



## Francisco da Silva Borba

Le 28 Juin 2013, j'ai rencontré Francisco da Silva Borba, professeur retraité de la Faculté de Sciences et Lettres de l'Université de l'Etat de São Paulo - UNESP, campus d'Araraquara, et un des plus grands linguistes et lexicographes brésiliens. Pendant une heure et demie nous avons parlé de son travail en lexicographie, de ces publications dans ce domaine et de ses projets. Nous pourrions en citer d'autres, comme son *Introdução aos Estudos Linguísticos* [Introduction aux études linguistiques], publiée chez Pontes Editores à Campinas, et dont la première édition date de 1971, un vrai manuel de linguistique pour les jeunes étudiants de Lettres ; ainsi que de ses importantes activités académiques, comme sa participation à la première équipe qui a dirigé le Groupe d'Etudes Linguistiques de l'Etat de São Paulo - GEL et d'avoir également organisé le premier Séminaire de ce groupe, à présent dans sa 61<sup>e</sup> édition, ou d'avoir été, en tant que directeur, le responsable de la construction du nouveau campus de l'UNESP d'Araraquara ; encore en dehors de l'académie, il participa à la création d'un centre d'éducation pour enfants autistes dans cette même ville, où il continue encore à jouer un rôle important. Lors de l'entrevue, il nous a beaucoup parlé de ses projets passés et à venir, cependant, par manque d'espace, nous avons dû en couper des parties, sous réserve de porter préjudice au sujet de la rencontre : la lexicographie.

**Interview organisée par Adriana Zavaglia et Philippe Humblé ; réalisée, transcrite et traduite par Adriana Zavaglia, avec révision de la traduction par Lélia Magalie Angole.**

**Synergies Brésil (SB) :** Comment votre parcours vous a-t-il mené à la lexicographie?

**Francisco da Silva Borba (FSB) :** D'abord cela a été une affaire de goût personnel et d'attirance. Dans les études linguistiques, je me suis toujours intéressé à la syntaxe, d'où mon rattachement premier à la syntaxe structurale de Tesnière. Je me suis spécialisé dans le structuralisme et j'ai eu une formation en France, à la Sorbonne - à l'époque, suivre une formation là-bas était quelque chose d'exceptionnel -, j'ai étudié la linguistique avec Martinet, qui était à l'époque le grand chef du structuralisme français. Je pense aussi au structuralisme européen, provenant de Prague. C'est dans ce contexte que je suis devenu professeur agrégé. J'ai préparé ma thèse de doctorat en syntaxe, en France, et que j'ai ensuite soutenue au Brésil, où j'étais déjà professeur.

En tant que professeur de linguistique, je me suis toujours intéressé à tous les sujets de la linguistique, mais je voulais trouver une méthodologie d'analyse des langues, car je pense que c'est plus intéressant pour les étudiants en licence à leur début de formation. Au niveau de la licence, on doit pouvoir leur offrir les bases du travail en linguistique, afin qu'ils puissent décider de leurs choix d'avenir. C'est ainsi que j'ai débuté. Puis, en syntaxe et en linguistique d'une manière générale, j'ai commencé à réfléchir sur la circulation du lexique, me focalisant sur la langue écrite, puisque c'est ce que l'on enseigne à l'école. J'ai commencé à enseigner très tôt, vers 16, 17 ans, lorsque je donnais des cours de portugais. J'ai toujours porté un vif intérêt à la langue écrite. En effet, l'usage attire l'attention, non pas des mots en soi, mais sur les diverses combinaisons possibles de mots, de jeux de mots, et de la syntaxe. Cette dernière est un réseau d'interrelations remplies par le lexique, cela est très important à observer. D'abord il y a le jeu de la langue, une opération interne, puis, avec le lexique, il y a des interrelations dans une partie externe, parce que c'est le lexique qui établit le pont entre l'intérieur et l'extérieur, entre le système linguistique et ce qu'on doit exprimer. Le lexique est alors un dépôt de toutes les données culturelles. [...] Quand j'ai pris ma retraite il y a 28 ans, je pensais à la question de l'inclusion sociale. Comment partager avec les autres toutes mes préoccupations de la linguistique et de la méthodologie d'analyse ? J'ai toujours été un théoricien et un praticien, alors je me suis dit que c'est par le lexique que j'allais opérer. Si j'allais faire un manuel ou quelque chose de similaire, la diffusion serait moindre que celle des dictionnaires, dont la diffusion est plus grande, parce qu'il y a encore cette idée du dictionnaire comme un réservoir et qu'il enseigne tout. La « bible des nuls », comme on dit, n'est-ce pas ? On cherche tout dans le dictionnaire, c'est vrai qu'on y trouve beaucoup de choses, mais il ne mentionne pas tout. En effet, quand j'allais prendre ma retraite, j'assistais à une émission télévisée qui s'appelait « O bem amado » (Le bien-aimé), de Dias Gomes, qui est un excellent écrivain au niveau des jeux de mots, il jouait beaucoup sur le système dérivationnel de la langue. Le personnage principal inventait des mots, mais rien en dehors du système linguistique. Et tous comprenaient : par exemple, au lieu de « dorminhoco » (dormeur), il disait « \*dormideiro » (\*dormirier). Et cela montre aussi la fonction du lexique, il sert à tout, à condition qu'on ait la syntaxe de la langue en main, sans mélanger les choses.

