

CLASSIFICAÇÃO DAS LÍNGUAS INDÍGENAS DO BRASIL (*)

J. Mattoso Câmara Jr.
Rio

1. AS CLASSIFICAÇÕES EXISTENTES

As línguas indígenas do Brasil, no quadro geral das línguas indígenas sul-americanas, têm sido objeto de várias classificações. Destacam-se, especialmente, como as mais modernas e conduzidas em linhas mais seguras, as que apresentam — J. Alden Mason (“The Languages of South American Indians”) no **Handbook of South American Indians**, Washington D. C., 1950, vol. 6; Wilhelm Schmidt (“Die Sprachen Sudamerikas”) em **Die Sprachfamilien und Sprachenkreise der Erde**, Heidelberg, 1926; Chestmir Loukotka (“Línguas Indígenas do Brasil”) em **Revista do Arquivo Municipal**, São Paulo, 1939, n.º LIV, que está ligada à classificação geral das línguas indígenas sul-americanas constante de **Les Langues du Monde**, Paris, 1952; 2.º ed., de Meillet-Cohen.

Tôdas essas classificações procuraram ser, em princípio, GENÉTICAS, isto é, dividir as línguas pelo seu parentesco, na base de uma hipotética origem comum, o que tem sido o critério essencial da ciência da linguagem desde que abandonou a tentativa de classificação tipológica esboçada por Schleicher e seus continuadores (línguas isolantes, aglutinantes e flexionais; ou, em parte noutro sentido, línguas analíticas, sintéticas e polissintéticas). A preferência se impôs, não só em virtude da tendência dominante, e até avassaladora, para as classificações genéticas em lingüística, mas também porque o critério tipológico rudimentar, então existente, não propiciava fundamentos aproveitáveis, uma vez que as

(*) Documento de trabalho n.º 2 apresentado à IV Reunião Brasileira de Antropologia, Curitiba, 15 a 18 de julho de 1959.

línguas indígenas americanas, em geral, eram aí consideradas indistintamente aglutinantes, ou incorporantes, ou polissintéticas.

O critério genético, ao contrário, permitia numerosos quadros e divisões, favorecendo uma ordenação metódica dentro da multiplicidade intrincada e confusa com que se deparavam os pesquisadores.

Foi êle assim de apreciável utilidade, mas (digamo-lo sem ambages) de utilidade meramente prática e sem maior valor científico.

A rigor, o princípio lingüístico genético vai aí de par com critérios puramente geográficos ou etnográficos, e, em si mesmo, se mostra reticencioso, apelando a cada passo para o conceito de “língua mista”, como ressaltai alhures.

Nem podia ser de outro modo, em vista dos dados precários e incertos de que em regra dispuseram os classificadores.

Com efeito, a pesquisa nas línguas indígenas sul-americanas tem sido conduzida, quase exclusivamente, nas velhas e obsoletas linhas de Adelung e Hervás, que no século XVIII se voltaram para a coleta lingüística objetiva, mas ainda sem um corpo de doutrina específica para se orientarem. Trata-se de meros levantamentos de vocabulários de algum interêsse etnográfico e onde, em regra, a grande atenção foi dada ao valor geral semântico dos itens e não à forma lingüística, quer do ponto de vista fonético, quer do ponto de vista gramatical. Aquilo que caracteriza realmente uma língua — o quadro de fonemas, os traços prosódicos, os tipos de morfemas, a sua estruturação nos vocábulos e a dêstes na frase — só indireta, confusa, fragmentária e não raro inseguramente se deixa depreender dessas listas lexicais.

2. A CLASSIFICAÇÃO GENÉTICA

É conveniente, portanto, confrontarmos de início essa situação e a natureza das classificações, daí decorrentes, com o que estritamente se entendeu por classificação genética na época em que se processava todo êsse trabalho classificatório das línguas indígenas sul-americanas.

