

## SOBRE AS LÍNGUAS INDÍGENAS E SUA PESQUISA NO BRASIL

Aryon Dall'Igna Rodrigues

**DIVERSIDADE E MULTIPLICIDADE LINGÜÍSTICA NO PASSADO** A única estimativa de que dispomos sobre a diversidade das línguas indígenas existentes no Brasil há 500 anos, antes do início da colonização desta parte da América do Sul pelos europeus, é a que foi apresentada, em 1992, na Reunião Anual da Sociedade Brasileira para o Progresso da Ciência (Rodrigues, 1993a, 1993b). Segundo essa estimativa, teria sido de cerca de 1,2 mil o número de diferentes línguas faladas em nosso atual território pelos povos indígenas. O ponto de partida para essa estimativa foi uma relação de 76 povos indígenas que se encontravam numa estreita faixa paralela à costa leste, desde o rio São Francisco, ao norte, até o Rio de Janeiro, ao sul, feita pelo padre jesuíta Fernão Cardim no século XVI (Cardim, 1978 [manuscrito de 1584]). Nessa lista, Cardim referiu-se explicitamente à identidade ou à diferença das línguas faladas por esses povos, deixando claro que, ao todo, se tratava de 65 línguas distintas entre si e distintas da língua dos índios da costa, que eram os tupinambás (que incluem os tupiniquins, caetés, potiguaras, tamoios etc.), com os quais os portugueses mantinham contacto. Como alguns nomes na lista estão claramente na língua dos tupinambás – a mesma que hoje também é chamada de tupi antigo e que no século XVII foi denominada *língua brasileira* – e os demais estão grafados à maneira como os jesuítas escreviam essa língua, pode-se supor que as fontes de informação tenham sido os índios tupinambás e que aquela enumeração representasse o conhecimento destes sobre seus vizinhos mais imediatos. Ela deve ser bastante representativa para a área coberta, embora possa não ser exaustiva.

Apesar da grande diversidade de povos nativos no interior mais imediato à costa atlântica, uma característica da colonização européia do Brasil, não só da portuguesa, mas também das tentativas francesas, foi a de privilegiar o conhecimento do idioma dos tupinambás que era, como já no fim do século XVI foi consignado no título da gramática feita por José de Anchieta, “a língua mais usada na costa do Brasil” (Anchieta, 1595). Para a comunicação com os outros povos recorria-se a intérpretes indígenas. Uma consequência dessa situação é que, durante os três séculos do período colonial fizeram-se gramáticas e dicionários de somente três línguas indígenas: do próprio tupinambá, de que foram feitas duas (Anchieta, 1595, e Figueira, 1621), da língua kirirí (Mamiani, 1699) e da língua dos maramonins ou guarulhos. Desta última, elaborada pelo Padre Manuel Viegas com o auxílio do Padre Anchieta, estão, entretanto, perdidos todos os manuscritos (gramática, vocabulário e catecismo). As gramáticas do tupinambá e do kirirí foram publicadas nos séculos XVI e XVII e, assim, sobreviveram e puderam ser reproduzidas em novas edições (a de Figueira já no século XVII, as de Anchieta e Mamiani

só no século XIX), mas a de Viegas não foi publicada e se perdeu, assim como seus outros trabalhos sobre a língua. Também não foi publicado o dicionário da língua tupinambá, mas deste foram felizmente preservadas umas poucas cópias manuscritas, uma delas datada de 1621. Também o manuscrito do dicionário do kirirí, feito por um padre de nome João de Barros, até hoje não foi encontrado, da mesma forma como ainda não se tem notícia sobre a possível sobrevivência de manuscritos de gramática e dicionário dos capuchinhos franceses que atuaram no fim do século XVII e no início do XVIII no rio São Francisco, junto a um povo estreitamente aparentado aos kirirís, o povo dzubukuá, em cuja língua Frei Bernardo de Nantes publicou um catecismo (Nantes, 1709).