**SB :** Ah, vous êtes un passionné du lexique...

**FSB :** Oui. Voyez-vous, un jour, je regardais une bibliothèque et j'ai vu le dictionnaire de Francisco Fernandes. Un dictionnaire de régime verbal, [Francisco Fernandes, *Dicionário de regimes e verbos*, Porto Alegre, Globo, 1940]. Il était un amateur du langage, et il était *gaúcho* [originaire du Rio Grande do Sul, Brésil]. Il a publié ce dictionnaire [...] en faisant des recherches à sa façon, prélevant les prépositions qui régissent les verbes et les enregistrent. Son dictionnaire comptait environ 8 mille verbes et ce même dictionnaire

a reçu un prix [en 1942, le prix Francisco Alves, de l'Académie Brésilienne de Lettres] et a été pendant plus d'une trentaine d'années, jusqu'en 1975/76, le seul dictionnaire de verbes qui circulait au Brésil et que tout le monde consultait. Alors, je me suis dit que j'allais faire quelque chose sur des bases plus récentes, faire une recherche dans le lexique, une recherche dans les textes pour créer un autre dictionnaire de régimes. J'ai alors convoqué tous mes collègues du département de linguistique pour une réunion, j'étais directeur de la faculté à cette époque et leur ai dit : allons faire un dictionnaire de notre département, mais les intérêts étaient très variés. J'avais ma méthodologie, ma façon de travailler, et les autres, la leur. Finalement, nous sommes restés à 3 ou 4.

**SB :** Etait-ce là, la première équipe de votre dictionnaire de verbes ?

**FSB :** En effet, c'était la première équipe. Tereza [Maria Tereza Camargo Biderman] rejoignit, à ma demande, notre groupe quelque temps plus tard, quand l'UNESP s'est restructurée. Mais elle était plus intéressée par la recherche diachronique, de lexicologie, alors elle m'a dit qu'elle préférerait travailler à part. J'ai été motivé à élaborer un dictionnaire de verbes car celui de Francisco Fernandes devait être remplacé. Il n'était pas mauvais, mais il n'avait pas été mis à jour. C'était, certes, un dictionnaire contextualisé, muni d'exemples pour chaque régime avec des citations, mais d'auteurs anciens.

**SB :** Et c'est à ce moment que vous avez commencé à créer le corpus d'Araraquara ?

**FSB :** D'une certaine manière, oui. C'était une salle comme celle-ci. Je voulais qu'on fasse le prélèvement dans les textes, à partir d'une liste d'œuvres, d'un échantillon de 50 pages chacune et qu'on enregistre le tout sur de petites fiches. Nous n'avions pas encore d'ordinateur ici. L'ordinateur n'était pas encore une chose commune et nous avons travaillé à l'ancienne. Nous avons travaillé ainsi avec les collègues qui sont restés. Si je voulais savoir comment la langue fonctionnait, je devais faire le prélèvement dans des textes réels. Des textes littéraires, comme avant.

**SB :** Avez-vous choisi des textes plus récents ?

**FSB :** Des textes à partir du XX<sup>e</sup> siècle. J'ai alors fait une demande de financement à la FAPESP [Fundação de Amparo à Pesquisa do Estado de São Paulo, Fondation d'aide à la recherche de l'état de São Paulo] - j'ai toujours fait ça -, on m'a aidé et j'ai embauché moi-même des étudiantes en lettres pour le prélèvement, qui était transcrit sur des fiches. Ensuite, j'ai instauré un système de boîtes où les fiches étaient classées. Et c'est dans cette salle que la mise en place du corpus a commencée. C'était extrêmement précaire. On ne pouvait pas retrouver de nouveau une information. Je leur disais de faire le prélèvement par la fréquence, de prélever les constructions différentes - les plus communes apparaissant dans 2 ou 3 exemples, pour varier le plus que possible. J'ai aussi réalisé un schéma d'abréviation pour chaque œuvre, un schéma que j'utilise jusqu'à ce

jour, où la citation est suivie d'un sigle. Auparavant, on donnait l'exemple, ensuite le nom de l'auteur, l'œuvre, la maison d'édition, l'année et la page.