A classificação genética é precìpuaemente o resultado de uma comparação gramatical, que exige o conhecimento fonético cabal

das línguas comparadas, o levantamento das suas formas gramaticais e uma perspectiva no tempo que permita estabelecer cadeias de formas na evolução do estado antigo para o atual. Com êsses dados é possível traçar, para cada língua, a linha evolutiva do seu sistema de sons, depreendendo-se o que se convencionou chamar “leis fonéticas”. O método foi especialmente fecundo no domínio das línguas indo-européias e semíticas, mas alhures mostrou-se sempre um tanto precário. Em referência às línguas americanas, um grande teorista da gramática comparativa se mostrou muito cético a respeito do “estabelecimento de gramáticas comparativas precisas e completas”: “as sondagens feitas até agora prometem pouco”, escrevia êle em 1924 (A. Meillet na “Introdução” da 1ª edição de **Les Langues du Monde**, Paris, 1924).

Hoje, no entanto, já não tem pròpriamente cabida um ceticismo dessa ordem.

A razão não está apenas num maior aprofundamento das pesquisas e no aperfeiçoamento das técnicas comparativas, senão também numa nova conceituação mais elástica do comparativismo lingüístico.

Em relação à pesquisa e à técnica, a etnologia lingüística norte-americana desenvolveu para as línguas da América do Norte um trabalho descritivo rigoroso e amplo, numa já longa experiência de cujos ensinamentos está começando a se beneficiar a América do Sul e muito recentemente o Brasil. É de esperar que com o correr do tempo tenhamos descrições plenas e satisfatórias das línguas indígenas ainda existentes, que deixarão a perder de vista os sumários e inadequados vocabulários com que quase exclusivamente trabalharam os autores das classificações até agora feitas. Por outro lado, os princípios da geografia lingüística, transpostos para o âmbito ameríndio, permitem compensar, até certo ponto, a falta de documentação temporal, que tem impedido estabelecer as cadeias evolutivas: as formas distintas cognatas podem ser agrupadas numa ordem de sucessão que dará, na base das línguas atuais, a perspectiva da evolução das formas.

Em relação à nova conceituação do comparativismo, é importante atentar que a relutância a operar primacialmente com dados

léxicos, por parte dos comparativistas, só confiantes nos dados fonéticos e morfológicos, cedeu lugar a uma atitude muito mais liberal. Pensa-se hoje que o acervo de palavras não-culturais, referentes a coisas e atividades básicas da vida social, apresenta um índice apreciável de manutenção, através do tempo, e só são substituídas de maneira muito paulatina, e até com certa velocidade regular, essas palavras. Isso conduz naturalmente ao levantamento de um novo tipo de vocabulário, que poderá ser utilizado com eficiência no trabalho comparativo e na classificação das línguas.

O fator da velocidade regular, na substituição dos itens, permite a mais uma avaliação temporal das línguas, ou seja, uma cronologia comparativa entre elas, ou “glotocronologia”.

Acresce a circunstância de que a lingüística já não se atém, exclusivamente, ao ideal de uma classificação genética perfeitamente comprovada e indubitável. Neste particular, os estudiosos das línguas indígenas brasileiras podem recorrer ao conceito do “parentesco provável”, que estabelece Sydney Lamb como um objetivo perfeitamente válido para a classificação lingüística (“Some Proposals for Linguistic Taxonomy”, em **Anthropological Linguistics**, 1-2, Bloomington Ind., 1959): “uma ERC” (i. e., Established Relationships Classification) “indica quais os parentescos que até um dado momento se conseguiu estabelecer, enquanto uma “PTC” (i. e., Probable Truth Classification) “representa o que de melhor se pôde deduzir sobre a verdade provável. E, enquanto a ERC exige certeza e em algumas de suas versões se baseia em tipos limitados de prova, a PTC só pede que se considerem as possibilidades e recorre a qualquer tipo de prova ao seu alcance” (p. 34). E mesmo dentro da ERC, ou classificação na base de parentesco firmado, se pode assumir uma atitude “liberal” (idem, p. 35) que nos desvencilhará do quadro comparativista estrito e rígido em que se colocava Meillet para duvidar da possibilidade das classificações genéticas nas línguas americanas, como citamos há pouco.

Na realidade, as classificações ora vigentes para as línguas indígenas da América do Norte assentam nesses critérios de comparativismo liberal e de possibilidades de parentesco. Ora, os seus resultados têm sido tão úteis como ordenação de conhecimentos lingüísticos e ponto de partida e orientação para o estudo, que ple-

namente se justifica procurarmos seguir o exemplo para as línguas indígenas do Brasil.