**DIVERSIDADE E MULTIPLICIDADE LINGÜÍSTICA NO PRESENTE** Presentemente, são faladas no Brasil 181 línguas indígenas. Esse número admite pequena margem de erro para mais ou para menos, devido principalmente à imprecisão, em alguns casos, da distinção entre línguas e dialetos (estes são variedades de uma língua tão pouco diferenciadas, que não dificultam a comunicação entre seus respectivos falantes). Nesse número podem estar incluídas duas ou três línguas que deixaram de ser faladas nos últimos cinco anos. Por outro lado, o Departamento de Índios Isolados da Funai, que monitora as informações sobre a existência de povos indígenas ainda sem contacto aberto com segmentos da nossa sociedade, admite que são perto de 20 os grupos de pessoas nessa situação. Alguns desses grupos podem falar línguas compartilhadas com outros já conhecidos, mas vários deles podem ser detentores de idiomas ainda desconhecidos.

A classificação científica das línguas é de natureza genética: incluem-se em uma mesma classe línguas para as quais há evidências de serem provenientes de uma mesma língua ancestral, analogamente à situação das línguas românicas ou latinas, que provêm do latim falado na Europa ocidental há cerca de 2.000 anos. Um conjunto de línguas que compartilham assim a mesma origem é o que tecnicamente se chama uma *família lingüística*. Na medida em que progride o seu conhecimento, as línguas indígenas brasileiras vêm sendo classificadas em famílias genéticas. Presentemente são distinguidas 43 famílias, algumas das quais consistem em uma só língua e caracterizam o que também se chama de “língua isolada”, termo pouco significativo, uma vez que freqüentemente esse isolamento decorre de acidentes históricos e, no caso das línguas do Brasil, mais provavelmente do processo colonizador, que exterminou os povos que falavam outras línguas de uma mesma família. De algumas famílias, embora haja documentação dos séculos passados que permite determiná-las ao menos aproximadamente, já morreram todas as línguas e, portanto, a própria família está morta. Esse é o caso de várias famílias lingüísticas do Brasil oriental, como a karirí, a kamakã e a purí. Por outro lado, entre algumas famílias têm sido reconhecidas propriedades comuns de natureza tal que só podem ser explicadas por uma origem comum mais remota do que as que justificaram a constituição de cada família. Nesse caso postula-se uma classe genética mais abrangente e de maior profundidade temporal, o *tronco lingüístico*. No Brasil reconhece-se um tronco bem estabelecido, o tupi, que compreende dez famílias, e outro de caráter ainda bastante hipoté-

tico, o macro-jê, abrangendo doze famílias. No **quadro 1** figuram as famílias lingüísticas segundo o estado atual do conhecimento, com as respectivas línguas, as siglas dos estados em que estas são faladas e o número de falantes.

**A PERDA DA DIVERSIDADE** A redução de 1200 para 180 línguas indígenas nos últimos 500 anos foi o efeito de um processo colonizador extremamente violento e continuado, o qual ainda perdura, não tendo sido interrompido nem com a independência política do país no início do século XIX, nem com a instauração do regime republicano no final desse mesmo século, nem ainda com a promulgação da “Constituição Cidadã” de 1988. Embora esta tenha sido a primeira carta magna a reconhecer direitos fundamentais dos povos indígenas, inclusive direitos lingüísticos, as relações entre a sociedade majoritária e as minorias indígenas pouco mudou. Graças à Constituição em vigor está havendo diversos desenvolvimentos importantes para muitas dessas minorias em vários planos, inclusive no acesso a projetos de educação mais específicos e com consideração de suas línguas nativas. Entretanto, ainda são grandes a hostilidade e a violência, alimentadas não só por ambições de natureza econômica, mas também pela desinformação sobre a diversidade cultural do país, sobre a importância dessa diversidade para a nação e para a humanidade e sobre os direitos fundamentais das minorias.

Os dados demográficos apresentados no quadro 1 têm diferentes graus de precisão. Alguns são exagerados para mais, outros para menos, em geral para mais, pois grande parte deles se refere à população e não ao número de falantes da língua indígena, que em muitos casos é inferior. Assim mesmo esses dados são reveladores da situação extremamente grave em que se encontra a maioria das línguas indígenas. No plano mundial tem-se considerado que hoje qualquer língua falada por menos de 100 mil pessoas tem sua sobrevivência ameaçada e necessita de especial atenção. Todas as línguas indígenas no Brasil têm menos de 40 mil falantes, sendo que a mais forte, a tikúna, falada no alto Solimões, apenas ultrapassa a marca de 30 mil. O aspecto mais grave está, porém, no outro lado do espectro demográfico, nas línguas infimamente minoritárias, com populações que não vão além de 1 mil pessoas. Essa é a situação de três quartos (76%) das nossas línguas indígenas e significa que é tarefa de alta prioridade e urgência a pesquisa científica que visa à documentação, análise, classificação e interpretação teórica dessas línguas, que em sua grande maioria só existem aqui. Igualmente prioritária é a promoção de ações que visem a assegurar aos povos indígenas as condições necessárias para continuar transmitindo suas línguas às novas gerações.