**SB :** Oui, pas de moyens, ça demandait beaucoup d'espace, n'est-ce pas ?

**FSB :** Oui, en quelque sorte, après beaucoup d'erreurs, j'ai réussi - je suis très têtu et je ne lâche pas - à publier le *Dicionário Gramatical de Verbos do Português Contemporâneo do Brasil* [Dictionnaire grammatical de verbes du portugais contemporain du Brésil] (São Paulo, Editora Unesp, 1990), qui a déjà eu beaucoup d'éditions, avec la description de chaque verbe, composé d'environ 5 mille verbes.

**SB :** Et tous ces verbes, on les a analysés par la grammaire de valences, n'est-ce pas ?

**FSB :** Oui, plus précisément avec la syntaxe de Tesnière. Peu après, comme la valence m'intéressait, j'ai travaillé là-dessus, en publiant *Uma gramática de valências para o português* [Une grammaire de valences pour le portugais] (São Paulo, Editora Ática, 1996). A cette époque-là, j'ai pris ma retraite. Alors je me suis dit que j'allais diffuser mon travail par les dictionnaires, parce qu'un dictionnaire, tout le monde en achète, quel que soit le sujet, il occupera toujours une position privilégiée par rapport à un livre quelconque. Lorsque l'ordinateur est apparu, j'ai eu l'idée de réaliser un dictionnaire plus général.

**SB :** Et votre université, l'UNESP, a-t-elle soutenu vos projets ?

**FSB :** A vrai dire, l'appui ne venait pas à proprement parler de l'université. Elle nous a offert l'infrastructure du département, les personnes qui ont voulu y participer, mais l'appui ne venait que de la FAPESP. L'UNESP, rémunère le professeur à temps plein pour qu'il fasse aussi des recherches. J'ai suivi divers cours et j'ai fait des stages à l'étranger, notamment en France, ainsi qu'aux Etats-Unis étant donné que mon intérêt penchait vers la syntaxe, où je fus invité en tant que professeur invité.

**SB :** Et aux Etats-Unis, auprès de qui avez-vous travaillé ?

**FSB :** En Californie, à Berkeley, avec Fillmore, sur la question des cas, très liée à la valence. C'est ce qui m'a conduit à une méthodologie de la description syntaxique et qui m'a donné envie de me dédier aux dictionnaires. Dès cette époque, j'ai commencé à faire mes propres observations sur les dictionnaires brésiliens et j'ai constaté que dans ces dictionnaires la citation avait peu d'espace, à commencer par celui d'Aurélio [Buarque de Holanda Ferreira, auteur du *Novo Dicionário Aurélio da Língua Portuguesa*, Nouveau dictionnaire Aurélio de la langue portugaise, publié aujourd'hui chez Editora Positivo, Curitiba], qui est notre plus grand auteur de dictionnaires et l'un des plus respectables. On a aussi Michaelis [*Moderno Dicionário da Língua Portuguesa*, Dictionnaire moderne de la langue portugaise], chez Melhoramentos. Ces deux dictionnaires sont deux grands

réservoirs lexicaux. A une époque où la lexicographie n'était pas une discipline dans les facultés brésiliennes, à partir de laquelle on aurait pu étudier les principes de la lexicographie, les technologies... mais on n'avait pas ça, les dictionnaires étaient élaborés par des érudits, indépendamment de leur formation. Mais nous avons eu le professeur [Antônio] Houaiss, qui était à la fois un grand érudit et un philologue. Vous pouvez voir que son dictionnaire [*Grande Dicionário Houaiss da língua Portuguesa*, publié chez Objetiva] a une orientation philologique, il présente beaucoup de problèmes, mais c'est une référence philologique importante. C'est plutôt un dictionnaire pour spécialistes, parce que nous savons sélectionner ce que nous voulons ou ne voulons pas. De même pour celui d'Aurélio.

**SB :** C'est-à-dire que vous avez introduit les premiers dictionnaires brésiliens élaborés à partir d'une approche lexicographique plus moderne.

**FSB :** Non, j'ai commencé par la lexicographie traditionnelle, en suivant ses principes. Et quand je dis traditionnelle, au Brésil, je dis « d'extraction française ». C'est par là que j'ai commencé. Et c'est justement la raison pour laquelle je suis allé en France. Alors j'ai voulu rencontrer Madame Rey-Debove, très sympathique, qui était chef de rédaction au Robert. Je l'ai rencontrée à plusieurs reprises - et ceci s'est passé après la production du dictionnaire de verbes, lorsque je commençais à élaborer le dictionnaire général.