3. OUTROS CRITÉRIOS DE CLASSIFICAÇÃO

É preciso, entretanto, ressaltar que a classificação genética não é a única diretriz no problema classificatório das línguas. A ciência contemporânea já ultrapassou a atitude, ainda enfaticamente assumida por Meillet na “Introdução” a **Les Langues du Monde**, há pouco citada, segundo a qual a única classificação cientificamente válida, em matéria de línguas, é a que assenta no parentesco, ou seja, em famílias cujos membros se ligam por uma origem comum.

A lingüística hodierna aceita, ao contrário, como fêz recentemente J. Greenberg (**Essays in Linguistics**, New York, 1957, págs. 66-85), 3 espécies de classificação para as línguas, cada qual com as suas vantagens, aplicações e oportunidades distintas: a genética, a areal, a tipológica.

A classificação areal não é propriamente fundamentada no simples critério geográfico, que em si mesmo não pode interessar a lingüística. Fundamenta-a antes o princípio da difusão dos traços lingüísticos através de línguas variadas em contacto geográfico. Visa-se, portanto, a reunir em grupos as línguas contíguas que apresentam, em virtude da difusão, semelhanças de estrutura gramatical ou de formas vocabulares básicas.

A sua importância para a antropologia cultural é muito grande, como insiste Greenberg. Há uma evidente correlação entre as áreas lingüísticas, assim obtidas, e as áreas culturais, pois umas e outras decorrem de influências por contacto, e os critérios lingüísticos, que serviram para o estabelecimento de um grupo areal de línguas, pode ser utilizado pelo antropólogo, ao lado dos seus critérios não-lingüísticos, para determinar uma área cultural. “É evidente, portanto, que é a área lingüística, e não a família lingüística, que corresponde à área cultural, tanto nos processos históricos de formação como nas mais importantes características formais” (Greenberg, cit., pág. 68).

A adoção de classificações areais não exclui a de classifica-

ções genéticas. A rigor, é até conveniente que existam estas últimas para se poder determinar com mais precisão os traços em comum provenientes da difusão, os quais são em princípio a justificação das línguas (**Language, an introduction to the study of**

Por outro lado, o levantamento de traços lingüísticos comuns, independentes do parentesco lingüístico, leva-nos naturalmente ao critério tipológico para a classificação das línguas, abandonado até bem pouco em virtude da avassaladora importância dada ao critério genético.

A diferença básica e essencial entre uma classificação por áreas e uma classificação por traços lingüísticos típicos é que, nesta última, não entra o fator difusão. As classificações tipológicas partem do pressuposto de que há tipos gerais de estrutura lingüística e que por êles é possível separar ou reunir as línguas.

4. AS CLASSIFICAÇÕES TIPOLÓGICAS

Edward Sapir foi o primeiro teorista a oferecer à lingüística um quadro geral de traços tipológicos para base de uma classificação das línguas (**Language, an introduction to the study of speech**, New York, 1921). É um quadro extremamente complexo, todavia, e que exige um enorme conhecimento da estrutura de cada língua, em amplitude e profundidade; seria utópico procurar aplicá-lo às línguas indígenas do Brasil, onde só agora se começa a dirigir a pesquisa em linhas verdadeiramente rigorosas.

É muito mais operacional o critério que apresenta Greenberg, como reformulação do quadro de Sapir (**Methods and Perspective in Anthropology**, Minneapolis, 1954, pág. 192 ss.). A sua maior vantagem de aplicação, para nós, é que aí se parte de um texto limitado e não de um estudo geral e sistemático da língua; é, nos termos do autor, uma classificação de base “pragmática” e não “sistêmica” (cf. **Revista de Antropologia**, São Paulo, 1958, págs. 5,7). Se nos dispusermos, portanto, a utilizar as fórmulas de Greenberg, bastará programar uma coleta de um texto limitado, para cada língua a classificar, o que está perfeitamente dentro das nossas possibilidades atuais.

Teremos então uma classificação de bases morfológicas, as-

sentando essencialmente na estrutura do vocábulo e no comportamento dos afixos em relação à raiz.

