

A classificação do tronco linguístico Tupí

Aryon Dall'Igna Rodrigues

Uma comunicação de Morris Swadesh publicada nos Anais do XXXI Congresso Internacional de Americanistas aventou um possível aperfeiçoamento da classificação das línguas americanas através do emprego da estatística lexical¹. Na mesma ocasião, apresentamos um trabalho sobre as chamadas línguas “impuras” do Tupí-Guaraní, servindo-nos do método léxico-estatístico².

A estatística já fora empregada anteriormente no estudo comparativo das línguas americanas; Chestmír Loukotka, por exemplo, utilizou-a como fundamento de seus estudos gerais de classificação. Baseando-se na contagem de vocábulos, ele distingue em suas famílias linguísticas de um lado línguas com “vestígios” estranhos, línguas com “intrusões” estranhas, línguas “mistas”, e de outro lado as línguas “puras”. Existe, entretanto, uma diferença entre o método de Loukotka e o da estatística lexical que empregamos. Loukotka baseou sua classificação sobretudo na comparação de 45 vocábulos “típicos”, incluindo diversas “palavras culturais”. Nós, ao contrário, comparamos, sob condições favoráveis, uma lista de 200 palavras, na maioria das vezes, porém, de 100 a 200 e, nos casos de material escasso, também listas com menos de 100 vocábulos. Mas a característica principal de nossa lista, que foi construída por Swadesh³, é a de incluir palavras pouco influenciadas pela mudança cultural e que, por isso mesmo, se conservam em todas as línguas relativamente

1. Morris Swadesh, *Towards a satisfactory genetic classification of Amerindian languages*. Anais do XXXI Congresso Internacional de Americanistas, São Paulo, 1955. Págs. 1001-1012.

2. Aryon Dall'Igna Rodrigues, *As línguas “impuras” da família Tupí-Guaraní*. Anais do XXXI Congresso Internacional de Americanistas, São Paulo, 1955. Págs. 1055-1071.

3. Morris Swadesh, *Amerindian non-cultural vocabularies* (mimeographed sheet). Cf. M. Swadesh, Lexico-statistic dating of prehistoric ethnic contacts. Proceedings of the American Philosophical Society, vol. 96, n. 1952. Págs. 452-463.

inalteradas no decorrer do tempo, podendo ser consideradas, portanto, como vocabulário básico. Em seus trabalhos comparativos, além disso, Loukotka procura os elementos estranhos, enquanto nós nos preocupamos em encontrar cognatos no sentido histórico. Em consequência disso, as línguas mistas desempenham papel preponderante em Loukotka (em um de seus trabalhos constatou a existência de 90 na América do Sul⁴), enquanto para nós é praticamente impossível admitir uma língua mista.

Acreditamos que, através da comparação de vocabulários básicos, algumas das maiores dificuldades da classificação de línguas sul-americanas poderão ser superadas, inclusive a questão das chamadas línguas mistas. À guisa de exemplo, citamos o caso da língua dos Maué: em 1942 Loukotka a considerou língua da família Tupí mesclada com Karíb e Aruák⁵; em 1950 denominou-a “langue très altérée” da família Tupí-Guaraní, com elementos aruák, karíb, além de outros, acreditando possível serem os “elementos tupí nessa língua simplesmente palavras emprestadas da língua geral, podendo-se considerá-la como isolada”⁶. Mas uma análise léxico-estatística do vocabulário básico do Maué demonstra a presença de quase 50% de elementos Tupí-Guaraní, o que depõe decisivamente em favor de sua filiação à família Tupí-Guaraní. Este resultado confirma a opinião de Curt Nimuendajú sobre o Maué, isto é, de que se trata fundamentalmente de língua Tupí, embora se diferencie do Guaraní-Tupinambá⁷. Outros exemplos seriam o Mundurukú, o Kuruáya, o Yurúna e o Xipáia, que, como o Maué, já foram por nós examinados em 1954.

Outra medida metodológica que também queremos salientar aqui e que pertence ao método léxico-estatístico de Swadesh é a distinção de determinados graus de parentesco linguístico. As classificações de Rivet e Loukotka⁸ agrupam as línguas em “famílias”, mas não estabelecem o âmbito de cada família, ou seja, quão diferentes duas línguas podem ser, sem que devam ser classificadas em duas famílias distintas. Para isso seria necessário limitar o âmbito das famílias. É óbvio que os supostos limites não podem ser rígidos, e devem ser estabelecidos mais ou menos empiricamente e em parte até ar-

4. Chestmír Loukotka, *Klassikation der südamerikanischen Sprachen*. Zeitschrift für Ethnologie, vol. 74 (1942), Berlim, 1944. Págs. 1-69.