**SB :** C'est-à-dire le *Dicionário de Usos [do Português Contemporâneo do Brasil*, Dictionnaire d'usages du portugais contemporain du Brésil] ?

**FSB :** Oui. J'ai fait comme vous, j'ai préparé une série de questions à lui poser. Alors, elle m'a dit, à part les questions générales concernant l'organisation des dictionnaires et le reste : « Je vois que vous êtes très sûr de ce que vous voulez faire », et elle n'a pas dit, « préparé », mais « très sûr », et en effet j'étais vraiment convaincu de mon projet. Et elle a ensuite continué : « Vous voulez élaborer un dictionnaire du portugais, mais vous voulez trouver une méthodologie différente de celle que vous avez, vous la cherchez ici et ailleurs, alors c'est une chose que vous devez faire dans les deux langues, n'est-ce pas, une analyse contrastive ? » Elle n'a pas exactement utilisé ces termes, à peu près, . Je connais aussi le français, j'ai été également professeur de français au Brésil avant d'aller en France. Non pas à l'université, mais au secondaire. De toute façon, je connaissais davantage le système verbal, aussi bien en portugais qu'en français, alors j'ai pensé à quelques questions mettant en rapport les articles de verbes du Robert avec ceux que je voulais préparer et ainsi je passerais aux autres classes de mots. Petit à petit, je l'interrogeais sur ces points. Durant cette période, je commençais la préparation du corpus d'Araraquara.

**SB :** Vous parlez du corpus électronique ?

**FSB** : Oui.

**SB** : Et c'était environ à quel moment ?

**FSB** : Au début des années 1980, vers 1981, 82. Je n'étais plus à la direction [de la faculté à Araraquara, au Brésil] en 1979... A cette époque, on a commencé de la même manière : au lieu de copier sur des fiches, j'ai embauché du personnel afin de copier les textes sur le support informatique, à l'aide de ces disquettes, vous rappelez-vous ? A une époque où l'ordinateur n'était pas adapté au portugais, on ne pouvait donc mettre ni d'accents ni cédilles. L'horreur !

[...]

**SB** : Et ces petites fiches en papier, se sont-elles perdues ou les a-t-on récupérées ?

**FSB** : Non, elles se sont perdues. [...] Mais j'ai tout refait.

**SB** : Et ce corpus, provenait-il des journaux ou venait-il encore de la littérature ?

**FSB** : A ce moment-là, j'ai commencé à étudier la linguistique de corpus, le montage des corpus, ainsi que le Brown Corpus américain en anglais. C'était un corpus de 5 millions de mots, à partir duquel on a monté l'American Heritage Dictionary, avec 80 mille entrées, dont les textes avaient une origine variée. Alors, je me suis dit, je vais faire la même chose, je vais compiler un corpus de 5 millions de mots, à partir de techniques de montage de corpus que j'avais apprises en France, avec Martinet ; mais des corpus variés, avec une seule date de départ, c'est-à-dire un corpus d'usages du portugais contemporain du Brésil. Et qu'entends-je par contemporain ? Arbitrairement, à partir de 1950. Alors les textes s'étendent de 1950 à 1980, plus ou moins, l'année où nous nous trouvions. Des textes du Brésil entier, basés sur la prose et non pas la poésie, mais aussi des textes techniques provenant de collections comme « Qu'est-ce que la chimie ? la pharmacie ? la mythologie ? le climat ? », de Melhoramentos, de Ática, la collection Princípios [Principes], des livres techniques regroupant tous les domaines que j'ai pu trouver. De même que des textes journalistiques, issus de journaux de toutes les capitales du pays - j'ai des amis partout, alors on m'envoyait des journaux tous les trois mois -, et les techniciens les recopiaient... Alors je prenais ces copies et je mettais en place les recherches, de façon arbitraire, car il n'y avait pas d'autres moyens, on n'avait que l'ordinateur, sans autres ressources ou outils. En arrivant à presque 5 millions de mots, toujours avec le financement de la FAPESP, je me suis dit : il faut que j'apprenne à travailler avec ce considérable corpus. Ce n'est pas possible que ça soit comme ça ! C'était très dur ! J'ai alors fait un prélèvement sur ce matériel dont le résultat a été de 30 mille occurrences de mots... J'avais bien sûr tous les mots grammaticaux, parce que ceux-ci apparaissent dans tous les contextes. Mais comment faire avec les mots lexicaux, ceux qui font la connexion avec le monde externe ? Leurs occurrences