Compreende-se, não obstante, que há outras linhas tipológicas possíveis de adoção.

A estrutura fônica, por exemplo, é um dado muito importante, que pode servir de roteiro para a nossa classificação tipológica. Pode-se levar em conta a natureza do acento, a estrutura da sílaba, os fonemas preferenciais no início ou no fim dos vocábulos, as classes de consoantes que funcionam na língua ou seu vocalismo. É assim que o português se distingue tipologicamente do italiano pela ausência de consoantes geminadas, do francês pela presença de paroxítonos e proparoxítonos e de /r/ e /s/ finais, do espanhol por um vocalismo distinto conforme a posição da vogal é átona ou tônica.

O próprio léxico pode ensejar uma classificação tipológica de natureza vocabular, com implicações antropológicas interessantes, porque aí se revelarão associações semânticas de interesse cultural em certos campos semânticos como o dos têrmos de parentesco, o dos fenômenos atmosféricos, os de certas atividades sociais.

Na verdade, pode-se conduzir uma classificação tipológica, dentro do que Charles Hockett chama “o núcleo gramatical” das línguas (“the grammatical core”), focalizando-se em dêsse 4 aspectos: 1) o sistema de partes do discurso (p. ex. — verbos distintos de nomes, substantivos distintos de nomes, substantivos distintos de adjetivos, espécies de pronomes etc.), 2) as categorias gramaticais (espécie de gênero, tempo versus aspecto verbal, espécies de modos verbais etc.), 3) os “functores” ou elementos gramaticais, 4) os tipos frasais e as construções léxicas (cf. C. Hockett, **A Course in Modern Linguistics**, New York, 1958, pág. 265).

A êsse respeito, é interessante lembrar o critério de Wilhelm Schmidt no seu livro clássico sôbre **Die Sprachfamilien und Sprachenkreise der Erde** (Heidelberg, 1926). Para estabelecer os seus “círculos lingüísticos”, que são no fundo agrupamentos tipológicos, embora o autor lhes empreste um sentido genético

(admitindo a passagem de um círculo para outro por transformação no tempo), focalizam-se aí certos traços fonéticos, mórficos e sintéticos em número restrito mas de valor expressivo: os tipos de vogais, os tipos de consoantes, a natureza fonética do início e do fim do vocábulo, a estrutura dos numerais e a base de numeração, o número gramatical nos pronomes e nomes, o gênero gramatical, a posição do termo determinante em face do determinado (ob. cit. p. 272 ss.). É verdade que Schmidt, dentro das idéias, ou antes dos preconceitos, da sua escola histórico-cultural, procura associar por um elo de essência os “círculos lingüísticos”, assim obtidos, com os “círculos culturais”, “para ver se e até que ponto os grandes agrupamentos lingüísticos coincidem, externamente, e, internamente, se integram com os etnológicos” (ob. cit. p. 271).

O que hoje se nos apresenta como pueril, nessa tentativa, não nos deve impedir, entretanto, de reconhecer o cabimento dos traços lingüísticos assim levantados nem de aproveitá-los como um ponto de partida, pelo menos, num trabalho de classificação tipológica das nossas línguas indígenas.

5. AS CLASSIFICAÇÕES DOS SUBGRUPOS.

Ao lado da classificação dos grandes grupos lingüísticos, há o problema dos subgrupos, que sob certos aspectos é bem mais árduo do que a apreensão das famílias. Até no âmbito indo-europeu nota-se essa dificuldade, como testemunham as controvérsias a respeito da relação do itálico com o celta ou mesmo do osco-umbro com o latim.

Nas classificações das nossas línguas indígenas, de início aqui citadas, estabeleceram-se, de maneira geral, os subgrupos por um critério quase exclusivamente geográfico. É o que logo ressalta do exame da classificação de Loukotka e o que reconhecem explicitamente tanto Mason (p. 289) como Schmidt (p. 236). Quando muito, apela-se também para um critério etnológico ou ainda político, como declara Mason (ibidem).

É, entretanto, plenamente possível proceder à classificação dos subgrupos por um critério puramente lingüístico.

