose sur la base des critères suivants :

- **sexe** : deux hommes et deux femmes ;
- **âge** : deux locuteurs ayant entre 18 et 30 ans et deux autres entre 40 et 70 ans ;
- **scolarité** : non scolarisés et, pour les scolarisés, on définit leurs niveaux et le lieu de leur scolarisation.

## 10. Informatisation des données

Un logiciel interactif sera développé pour permettre la consultation des données sonores qui illustrent les variations phonétiques de chaque langue. La base de données informatisée fournira à l'utilisateur l'accès à la voix des locuteurs sans passer par la subjectivité du linguiste, si commune dans les transcriptions phonétiques. La base de données sonore sera accessible sur internet afin de permettre la socialisation des connaissances sur les peuples autochtones du Brésil.

Le logiciel devra également permettre le traitement des données du point de vue acoustique. Le logiciel PRAAT (version 5.3.51) de Bovesma et Wenink (2013), qui est du domaine public, intégrera l'interface interactive. Le but ici est de permettre une recherche acoustique de telle sorte que dans un proche avenir on puisse penser à la cartographie acoustique des segments phonétiques soumis à la variation des langues indigènes du Brésil. Le programme informatique créé devrait également intégrer et fournir des données lexicales, grammaticales et sémantiques ainsi que des données historiques, culturelles, archéologiques, entre autres sur les langues autochtones du Brésil et sur leurs locuteurs.

## 11. Quelques conclusions

L'*Atlas sonore des langues indigènes du Brésil* sera une importante source de connaissance linguistique sur les langues indigènes du Brésil. À cause de l'inclusion des études lexicales, il y aura aussi une importante interface avec la botanique, la zoologie, la géographie (toponymie), et avec d'autres disciplines. L'incorporation des informations historiques, rapportés aussi par des locuteurs des langues étudiées, sera aussi un facteur fondamental pour le dialogue avec l'histoire. L'atlas sera donc d'une importance fondamentale pour stimuler la discussion sur les différences entre langues et dialectes, mais aussi sur le changement des langues et sur les situations de contact vécues par les locuteurs des langues documentées. L'ALSLiB sera une innovation positive pour la divulgation et l'enseignement des langues indigènes à un public plus large que ceux des linguistes. Pour utiliser des technologies informationnelles, il sera aussi un produit technologique au service de stockage des données et de l'éducation ouverte à tous.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- AGUILERA Vanderci de Andrade, 1994, *Atlas lingüístico do Paraná (ALPR)*, Curitiba, Imprensa Oficial do Estado.
- ARAGÃO Maria do Socorro Silva de & PEREIRA Maria das Neves, 2005, «Atlas Lingüístico do Rio Grande do Norte – um projeto em desenvolvimento», dans V. d. A. Aguilera, *A Geolingüística no Brasil: trilhas seguidas, caminhos a percorrer*, Londrina, EDUEL.
- ARAGÃO Maria do Socorro da Silva & MENEZES Cleusa Palmeira Bezerra de, 1984, *Atlas lingüístico da Paraíba*, Brasília, UFPB/CNPq, 2 vol.
- BERLIN Brent, KAUFMAN Terrence, CARSON Neusa & RODRIGUES Aryon, 1986, «Diagnostic vocabulary», dans *Projeto de documentação das línguas indígenas da América do Sul [South American indian languages documentation project (SAILDP)]*, Berkeley, University of California, Campinas, SP, Universidade Estadual de Campinas.
- BORESMA Paul & WEENINK David, 2013, *Praat: Doing Phonetics by Computer* [logiciel, version 5.3.51]. Disponible sur <[www.praat.org](http://www.praat.org)> [dernier accès en avril 2015].
- CABRAL Ana Suelly Arruda Câmara & SOLANO Eliete de Jesus Bararuá, 2006, «Mais fundamentos para a hipótese de proximidade genética do araweté com línguas do subramo V da família Tupí-Guaraní», *Estudos da Língua(gem)*, vol. 4, Vitória da Conquista/BA, p. 41-65.
- CABRAL Ana Suelly Arruda Câmara *et al.*, 2011, «Línguas entrelaçadas, uma situação sui generis de línguas em contato», *Papia*, vol. 21, n° 2, p. 221-230.
- CABRAL Ana Suelly Arruda Câmara *et al.*, 2014, «A linguística histórica das línguas indígenas do Brasil, por Aryon Dall'igna Rodrigues: perspectivas, modelos teóricos e achados», *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, vol. 30, número spécial, p. 513-542. Disponible sur <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-44502014000300513&lng=pt&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502014000300513&lng=pt&nrm=iso)> [dernier accès en avril 2015].
- CAMARGOS Lidianie Szerwinsk, 2013, *Consolidando uma proposta de Família Linguística Boróro*, thèse de doctorat, PPGL/LIP/IL/UnB, Brasília.
- CARDIM Fernão, 1978, *Tratados da terra e gente do Brasil*, São Paulo, Cia. Editora Nacional/MEC, 3<sup>e</sup> éd.
- CARDOSO Suzana, 1996, «O(s) atlas lingüístico(s) do Brasil: atualidade e importância», dans *Anais do Encontro de estudos do bilingüismo e variação lingüística na região sul*, Curitiba, Universidade Federal do Paraná.
- COMITÊ NACIONAL DO PROJETO ALiB, 1998-2003, *Projeto atlas lingüístico do Brasil*, Salvador, Instituto de Letras, manuscrit.
- FERREIRA Carlota *et al.*, 1994, *Diversidade do Português do Brasil: estudos de dialetologia rural e outros*, Salvador, UFBA, 2<sup>e</sup> éd.
- FERREIRA Carlota, MOTA Jacyra, FREITAS Judith, ANDRADE Nadja, CARDOSO Suzana, ROLLEMBERG Vera & ROSSI Nelson, 1987, *Atlas Lingüístico de Sergipe (ALS)*, Salvador, UFBA-FUNDESC.