**A DOCUMENTAÇÃO CIENTÍFICA** As línguas são objetos fugidios, cujas manifestações faladas são momentâneas e se sucedem em alta velocidade, de modo que sua observação e sua análise científica dependem normalmente de uma fixação. Esta é obtida mediante a escrita, que se desenvolve primeiramente como fiel transcrição fonética dos enunciados ouvidos pelo pesquisador, freqüentemente gravados eletronicamente para poderem ser repetidos com precisão, e progride para uma representação fonológica mais abstrata, que permite registrar com fidelidade os dados relevantes para a análise gramatical e a

interpretação semântica. O lingüista treinado para efetuar essas operações tem de ter acesso a falantes nativos da língua e, para línguas ainda desconhecidas, esse é um processo que demanda, em condições boas de pesquisa, quatro ou mais anos para produzir uma boa descrição gramatical e um dicionário com registro amplo do vocabulário que cubra todos os domínios semânticos relevantes da cultura nativa. Para documentar adequadamente a língua de um povo culturalmente tão diferente quanto os índios, o lingüista precisa de uma percepção etnológica, razão por que sua pesquisa é também denominada etnolingüística.

Devido ao grande número de línguas faladas pelos povos indígenas do Brasil e à grave ameaça de desaparecimento que incide sobre a maioria delas, sua documentação científica requer grande número de pesquisadores em condições favoráveis de dedicação continuada. Em 1987, quando só uma universidade brasileira, a Unicamp, estava formando pesquisadores para trabalhar com línguas indígenas, foi criado no CNPq, como programa especial, o PPCLIB - Programa de Pesquisa Científica das Línguas Indígenas Brasileiras, destinado a fomentar tanto a pesquisa como a formação de pesquisadores. Não foi destinada nenhuma verba específica para esse fim, mas foi adotada uma política de preferência a projetos que se enquadrassem nas prioridades definidas pelo programa. Também foi obtida a colaboração da Finep, que se dispôs a apoiar ajustes de instalações. Deste apoio só pôde beneficiar-se uma instituição, já que outras não o solicitaram, ou não puderam satisfazer as condições mínimas da Finep. Entretanto, com o apoio do CNPq foi possível motivar alguns programas de pós-graduação em letras a abrirem espaço para pesquisas sobre línguas indígenas, além do programa de lingüística da Unicamp: UFG, UFPE, UFRJ, UFSC, UnB. Além de bolsas concedidas a alunos dessas universidades, principalmente em nível de mestrado, foi possível também outorgar bolsas para o doutorado no exterior, tanto nos Estados Unidos, como na Europa. Durante o governo Collor foram cancelados os programas especiais, inclusive o PPCLIB, mas felizmente a área de Ciências Humanas manteve informalmente as definições de prioridade para bolsas e auxílios na área de línguas indígenas. Assim, novos estudantes continuaram recebendo apoio do CNPq para investigar essas línguas. O clima criado pelo PPCLIB estimulou também apoio de algumas fundações estaduais, como a Faperj e a Fapesp.

**A FORMAÇÃO DOS PESQUISADORES** Entretanto, o número de pesquisadores é ainda muito pequeno, e isso não se deve ao CNPq, mas a outros fatores. Dentre os mais fortes cabe mencionar a organização das universidades brasileiras e a posição que tem nelas a lingüística. No nível de graduação a lingüística é ensinada nos cursos profissionalizantes de licenciatura em Letras, destinados basicamente à formação de professores para o ensino médio, e na maioria dos currículos seu papel é apenas ancilar. No nível de pós-graduação, poucas universidades têm programas específicos de lingüística, pois a maioria dos programas são projeções mais avançadas dos cursos de graduação, são programas simplesmente “de letras”. Isto reflete em boa parte a organização departamental das faculdades de letras, das quais muito poucas têm departamen-