dépendent de toute sorte de sujets, et même si j'avais beaucoup de textes de journaux, cela ne suffisait pas pour concrétiser mes objectifs. Mais comment, avec 5 millions et quelques, l'American Heritage était-il sorti avec 80 mille entrées ? C'est alors que je me suis rendu compte, après avoir réfléchi et beaucoup lu à ce sujet, qu'il s'agissait des natures des deux langues. L'anglais est essentiellement une langue syntagmatique et le portugais est paradigmatique. Cette dernière est une langue flexionnelle, alors qu'en anglais tout se résout dans les textes avec les particules. On dit tout avec des mots bien déterminés. Le portugais non, il n'a pas ça, il y a toutes les dérivations et les flexions, et il faut alors avoir beaucoup de textes pour voir la représentativité. A ce moment-là, j'ai été invité en tant que professeur à Provo, dans l'état d'Utah, une ville à côté de Salt Lake City, la capitale. C'est une région Mormon. La Brigham Young University est une grande université où on travaille sur les langues des Mormons, qui ont un grand intérêt pour le prosélytisme. C'est un important centre d'études de langues et d'informatique, de montage de textes sur ordinateur. La personne qui m'avait invité avait un fils qui travaillait justement sur ce sujet. Et il a tout arrangé pour moi, il parlait bien portugais, car sa mère était brésilienne. Il a tout arrangé et m'a appris à travailler sur le corpus.

**SB :** Avec les mots de recherche.

**FSB :** Oui, je cherchais les mots et élaborais à la fois l'article sur l'ordinateur, en alternant recherche et montage... Aujourd'hui c'est plus rapide, à ce moment-là c'était... Bon, c'est là que j'ai appris à utiliser ce système. Lors de mon retour au pays, je savais déjà comment me débrouiller.

**SB :** C'était dans les années 1980 ?

**FSB :** Oui, à la fin des années 1980. Je crois que j'ai été là-bas vers 1984, 1985... alors ceci s'est passé vers 1988, 89. Mais, quand je suis arrivé au Brésil, on avait commencé ce qu'on appelait la lecture optique. Alors, au lieu de taper, on pouvait scanner. Et puis après on devait tout corriger, car de nombreuses fautes demeuraient... C'était de la folie ! Maria Tereza [Camargo Biderman] m'avait dit qu'elle allait demander de l'argent à la FAPESP pour acheter cette machine. Et elle l'a achetée, et a commencé à travailler seule là-dessus. Je lui ai dit : Tereza, ce n'est pas comme ça, ce n'est pas un travail que toi-même tu dois faire, il faut former quelqu'un qui puisse le faire à ta place, afin que toi, tu puisses travailler sur le texte déjà numérisé.

**SB :** Concernant les analyses.

**FSB :** Concernant les analyses, ce travail devait être fait par un technicien. Et alors, après un an et demi, d'un corpus de 5 millions, je suis passé à environ 70 millions... Voyez-vous, c'est un bond ! J'ai mis presque 5 ans pour en faire un de 5 millions... Ce

sont les avancées de l'informatique, n'est-ce pas ? Après, c'était beau. Et aujourd'hui n'importe qui peut numériser un texte. Il y a eu aussi un moment où les journaux *Folha de S.Paulo*, *Estado de S.Paulo* et la revue *Veja* - non, celle-ci c'est moi qui l'ai obtenue - publiaient leurs éditions en disquette et on pouvait les acheter. Je les ai ajoutées aussi au corpus. Actuellement nous avons aussi les bibliothèques virtuelles, desquelles on peut tout copier après autorisation - l'œuvre de Machado de Assis, par exemple, je l'ai retirée de la bibliothèque virtuelle. C'est pourquoi le corpus a aujourd'hui 220 millions de mots de textes en portugais brésilien, de la prose, de la poésie, etc.

**SB :** A partir de 1950 ?

**FSB :** Non, à partir de 1500.

**SB :** Parce que vous avez aussi ajouté le corpus du DHPB [Dictionnaire historique du portugais brésilien] ?

A vrai dire, que s'est-il passé ? Le *Dicionário de usos* a été publié en 2002 ; son corpus avait environ 70 millions de mots. On a mis 4 ans pour le publier, à cause de questions concernant la maison d'édition Il était prêt vers 1996,97, mais il n'a été publié qu'en 2002. C'était aussi une édition assez difficile à réaliser étant donné que ce dictionnaire est sémantico-syntaxique, et c'est là où réside son caractère novateur.