5. Chestmír Loukotka, *Klass. ders üdam. Sprachen*, pág. 31.

6. Chestmír Loukotka, *Les langues de la famille Tupi-Guarani*. São Paulo, 1950. Págs. 27 e 29.

7. Curt Nimuendajú, *The Maué and Arapium*. Handbook of South American Indians, vol. 3. Pág. 246.

8. Chestmír Loukotka, *Klass. der südam. Sprachen*, op. cit. Paul Rivet et Chetsmír Loukotka, *Langues de L'Amérique du Sud et des Antilles*, in Meillet-Cohen, *Les langues du Monde*, 2a. edição, Paris, 1952. Págs. 1099-1160.

bitrariamente; do ponto de vista metodológico, entretanto, tais limites são muito importantes, uma vez que permitem melhor agrupamento linguístico dos idiomas. Sua importância se acentua quando somos obrigados a trabalhar sobretudo com material puramente lexical, como de ordinário se dá na América do Sul. Além disso, para melhor representar a diferenciação interna de cada grupo de línguas, torna-se necessário aplicar não apenas uma, mas várias unidades de classificação. Até agora, a estruturação da maioria dos grupos, tanto nos trabalhos de Loukotka e Rivet, como nos de Mason⁹, obedece a critérios que em grande parte são de ordem geográfica e não linguística. É verdade que Mason propõe algumas unidades maiores, os *phyla* “Macro-Tupí-Guaraní”, “Macro-Guaikurú”, etc., reconhecendo, entretanto, tratar-se em parte de meras suposições. Ao “tronco linguístico Tapúya-Jê” de Loukotka subordinam-se nove famílias linguísticas¹⁰; a este tronco linguístico corresponde o “Phylum Macro-Jê” de Mason. De maneira geral, porém, falta um critério adequado para distinguir determinados graus de parentesco linguístico. Também a diferença entre “línguas” e “dialetos” é pouco precisa, a ponto de às vezes dois idiomas estreitamente aparentados serem considerados duas “línguas”, enquanto outros, que não se assemelham tanto, são classificados como “dialetos”.

A fim de classificar as línguas Tupí, utilizamos a mesma tabela de gradação proposta por Swadesh no último Congresso de Americanistas¹¹. Embora a tabela se fundamente na glotocronologia, de maneira que cada grau corresponde a um período de tempo de desenvolvimento divergente das respectivas línguas, não nos propomos empreender qualquer datação, especialmente porque o material examinado ainda apresenta muitas lacunas. Trabalhamos apenas com as porcentagens de cognatos (isto é, vocábulos conservados em comum), que implicam, sem dúvida, distância temporal. Swadesh propôs os seguintes graus:

Família (family)	–	mais de 36% de cognatos
Tronco (stock)	–	de 36% a 12% de cognatos
Microphylum	–	de 14% a 4% de cognatos
Mesophylum	–	de 4% a 1% de cognatos
Macrophylum	–	menos de 1% de cognatos

9. J. Alden Mason, *The languages of South American Indians*. Handbook of South American Indians, vol. 6, págs. 157-317.

10. Chestmír Loukotka, *Klass. der süd. Sprachen*, págs. 2-6.

11. Morris Swadesh, *Towards a satisfactory genetic classification of Amerindian languages*, op. cit., pág. 1010.

Swadesh ainda sugeriu que sejam denominados dialetos os idiomas que apresentem um vocabulário com mais de 81% de cognatos.

Uma vez que tratamos apenas de um tronco linguístico, utilizamo-nos somente dos conceitos de dialeto, língua, família e tronco propostos por Swadesh. Acrescentamos, porém, o conceito de “subfamília” para designar um agrupamento interno dentro do âmbito de uma família que compreende línguas com um vocabulário de mais de 60% de cognatos.

Quando várias línguas apresentam várias semelhanças entre si, não é fácil estabelecer as relações mútuas apenas com o auxílio de um critério lexical como o da léxico-estatística. Isto se dá principalmente com relação aos dialetos, onde os aspectos fonológicos, morfológicos e outras características linguísticas também devem ser levados em conta. Este seria o caso dos idiomas incluídos em A.a.1. na classificação que se segue, e que ainda não foram examinados neste sentido.

Vejam agora a classificação do tronco linguístico Tupí (ingl. *Tupí stock*, alem. *Tupí-Sprachstamm*). Para muitas línguas deste tronco, e mesmo para a maioria delas, o material até agora conhecido é muito pobre; prevendo-se naturalmente um aperfeiçoamento futuro da classificação; acreditamos, porém, que as linhas fundamentais do esquema possam ser tidas como certas.

Denotamos as famílias com letras maiúsculas, as subfamílias por minúscula, as línguas por números e os dialetos por letras gregas. Alguns idiomas, cujos vocabulários publicados não puderam ser examinados, foram marcados com ponto de interrogação ou então nem tiveram menção (isto se refere principalmente aos dialetos).