## QUADRO DAS LÍNGUAS INDÍGENAS AINDA FALADAS NO BRASIL

## Famílias e línguas indígenas do Brasil | Localidade | nº de falantes

**Aikanã**

Aikanã, RO, 264

**Arawá**

Banawá, AM, 215; Deni, AM, 736; Jamamadí (Kanamanti), AM, 800; Jarawára, AM, 160; Kulína (Madihá), AC e AM, 2318; Paumari, AM, 870; Zuruahá, AM, 143

**Arikém**, tronco Tupi

Karitiána, RO, 206

**Aruák**

Apuriná (Ipuriná), AC e AM, 2779; Baniwa do Içana, AM, 5141; Baré, AM, 0; Kámpa (Aváninka), AC e AM, 813; Kuripáko, AM, 1.115; Maxinéri (Manchini), AM, 459; Mehináku (Meinacó), MT, 199; Palikúr, AP, 918; Paresi, MT, 1.293; Salumá (Enawenê-nawê), MT, 320; Tariána, AM, 55; Teréna, MS e SP, 15.795; Wapixána, RR, 6.500; Warekéna, AM, 491; Waurá, MT, 321; Yawalapiti, MT, 208

**Awetí**, tronco Tupi

Awetí, MT, 138

**Bora**

Miranha, AM, 613

**Boróro**, tronco Macro-Jê

Boróro, MT, 300; Umutina, MT, 0

**Chiquito**

Chiquito (Chiquitano), MT, 2.000

**Guaikurú**

Kadiwéu, MS, 1.592

**Guató**, tronco Macro-Jê

Guató, MS, 5?

**Irántxe**

Irántxe, MT, 326; Mynký (Münkü, Menky), MT, 78

**Jabuti**

Arikapú, RO, 19; Jabuti (Jeoromixi), RO, 123

**Jê**, tronco Macro-Jê

Kaingang, PR, SC, SP e RS, 25.000; Kayapó (Mebengokré), MT e PA, 8.148; Panará, MT e PA, 202; Suyá, MT, 334; Tapayúna, MT, 58; Timbira, MA, PA e TO, 4.445; Xavante, MT, 9.602; Xerente, TO, 1.814; Xokléng, SC, 757

**Jurúna**, tronco Tupi

Jurúna (Yudjá), MT, 278; Xipáya, PA, 2

**Kanoê**

Kanoê, RO, 6

**Karajá**, tronco Macro-Jê

Javaê, TO, 919; Karajá, GO e TO, 2.500; Xambioá, TO, 185

**Karib**

Arara do Xingu (Ukarangmã), PA, 195; Bakairi, MT, 950; Galibi (Kariña), AP, 28; Hixkariána, AM, 308?; Ikpéng (Txião), MT, 319; Ingarikó, RR, 675; Kalapálo, MT, 415; Kapóng (Patamóna), 50; Kaxuyána, PA, 69; Kuikúru, MT, 417; Makuxi, RR, 16.500; Matipú, MT, 119; Nahukwá, MT, 105; Taulipáng (Taurepá, Pemóng), RR, 532; Tiriyo (Tirió), PA, 735; Waimiri (Waimiri-Atroari), AM, 931; Waiwái, AM e RO, 2.020; Wayána, PA, 415; Yekuána (Mayongóng), RR, 426

**Katukina**

Kanamari, AM, 1327; Katawixi, AM, ?; Katukina, AM, 289; Txunhuá-djapá (Tsohom-djapá), AM, 100

**Kokáma** (Omágua, Cambeba),

língua mista, AM,

**Krenák**

Krenák (Botocudo), MG, 6

**Kwazá**

Kwazá (Kwayá, Coaiá), RO, 25

**Máku**

Máku, RR, 1

**Makú**

Dáw (Kamã), AM, 83; Húpda, AM, 1.431?; Nadéb, AM, 480?; Yuhúp, AM, 400

## Famílias e línguas indígenas do Brasil | Localidade | nº de falantes

**Maxakali**, tronco Macro-Jê

Maxakali, MG, 802

**Mawé**, tronco Tupi

Mawé (Sateré), AM e PA, 7.134

**Mondé**, tronco Tupi

Arara do Beiradão ou do Aripuanã, MT, 5?; Aruá, RO, 58; Cinta-larga, MT e RO, 1.900; Gavião (Ikóro, Digüt), RO, 436; Mondé, RO, Paitér (Surui de Rondônia), RO, 920; Zoró, MT e RO, 414