**SB :** Concernant les analyses, n'est-ce pas ?

**FSB :** Jusqu'alors, ceux qui créaient nos dictionnaires étaient des érudits préoccupés de la lexicologie et des principes généraux de la lexicographie. Puis, nous avons eu la préoccupation philologique, toujours diachronique, de l'histoire de la langue, pour enregistrer ces divers aspects. Tous ces travaux ont suivi la lexicographie traditionnelle. Quant à moi, j'ai apporté à la lexicographie nationale ce côté novateur des principes d'analyse de la description linguistique, d'analyse syntaxique et sémantique. Les trois grands dictionnaires cités auparavant, focalisent l'item lexical et la partie lexicologique, où l'entrée présente toute les possibilités du point de vue de la formation des mots. Tandis que la liste d'acceptions des articles que je propose suit un ordre sémantique.

**SB :** Et la fréquence ?

**FSB :** C'est Maria Tereza qui s'est toujours occupée de cette partie et qui y a travaillé. De la fréquence, je n'en utilise que les ressources de base. On m'a appelé l'année dernière [2012], en septembre, à l'ABL [Académie Brésilienne de Lettres] pour faire une conférence. J'étais ravi de l'invitation ; on voulait que je parle de mes travaux et sur les nouvelles directions de la lexicographie au Brésil.

**SB :** C'est justement une des questions que j'ai à vous poser.

**FSB :** C'est ce qu'ils pensent que j'ai réalisé, et je crois que j'ai réellement fait cela.



**SB :** Oui.

**FSB :** Dans le sens suivant : le lexicographe doit avoir une formation technique spécifique, suivie à l'université, et non pas au secondaire. Il est un technicien universitaire, puisque le lexicographe n'est pas un scientifique mais un technicien, parce que la lexicographie n'est pas une science, elle est une technique. Le lexicographe se préoccupe des études de langage et va chercher des principes linguistiques pour concevoir son dictionnaire. Comme dans mon cas, il doit connaître la syntaxe et la sémantique, et avant ça, la morphologie pour connaître la structure des mots, il doit alors être un scientifique du langage. Donc, le linguiste est un scientifique des sciences humaines.

**SB :** Je pense que vous n'avez pas seulement innové ici, mais aussi ailleurs, car même les grands dictionnaires français, par exemple, n'ont pas cette analyse de base qu'ont vos dictionnaires.[...] En fait, je ne connais pas de dictionnaire qui soit comme les vôtres. [...] Et concernant le Dicionário da UNESP [Dictionnaire de l'UNESP] ?

**FSB :** [...] C'est un dictionnaire fondé sur corpus et sur la fréquence, mais il ne comporte aucune citation, parce qu'il est adressé au lycéens. J'ai alors retiré [du dictionnaire d'usages] tout ce qui, à mon avis, n'avait pas d'intérêt au niveau du secondaire et j'ai contrôlé les entrées à partir du corpus d'Araraquara. Les mots-vedettes sont donc des mots d'usage de la langue écrite du pays. Mon modèle pour la réalisation de ce projet est le Longman, qui est un dictionnaire pour étrangers. Je m'intéresse aussi à cette perspective , c'est-à-dire, un dictionnaire non seulement pour l'étudiant monolingue, mais également pour le bilingue et l'étranger. [...] Le Dicionário da UNESP présente alors des phrases, non pas construites à partir de la seule expérience du lexicographe, mais adaptées d'un corpus. Le Longman est ainsi, c'est l'anglais de la BBC.

**SB :** Oui, adapté. Tout dictionnaire a un fond pédagogique. Mais pourrait-on dire que celui de l'UNESP a une fonction plus pédagogique ?

**FSB :** Il a été élaboré sur une demande du Ministère brésilien de l'éducation pour les trois classes de terminales de l'enseignement secondaire. C'était aux environs de 2004, 2005. Aujourd'hui l'équipe est formée de 5 personnes.

**SB :** Qui sont-ils ?

**FSB :** , Maria Helena de Moura Neves, Beatriz Nunes de Oliveira Longo, Marina Bortolotti Bazzoli, Sebastião Expedito Ignacio, mais ce dernier est malheureusement décédé, Alexandre et moi. [...] Trois de mes dictionnaires ont été choisis par le ministère, à partir du PNLD (Programme national du livre didactique).

**SB :** Vous parlez des dictionnaires de verbes, d'usages et de l'UNESP ?

**FSB :** Non, seulement celui de l'UNESP et de ceux qui ont été édités par la maison

d'édition Positivo, dont les noms sont Palavrinha Viva [Petit mot vivant] et Palavra Viva [Mot vivant]. Le Palavrinha Viva est très élémentaire et illustré ; son illustrateur est très doué.