- GILLIERON Jules & EDMONT Edmont, 1902-1910, *Atlas linguistique de la France*, Paris, Champion.
- KAMAIURÁ Wary, 2012, *Aweti e Tupi-Guarani: relações genéticas e contato lingüístico*, mémoire de master, PPGL/IL/LIP/UnB, Brasília.
- LOPES Jorge Domingues, 2014, *Uma interface da documentação lingüística e modelos lexicográficos para línguas indígenas brasileiras: uma proposta para o suruí-aikewára*, thèse de doctorat, PPGL/IL/LIP/UnB, Brasília.
- MARTINS Andrébio Márcio Silva, 2011, *Uma Avaliação da Hipótese de Relações Genéticas entre o Guató e o Tronco Macro-Jê*, thèse de doctorat, PPGL/IL/LIP/UnB, Brasília.
- MARTINS Andrébio Márcio Silva, 2007, *Revisão da família lingüística Kamakã proposta por Chestmir Loukotka*, mémoire de master, PPGL/IL/LIP/UnB, Brasília.
- OLIVEIRA Sanderson Soares de, 2014, *Contribuições para a reconstrução do Proto-Páno*, thèse de doctorat, PPGL/IL/LIP/UnB, Brasília.
- RAZKY Abdelhak, 1998, «O atlas geo-sociolingüístico do Pará: abordagem metodológica», dans V. d. A. Aguilera (éd.), *A Geolingüística no Brasil: caminhos e perspectivas*, Londrina, EDUEL, p. 155-164.
- RAZKY Abdelhak, 2003, «A construção de atlas sonoros: procedimentos metodológicos para o ALISPA», dans A. Razky (éd.), *Estudos geo-sociolingüísticos no estado do Pará*, Belém, Grafia, p. 173-183.
- RAZKY Abdelhak, 2003-2004, *Atlas lingüístico digital do Pará*, projet post-doctoral, manuscrit.
- RAZKY Abdelhak, 2005, «Atlas Lingüístico Sonoro do Pará: uma nova perspectiva para a organização de corpus geolingüístico», dans V. d. A. Aguilera (éd.), *A Geolingüística no Brasil: trilhas seguidas, caminhos a percorrer*, Londrina, EDUEL.
- RAZKY Abdelhak & CARVALHO Rosana, 2002, «Variation et changement linguistique : le cas du /s/ post-vocalique dans le nord du Brésil», *Revue Analyses*, n° 9, Presses universitaires de Toulouse.
- RODRIGUES Aryon Dall'Igna, 1992, «Diversidade Linguística na Amazônia», dans *Anais do Simposio internacional sobre meio ambiente, pobreza e desenvolvimento da amazonia*, Belém, PRODEPA, p. 367-370.
- RODRIGUES Aryon Dall'Igna, 2005, «Sobre as línguas indígenas e sua pesquisa no Brasil», *Ciência e Cultura* (SBPC), vol. 57, n° 2, p. 35-38. Disponible sur <<http://cienciaecultura.bvs.br/>> [dernier accès en avril 2015].
- RODRIGUES Aryon Dall'Igna, 1985, «The Present State of the Study of Brazilian Indian Languages», dans H. M. Klein et L. R. Stark (éds), *South American Indian Languages: Retrospect and Prospect*, Austin, University of Texas Press, p. 405-439.
- RODRIGUES Aryon Dall'Igna, 1986, *Línguas Brasileiras: para o conhecimento das línguas indígenas*, São Paulo, Edições Loyola.
- RODRIGUES Aryon Dall'Igna, 1993a, «Línguas indígenas : 500 anos de descobertas e perdas», *DELTA*, vol. 9, n° 1, p. 83-103.