Se o método comparativo permitiu o levantamento de proto-formas teóricas, de que as formas comprovadas das diversas línguas do bloco provieram, tem-se para logo uma possibilidade de agrupamento na base das mudanças fonéticas preferenciais que em cada língua se manifestaram. Outro critério é a distribuição dos afixos, que se apresenta diferente de língua a língua, associando-se estas pelas coincidências e opondo-se pelas discordâncias em tal distribuição. Assim, no subgrupo Akuê, da família Jê, aparece o prefixo possessivo **da-** e secundariamente **-su-**, que não figuram alhures, pelo menos com o mesmo relêvo e nitidez de funcionamento.

Por outro lado, o estudo comparativo em função de um vocabulário básico, que fundamenta a pesquisa “glotocronológica”, de início aqui referida, também pode propiciar métodos satisfatórios para sub-agrupar as línguas de uma família. Neste particular, merece atenção o artigo recente de H. A. Gleason, Jr (“Counting and Calculating for Historical Reconstruction”) em **Anthropological Linguistics**, 1-2, Bloomington Ind., 1959, sôbre a depreensão da maior ou menor associação genética de duas ou mais línguas, dentro uma família lingüística, partindo-se de uma contagem de semelhanças e diferenças entre itens vocabulares constantes de uma lista padrão: quanto mais restrita fôr a distribuição de um conjunto de termos-cognatos, tanto mais probantes serão êles do parentesco imediato das línguas que os possuem, dentro do parentesco geral de tôdas as famílias do grupo (cf. ps. 23 e 28).

Outra solução, para o problema do sub-agrupamento, é de consistir de estabelecer subgrupos dialetais distintos, mas ao contrário concatenar tôdas as línguas da família em “cadeia”, admitindo-se o que Morris Swadesh qualifica de “fluidez dialetal” (cf. “The Mesh Principle in Comparative Linguistics”, em **Anthropological Linguistics**, 1-2, Bloomington Ind., 1959).

Há finalmente o recurso, fora da classificação genética, aos critérios tipológicos e areais para sub-agrupar as línguas de um bloco, depreendido quer como família lingüística, quer como grupo tipológico, quer como conjunto areal. É, por exemplo, um

princípio areal que nos leva a dividir as línguas românicas em línguas do centro (italiano, rético, sardo), línguas de oeste (francês, provençal, castelhano, português, catalão) e línguas de leste (romeno e suas variantes, além do dálmata extinto). E, ainda nas línguas românicas, temos um princípio de tipologia das vogais, que opõe o italiano, o castelhano e o português, de um lado, com o francês, de outro lado (vogais “mistas” /ü/, /ö/), ou com o romeno (vogal /a/), ou um princípio de tipologia morfológica na maneira de se expressar o plural (cast. port. **rosas**, fr. **roses**), it. **rose**); e assim por diante.

6. CONCLUSÃO.

É tempo de finalizar êste Relatório, que apenas visou a sugerir temas para o debate que vamos aqui realizar.

Parece-me conveniente tornar explícitos êsses temas como um roteiro para o debate. Tentarei, pois, formulá-los em itens distintos.

1) Que espécie de classificação genética se pode realizar nas línguas indígenas brasileiras? Convém aí uma atitude liberal e mesmo o objetivo de um “parentesco provável”, em vez de firmemente estabelecido?

2) Até que ponto podemos nos beneficiar das técnicas da glotocronologia para a execução de uma classificação genética das nossas línguas indígenas?

3) Há vantagem em se tentar uma classificação areal das línguas indígenas brasileiras, ao lado dos esforços para a classificação genética?

4) Até que ponto uma classificação tipológica pode suprir a falta de uma classificação genética?

5) Quais os critérios tipológicos que se devem preferir? Convém contentarmo-nos com uma classificação “pragmática”?

6) Deve-se optar por uma classificação tipológica em vez de classificação genética, ou deve uma servir de complemento da outra?

7) Como se deve proceder à classificação dos subgrupos por um critério essencialmente lingüístico? Há vantagem relevante em se apelar para o método da estatística vocabular de Gleason? Até que ponto se pode postular uma dialeção em cadeia?

8) Quais as espécies de pesquisas lingüísticas que se tornam mais importantes para facilitar e apressar o trabalho de classificação das línguas indígenas brasileiras?

Curitiba, julho de 1959.