Tronco Tupí

A. Tupí-Guaraní

- a. 1. Tupí-Guaraní: α . Tupí (i. Tupinambá ou Tupí Antigo, ii. Nheengatú ou Tupí Moderno); β . Guaraní (i. Guaraní Antigo, ii. Avanheên ou Guaraní Moderno); γ . Kaiwá (i. Apapokúva, ii. Mbüá); δ . Txiriguâno; ϵ . Tapieté; ζ . Izozó (Txané); η . Guarayú.
- a. 2. Tenetehára: α . Tembê; β . Gwajajára; γ . Urubú; δ . Manajé; ϵ . Turiwára; ζ . Anambé.
- a. 3. Oyampí: α . Oyampí; β . Emerillon.
- a. 4. Kawaib: α . Wiraféd; β . Pawaté; γ . Parintintín.
- a. 5. Apiaká.
- a. 6. Kamayurá.
- a. 7. Awetü.
- a. 8. Tapirapé.
- a. 9. Xetá (Aré).
- a. 10. Pausérna.

- a. 11. Kayabí (?).
- a. 12. Canoeiro (Abá).
- a. 13. Takunapé (?).
- b. 1. Kokáma: α . Kokáma; β . Kokamíla.
- b. 2. Omágua.
- c. Guayakí.
- d. Maué.
- e. 1. Mundurukú¹²
- e. 2. Kuruáya.
- f. Sirionó (?)

B. Yurúna

- a. 1. Yurúna.
- a. 2. Xipáya.
- b. Manitsawá.

C. Arikêm

- 1. Arikêm.
- 2. Karitiána.
- 3. Kabixiána (?).

D. Tuparí

- 1. Tuparí
- 2. Guarategaya (Koarátira, Gaurátira, Amniapé, Mekén, Kanoé).
- 3. Wayoró: α . Wayoró (Ayurú); β . Apitxum.
- 4. Makuráp.
- 5. Kepkiriwát.

E. Ramarâma

- 1. Ramarâma; α . Ramarâma; β . Ntogapüd.
- 2. Urukú.
- 3. Urumí.
- 4. Arára.

F. Mondé

- 1. Mondé: α . Mondé; β . Sanamaikán (Salamáin).
- 2. Digüt.
- 3. Aruá: α . Aruá; β . Aruáxi.

G. Puruborá

12. Talvez antes uma família ao lado da Tupí-Guaraní do que uma subfamília deste.

Para finalizar, seguem-se algumas observações que julgamos importantes:

a) Segundo o critério aqui adotado, a relação entre o Tupinambá ou Tupí Antigo e o Guaraní Antigo é a de “dialetos”, e dialetos muito próximos (90% de cognatos), mas não a de “línguas”. Consequentemente, não se justifica, do ponto de vista linguístico, uma divisão da família Tupí-Guaraní em dois grupos principais, como o querem algumas classificações, em que um deles é subordinado ao Tupinambá e o outro ao Guaraní.

b) Tembé, Guajajara e demais dialetos do mesmo grupo (*A. a. 2.*) não podem ser considerados como derivados diretamente do Tupinambá, embora tenham parentesco muito próximo com o Tupinambá e o Nheengatú.

c) Digno de notas é o fato de quase todas as famílias linguísticas do tronco Tupí até agora reconhecidas se concentrarem na região do Guaporé, isto é, do Alto Madeira, particularmente entre os rios Guaporé e Ji-paraná (ou Machado). As famílias Arikê, Kanoé, Mondé, Ramarâma e Puruborá encontram-se exclusivamente nessa região, e também a família Tupí-Guaraní acha-se ali bem representada (pelas línguas registradas em *A. a. 4.*). Este fato sugere que talvez o centro de difusão do Proto-Tupí deva ser procurado na área do Guaporé.

d) Muito interessante é também a curiosa difusão da família Tupí-Guaraní, cujos representantes, em especial da subfamília *a.*, se apartaram de tal maneira que encontramos, por exemplo, o Avanheén (Guaraní) no norte da Argentina e o Oyampí na Guiana Francesa, o Kokáma no Peru oriental e o Tupinambá no litoral brasileiro. Um exame mais detido dos componentes dessa família fornecerá uma série de explicações para as migrações dos respectivos povos.

e) Apesar de nos termos absterido de qualquer datação glotocronológica devido à natureza do material disponível, podemos avaliar a idade de todo o tronco linguístico Tupí-Guaraní em uns 5000 anos (11% de cognatos aproximadamente) e a da família Tupí-Guaraní em aproximadamente 2500 anos (cerca de 35% de cognatos).

Tradução de Arnaldo Hauptmann e Theka Hartmann

Nota:

Este trabalho foi publicado em alemão nos Proceedings of the 32nd International Congress of Americanists, Copenhagen, 1959; págs. 679-684. Agradecemos ao Prof. Aryon Dall'Igna Rodrigues a permissão de divulgá-lo em tradução portuguesa. O esquema de classificação aqui reproduzido não é o

do trabalho original, mas uma versão corrigida, que o autor apresentou em seu artigo “Classification of Tupí-Guaraní”, *International Journal of American Linguistics*, vol. 24, págs. 231-234; 1958. — Por motivos de ordem tipográfica adaptamos a grafia dos etnônimos à que se estabeleceu por convenção firmada pela maioria dos participantes da 1ª Reunião Brasileira de Antropologia (Rio de Janeiro, 1953) e cujas normas se encontram na *Revista de Antropologia*, vol 2, págs. 150-152; São Paulo, 1954.