- RODRIGUES Aryon Dall'Igna, 1993b, «Línguas indígenas : 500 anos de descobertas e perdas», *Ciência e Cultura*, vol. 95, p. 20-26.
- RODRIGUES Aryon Dall'Igna, 1999, «Macro-Jê», dans R. M. W. Dixon et A. Y. Aikhenvald (éds), *The Amazonian Languages*, Cambridge University Press, p. 164-206.
- RODRIGUES Aryon Dall'Igna, 2001, «Biodiversidade e diversidade etnolingüística na Amazônia», dans M. S. Simões (éd.), *Cultura e biodiversidade: entre o rio e a floresta*, Belém, UFPA, p. 269-278.
- RODRIGUES Aryon Dall'Igna, 2007, «A ameaça à diversidade linguística no Brasil», conférence réalisée au cours des Seminários do LALI, Universidade de Brasília.
- RODRIGUES Aryon Dall'Igna & CABRAL Ana Suelly Arruda Câmara, 2012, «Tupían», dans L. Campbell et V. Grondona (éds), *The Indigenous Languages of South America*, Berlin/Boston, Mouton de Gruyter, vol. 2, p. 495-574.
- SILVA Ariel Pheula do Couto e, 2014, *Elementos fonológicos, morfossintáticos e sintáticos da língua Avá-Canoeiro do Tocantins*, mémoire de master, PPGL/IL/LIP/UnB, Brasília.
- SILVA Beatriz Carreta da, 2011, *Mawé/Awetí/Tupí-Guaraní: Relações Lingüísticas e Implicações Históricas*, thèse de doctorat, PPGL/IL/LIP/UnB, Brasília.
- SILVA Tabita Fernandes da & CABRAL Ana Suelly Arruda Câmara, 2006, «Desvendando a História Interna do Morfema Pə da Língua Tenetehára», *Estudos da Língua(gem)*, vol. 4, p. 77-86.
- SILVA Tabita Fernandes da, 2010, *História da Língua Tenetehára: contribuição aos estudos histórico-comparativos sobre a diversificação da família lingüística Tupi-Guarani do Tronco Tupí*, Brasília, Universidade de Brasília.
- SOLANO Eliete de Jesus Bararuá, 2004, *A posição do Araweté na família Tupí-Guaraní: considerações lingüísticas e históricas*, mémoire de master, PPGL/IL/LIP/UnB, Brasília.
- SOLANO Eliete de Jesus Bararuá, 2009, *Descrição gramatical da língua araweté*, thèse de doctorat, PPGL/IL/LIP/UnB, Brasília.
- THUN Harald, AQUINO Almidio, DIETRICH Wolf & SYMEONIDIS Haralambos, 2009, *Atlas lingüístico Guaraní-Románico, Tomo I: Léxico del cuerpo humano*, Kiel, Westensee.
- VOTRE Sebastião J., 1978, *Aspectos da Variação Fonológica no Rio de Janeiro*, thèse de doctorat, PUC, Departamento de Letras, Rio de Janeiro.