DO ESTUDO TIPOLÓGICO EM LISTAS DE VOCÁBULOS INDÍGENAS BRASILEIROS

J. Mattoso Camara Jr.
(Div. de Antropologia, Setor Lingüístico, Museu Nacional, Rio de Janeiro)

Desde os primórdios da lingüística, na primeira metade do século XIX, foi um dos objetivos da nova ciência uma classificação sistemática e racional das línguas do mundo. À classificação puramente geográfica, e sem maior sentido científico, substituíram-se outros critérios — os raciais de Friedrich Müller [1] e Finck [2], e o tipológico dos irmãos Schlegel [3] e depois de Schleicher [4]. Em breve, entretanto, a lingüística firmou-se numa solução que pareceu definitiva: a classificação genética, ou seja, a distribuição das línguas na base de uma origem comum.

E' óbvio que as classificações de fundo racial tinham de ser superadas com a progressão dos estudos lingüísticos, os quais evidenciaram na língua uma instituição puramente cultural, sem nenhuma dependência do tipo físico dos homens que as falam. O critério tipológico merecia, porém, a atenção da ciência e foi um mal ter sido minimizado e afinal abandonado. A maneira defeituosa por que foi adotado de início, partindo da estrutura isolante, aglutinante ou flexional do vocábulo, exclusivamente, e sem uma delimitação rigorosa entre aglutinação e flexão, poderia ter sido pouco a pouco melhorada. Mas sucedeu que o sentido tipológico, na medida em que a classificação tipológica deve servir para caracterizar descritivamente as línguas, foi desvirtuado pela apresentação das três estruturas vocabulares como uma súmula da evolução da linguagem humana do simples para o complexo. A ênfase emprestada a êste ponto de vista, obscurecendo todos os outros aspectos da classificação, prejudicou-a irremediàvelmente, quando a hipótese evolutiva, de que partia, se tornou insustentável.

Para fins de lingüística histórica é, em verdade, muito mais operacional a classificação genética. Ela se aplicou, com os brilhantes resultados que todos conhecemos, às línguas da Europa e da Ásia ditas, por isso, indo-européias. E tornou-se um propósito declarado da lingüística a extensão do método às demais línguas do mundo, a fim de reuni-las em grandes classes genéticas ou famílias.

* *
*

Ora, a técnica para a classificação genética exige certos dados preliminares, de que dispuseram os indo-europeístas, mas que nem sempre se encontram em outros âmbitos lingüísticos. Por isso, A. Meillet advertia que é desconhecer os princípios essenciais da classificação genética querer utilizá-la em quaisquer condições [5].

A base da comparação lingüística, de que emerge a classificação genética, está antes de tudo nos elementos gramaticais, que estruturam os vocábulos, e nas correspondências rigorosas entre consoantes e vogais; além disso, faz-se mister uma documentação de estados lingüísticos pretéritos que permitam estabelecer certas cadeias evolutivas. E' o que se pode resumir em três itens: 1) segmentos mórficos, de valor gramatical definido, abundantes e nítidos; 2) correspondências fonéticas, ou leis fonéticas, mostrando a evolução divergente, conforme a língua, de vogais e consoantes a partir de um som uno originário; 3) formas intermediárias, documentadas, permitindo associar duas ou mais atuais, que aparecem como desconexas.