**Mundurukú**, tronco Tupi

Kuruáya, PA, 5; Mundurukú, AM e PA, 7.500

**Múra**

Mura, AM, ?; Pirahã, AM, 360

**Nambikwára**

Lakondê, RO, 1; Latundê, RO, 20; Mamãindê, MT, 103; Nambikwára do Sul, MT, 663; Sabanê, RO, 15

**Ofayé**

Ofayé (Opayé, Ofayé-Xavante), MS, 5?

**Páno**

Amawáka, AM, 220?; Katukina, AC e AM, 318; Kaxarari, AM e RO, 269; Kaxinawá, AC, 3964; Korúbo, AM, 250; Kulino (Kulina), AM, 20; Marúbo, AM, 1.043; Matsés, AM, 829; Nukini, AC, 458; Poyanáwa, AC, 403; Xawanáwa (Arara), AC, 200; Yamináwa, AC, 618; Yawanáwa, AC, 450

**Puruborá**, tronco Tupi

Puruborá, RO, 2

**Ramaráma**, tronco Tupi

Káro, RO, 184

**Rikbák**, tronco Macro-Jê

Rikbák (Rikbaksá, Canoeiro), MT, 909

**Samuko**

Chamacoco, MS, 40

**Tikúna**

Tikúna (Tukúna), AM, 32.613

**Trumái**

Trumái, MT, 120

**Tukáno**

Arapáso, AM, 328; Bará, AM, 39; Barasána, AM, 61; Desána, AM, 1.531; Juriti, AM, 35; Karapaná, AM, 42; Kubéwa, AM, 287; Miriti-Tapúya, AM, 95; Pirá-Tapúya, AM, 1.004; Siriána, AM, 17; Tuyúka, AM, 593; Wanána, AM, 447

**Tupari**, tronco Tupi

Akuntsú, RO, 6; Kepkiriwát, RO, 0; Makuráp, RO, 2.67; Sakirabiát (Mequém, Mequéns), RO, 66; Tupari, RO, 338; Wayoró (Ajurú), RO, 15

**Tupi-Guaraní**, tronco Tupi

Amondáwa, RO, 83; Anambé, PA, 2; Apiaká, MT, 2; Araweté, PA, 278; Asuriní do Tocantins (Akwáwa), PA, 303; Asuriní do Xingu (Awaeté), PA, 106; Aurê-Aurá, MA, 2; Avá-Canoeiro, GO, TO, 16; Diahói (Diarroi), AM, 50; Guajá, MA, 280; Guajajára, MA, 13.100; Guaraní (Kaiwá, Mbyá, Nhandéva), ES, MS, PR, RJ, RS, SC e SP, 30.000; Juma, AM, 7; Ka'apor (Urubu), MA, 800; Kamayurá, MT, 355; Kariúna, RO, 1; Kayabi, MT, 1.000; Língua Geral Amazônica (Nheengatú), AM, 3.000; Parakaná, PA, 800; Parintinín, AM, 156; Surui do Tocantins (Mudjetire, Akewára), PA, 185; Tapirapé, MT, 438; Tembê, PA, 820; Tenharim, AM, 585; Uruewawau, RO, 87; Wayampí, AP, 525; Xetá, PR, 3; Zo'ê (Jo'ê, Puturú), PA, 152

**Txapakúra**

Kujubim, RO, 27; Orowin (Oro Win), RO, 50; Torá, AM, 51; Urupá, RO, 150; Wari (Pakaanóva), RO, 1.930