**SB :** C'est pour l'enseignement primaire ?

**FSB :** Oui, et Palavra Viva, pour les premières classes du secondaire.

**SB :** Palavrinha Viva et Palavra Viva dérivent-ils du Dictionnaire de l'UNESP ?

**FSB :** Plus ou moins. La conception de dictionnaires, quel que soit le niveau, a besoin d'un analyste de langues, pour qu'il puisse vérifier à qui s'adresse le dictionnaire et quels textes de base est utilisé. Pour les faire, on a dû constituer un autre corpus qui s'appelle Positivo et qui reste à la maison d'édition. Pour ce faire, nous avons utilisé des livres didactiques et tout ce qui concerne la littérature pour enfants et pour jeunes afin de relever quels étaient les mots les plus usités.

**SB :** Que faites-vous en ce moment ?

**FSB :** J'organise un dictionnaire général à partir d'un corpus de près de 230 millions de mots du portugais brésilien écrit depuis le XVI<sup>e</sup> siècle, depuis la Lettre de Caminha [premier document écrit sur le Brésil par Pero Vaz de Caminha]. Actuellement, il est en cours d'édition et je crois qu'il sera publié cette fin d'année [2013] ou au début de l'année prochaine [2014]. Le titre sera Dicionário Geral do Português Brasileiro [Dictionnaire général du portugais brésilien] ou Grande Dicionário do Português do Brasil [Grand dictionnaire du portugais du Brésil]. Cependant, le mot portugais « grande » a été mis en question, parce qu'avec « grande » on imagine la quantité de volumes. Mais on peut aussi s'imaginer « grande » par son côté exhaustif. Dans un sens...

**SB :** ...chronologique.

**FSB :** Oui, 230 millions... Il y a beaucoup de textes, de types divers, portant sur toutes les possibilités de structures syntaxiques du portugais brésilien. C'est exhaustif du point de vue syntaxique, parce que le lexique est inépuisable. Pour la nominata ou la nomenclature, nous avons déterminé deux conditions : tout mot ayant une fréquence différente de zéro et tout mot dont le contexte présente une structure syntaxique acceptable. [...] Dans ces conditions, nous en sommes arrivés à 113 mille entrées.

**SB :** Et ce dictionnaire, que vous allez publier dans quelques mois, il a aussi la même base analytique que les autres ?

**FSB :** Oui. Cette méthodologie d'analyse, je l'ai perfectionnée à partir du dictionnaire de verbes. Ce sont les mêmes critères de lecture. On ne peut pas tout simplement prélever un mot d'un texte, il faut l'interpréter dans le texte, comprendre sa pertinence, sa structure syntaxique, base de l'analyse. [...] A la différence du dictionnaire

d'usages, dont la description syntaxique et sémantique est conséquente, ce qui est parfait pour nous les spécialistes, mais encombrant pour le non spécialiste. Dans le dictionnaire de l'UNESP, d'une même orientation, l'analyse n'est pas explicitée. Dans le dernier qui est en préparation... il s'appellera peut-être Dicionário Documentado do Português do Brasil [Dictionnaire documenté du portugais du Brésil], parce que toutes les occurrences, tous les usages y sont répertoriés, il y a un sigle qui renvoi à l'œuvre source. On précise « du Brésil » dans le titre parce qu'il est limité à cet espace. Si on incluait le portugais européen, africain, etc... ce serait une tâche ardue. Pour les analyses que nous faisons, nous devons tout délimiter dans le temps et dans l'espace. Concernant le temps, on se base sur 500 ans de portugais au Brésil, comprenant tout ce qu'on a écrit au Brésil, sur le Brésil durant cette période. Les premiers textes que nous possédons, ceux qui ont été vraiment produits au Brésil, ont été, bien sûr, rédigés par des Portugais c'est-à-dire des jésuites et des provinciaux de l'Ordre jésuite qui sont venus catéchiser les indigènes, comme Fernão Cardim, Manuel da Nóbrega, José de Anchieta, pour ne pas en citer d'autres. Ces trois jésuites principaux envoyaient périodiquement des rapports à l'Ordre au Portugal. Alors on s'aperçoit que ce n'est plus que le portugais du Portugal. On y trouve des descriptions, par exemple, « il existe un petit fruit, rond et rouge, très beau, que les natifs appellent pitanga ». Ce mot n'existait pas encore là-bas.

**SB :** Et du point de vue de l'orthographe ?

**FSB :** Etant donné que mon intérêt n'est pas philologique, mais linguistique, [...] nous avons, du point de vue diachronique, deux informations. Tout d'abord l'origine, si le mot est un africanisme, s'il vient du français, du turc, du persan. C'est une référence toute simple. Puis il y a la question concernant la dernière occurrence temporelle du mot : les variantes, par exemple, « dous » et « dois » [deux] ; le dernier usage de ce mot apparaît dans une comédie d'Arthur de Azevedo dans la fin du XIXe. Et c'est très intéressant de relever ce fait, parce que c'était un jeu de mots, où l'usage de « dous » montrait un certain pédantisme du personnage, ce qui permet de constater que c'était déjà une forme de pédantisme de distinguer ces deux occurrences, dire « dous » au lieu de « dois ». [...]