**Yanomámi**

Yanomámi (Ninám, Sanumá, Yanomám, Yanomámi), RR, 11.000

**Yatê**, tronco Macro-Jê

Yatê (Camijó, Fulniô), PE, 2.930

tos de lingüística. E mesmo estes, onde existem, são compelidos a preocupar-se maximamente com a formação de professores para a língua majoritária, a portuguesa, e suas expressões literárias. Uma outra característica desfavorável da atual organização universitária é a compartimentação estanque dos departamentos e dos cursos, não só na graduação, mas também na pós-graduação. Em geral não existe espaço para currículos diferenciados e interdisciplinares. Um estudante de lingüística não pode integrar em seu currículo os créditos que necessita de antropologia ou de sociologia ou de estatística ou de história ou de informática, etc. Um orientador não pode dirigir o estudante para a integração de conhecimentos complementares, porque os requisitos dos cursos são estreitamente limitadores. Na pós-graduação, onde deveria haver mais flexibilidade, a imposição, pela agência reguladora do Ministério da Educação, de prazos mínimos rígidos com a previsão de penalidade aos cursos que os deixem ultrapassar, somada à exigência de estágio docente dentro desses mínimos, desestimula, se não impossibilita, a realização do trabalho de campo lingüístico que pode demandar um esforço e um tempo que não são levados em conta pelos avaliadores ministeriais.

Somando professores e estudantes, temos hoje no Brasil perto de uma centena de pessoas envolvidas em pesquisa de línguas indígenas, a maioria ainda estudantes, mas cerca de quarenta doutores. Com o apoio da Associação Nacional de Programas de Pesquisa e Pós-Graduação em Letras e Lingüística (Anpoll), veio consolidando-se, desde 1985, um Grupo de Trabalho sobre Línguas Indígenas (GTLI), o qual promoveu, em 2001, o primeiro encontro internacional, ao qual compareceram 90 pesquisadores, dos quais 12 estrangeiros (Cabral e Rodrigues, 2002). Dos brasileiros, 23 eram doutores e os demais, estudantes de graduação e pós-graduação. Destes últimos, alguns já concluíram o doutorado nos últimos dois anos, no Brasil ou no exterior, alguns destes ainda como reflexos do PPCLIB.

**CONCLUINDO** Embora, como se depreende deste histórico, o número de pesquisadores venha crescendo, esse crescimento não se correlaciona diretamente com a cobertura das necessidades mais urgentes da pesquisa das línguas indígenas. Os trabalhos realizados são em grande parte fragmentários, associados a dissertações de mestrado, sem que tenha havido oportunidade, para os respectivos autores, de dar continuidade à pesquisa. Também projetos de maior fôlego têm ficado inconclusos, ou passaram a alongar-se excessivamente, devido a contingências profissionais e outras, inclusive às dificuldades de financiamento. Certamente, dez anos após a desativação do PPCLIB, torna-se necessário pensar em novo programa especial de fomento à documentação, análise e descrição das línguas indígenas, que, por um lado, contemple não só o estímulo para o ingresso de novos pesquisadores nessa área, mas também a sustentabilidade dos bons projetos dentro de prazos razoáveis, em conjugação com as atividades de pós-graduação e com a necessidade de coordenar a pesquisa lingüística com o apoio a projetos de revitalização e promoção do uso das línguas nativas nas comunidades indígenas.

*Aryon Dall'Igna Rodrigues foi professor titular do Departamento de Lingüística da Unicamp; é professor do Laboratório de Línguas Indígenas da UnB; sócio fundador da Associação Brasileira de Lingüística (Abralín); e pesquisador 1A do CNPq.*

#### BIBLIOGRAFIA CITADA

- Anchieta, J. de. (1595). *Arte de grammatica da língua mais usada na costa do Brasil*. Coimbra.
- Nantes, Bernardo de. (1709). *Katecismo indico da língua Kariris*. Lisboa.
- Cabral, A. S. A. C., e A. D. Rodrigues. *Línguas indígenas brasileiras: fonologia, gramática e história*. Atas do I Encontro Internacional do Grupo de Trabalho sobre Línguas Indígenas da Anpoll. 2 volumes. Belém: Edufba. 2002.
- Cardim, F. *Tratados da terra e gente do Brasil*. 3ª edição. São Paulo: Companhia Editora Nacional/MEC. 1978.
- Mamiani, L. V. (1699). *Arte de grammatica da língua brasílica da naçam Kiriri*. Lisboa.
- Rodrigues, A. D. "Línguas indígenas: 500 anos de descobertas e perdas". *D.E.L.T.A.* 9.1:83-103. São Paulo. 1993a.
- Rodrigues, A. D. "Línguas indígenas: 500 anos de descobertas e perdas". *Ciência e Cultura* 95:20-26. 1993b.