**SB :** Y a-t-il une rubrique pour marquer des occurrences comme celle-là ? [...]

**FSB :** Il y aura un signe, une petite étoile par exemple, qui informera si la variante est peu utilisée ou très peu utilisée. [...] Le dictionnaire dispose d'une notation spéciale pour ces cas. [...] Ce dictionnaire cherche à documenter sur tous les types syntaxiques et se différencie des autres par le classement des acceptions dans les articles, lequel obéit en général à un ordre syntaxique, et non pas sémantique, surtout pour la catégorie verbale. On part de la structure la plus simple vers la plus complexe. Si c'est un verbe,

on présente d'abord l'intransitif, ensuite le transitif direct, après l'indirect et ainsi de suite, de sorte que le premier sens qui apparaît, dans une structure syntaxiquement plus simple, ne l'est pas forcément du point de vue sémantique. [...] Pour les noms, par exemple, c'est quelque peu similaire aux autres dictionnaires, car la syntaxe des noms est plus simple que celle des verbes. [...]

**SB :** La nomenclature du dictionnaire de verbes a-t-elle été utilisée dans l'élaboration des autres dictionnaires ?

**FSB :** Oui, nous en avons profité à de nombreux niveaux pour les autres dictionnaires. Et il y a cette histoire, où avant les dictionnaires se copiaient beaucoup entre eux. Comme on ne travaillait pas sur corpus, il apparaît dans ces ouvrages beaucoup d'apud. L'un s'inspirait de l'autre.

**SB :** Vous avez profité de vous-même !

**FSB :** Oui, mais à partir des textes. Il y a toujours un vocabulaire de base qui reste présent. Dans le dictionnaire de l'UNESP, nous avons mis une explication sur les niveaux lexicaux de base. Les 20 mille premiers mots sont ceux de haute fréquence que nous pouvons reconnaître par leur polysémie. Plus un mot est fréquent, plus la polysémie peut s'étendre. [...] Pour les adverbes, par exemple, j'ai utilisé les analyses réalisées par Ataliba Teixeira de Castilho, un travail bien construit.

**SB :** Quels conseils donneriez-vous aux jeunes lexicographes ? [...]

**FSB :** Non seulement en lexicographie, mais dans n'importe quel domaine de la connaissance, dans notre cas, en sciences humaines, on ne peut pas entrer en amateur ou comme un simple amateur. On doit avoir une attitude professionnelle et suivre des cours dans le domaine. Pour ceux qui vont travailler sur le lexique, il est indispensable de suivre des cours de description de langues et aussi de langues étrangères, car même si l'on va travailler sur sa propre langue, on doit s'appuyer sur une autre, proche ou non. Il faut aussi s'informer des nouveautés informatiques dont les outils permettent de découvrir des choses très intéressantes sur le lexique. De plus au Brésil il y a beaucoup à faire par rapport aux vocabulaires spécialisés, l'étude des domaines, des champs sémantiques...

**SB :** Et aussi en lexicographie bilingue. On en a vraiment besoin.

**FSB :** La lexicographie bilingue est un domaine presque vierge en portugais. Je ne connais personne qui fasse quelque chose de semblable, fondé sur la linguistique contrastive. Et nous sommes bien équipés pour diriger des travaux sur ce sujet. Le champ est libre, il y a beaucoup de matériel, mais il faut aussi être préparé pour travailler dessus et avec ce matériel. Car il est possible de travailler sur corpus en ne

faisant que des prélèvements, comme des listes de mots.

**SB** : Sans aucune analyse, n'est-ce pas ?

**FSB** : Aucune, alors...

**SB** : En bref, votre conseil aux jeunes lexicographes : soyez d'abord un linguiste et faites des analyses linguistiques.

**FSB** : Ils doivent aussi décrire, Ils doivent avoir d'abord une bonne formation en structuration du vocabulaire, en lexicologie et en morphologie. Premièrement dans leur propre langue et puis dans une autre, qu'elle soit une langue de tradition écrite ou pas, comme nos langues indigènes. Ainsi Ils doivent se rendre compte comment la langue fonctionne pour qu'ils puissent prélever le vocabulaire qui y circule. Afin que leur dictionnaire soit utile. Le dictionnaire est un instrument pédagogique de première catégorie.

**SB** : Je vous remercie beaucoup !

**FSB** : Je vous en prie.