Compreende-se, assim, como foi precária a utilização do método de classificação genética a um setor como o das línguas índias sul-americanas, onde faltam essas três espécies de dados. A pesquisa lingüística limitou-se aí, até recentemente, a coletas de palavras que particularmente consultavam o interêsse etnológico. Em tais listas não há preocupações de análise mórfica, que é indispensável para se conseguir destacar os elementos gramaticais aludidos; apenas se individualizam, e nem sempre explìcitamente, certas partículas aglutinadas como os possessivos nos nomes de parentesco e de partes do corpo. Também não há, muitas vêzes, a preocupação de rigor fonético: listas levantadas por Castelnau [6], Martius [7] e vários outros pesquisadores, para as línguas indígenas do Brasil, adotam a grafia francesa, alemã ou portuguêsa no registro dos vocábulos indígenas, procurando adaptar os sons nativos aos fonemas definidos que as letras latinas em cada uma dessas grafias representam. Quando, ao contrário, houve maior rigor de transcrição e o pesquisador usou um sistema gráfico especial, o registro foi muito impressionístico e subjetivo para uma cabal dedução de leis fonéticas. E' quase inútil frisar a inexistência de documentação de fases lingüísticas pretéritas para relacionar formas que no seu aspecto atual estão profundamente dissociadas.

Bem indica a precariedade de aplicação da classificação genética às línguas indígenas sul-americanas o fato de que até hoje não se tentou sèriamente uma reconstituição da protolíngua de cada uma das famílias encontradas. Ora, a reconstrução teórica da estrutura lingüística originária é essencial, nos primeiros estágios da pesquisa, para o estabelecimento de um ponto de partida concreto que possibilite ampliar a família lingüística, levando-a a abarcar línguas cuja evolução só se esclarece quando se remonta à situação primitiva.

Daí as classificações genéticas que se têm obtido para as línguas indígenas sul-americanas, e em particular para as línguas indígenas brasileiras, assentarem apenas na coincidência ou semelhança de forma de um dado conjunto de itens de vocabulário, o que é uma prova muito precária de associação genética. Por isso, cada família lingüística, nìtidamente firmada, é apenas, na realidade, uma única língua **lato sensu** com suas múltiplas variedades dialetais. Quando as diferenças se extremam um pouco mais, a interpretação da língua como pertencente a esta ou aquela família se torna para logo incerta e acompanhada de ressalvas, como faz Chestmir Loukotka ("família Kamakan, com intrusão de Jê", "Kaingang, com os vestígios de Aruak", "língua mesclada com Kaingang", etc.) [8].

A conseqüência, já assinalada por Mason [9], é o aumento, antes que a redução, do número de famílias lingüísticas nas classificações mais recentes, o que é paradoxal, dentro do espírito teórico da classificação genética, que visa justamente relacionar, cada vez mais, línguas aparentemente desconexas, simplificando a complexidade dos grupos na base de um parentesco que cada vez mais se amplia.

Só novas pesquisas, procedidas noutros moldes e com outra orientação técnica, poderão futuramente proporcionar melhores condições para a comparação lingüística e permitir uma classificação genética em linhas mais ortodoxas.

* *
*

A realidade atual é que se pediu às listas vocabulares, de que quase exclusivamente dispomos (pois os textos seguidos, as análises mórficas e os esboços gramaticais são em regra raros, falhos e até inadequados), um serviço muito acima do que elas poderiam fornecer, quando sôbre elas se erigiram sistemas de classificação genética.

Nesta conjuntura, é o caso de nos perguntarmos se a classificação genética é a única que deva ser colimada.

A importância dada atualmente em lingüística aos estudos descritivos, independentes da lingüística evolutiva ou histórica, pôs de novo em foco o critério da classificação tipológica como meio de se chegar a uma taxeonomia lingüística [10]. Sapir, como se sabe, tentou uma revisão da classificação tipológica de Schleicher, tirando-lhe as insustentáveis implicações evolutivas e dando maior nitidez e hierarquização aos conceitos fundamentais classificatórios [11]. O seu esfôrço não foi, a princípio, devidamente apreciado pelos descritivistas norte-americanos que não queriam ascender aos problemas taxeonômicos gerais [12]; mas recentemente Greenberg pôs no devido relêvo o trabalho sapiriano, e, inspirado nêle, procurou chegar a uma classificação tipológica, — de natureza mórfica — precisa, cômoda e fàcilmente manipulável [13]. O mesmo Greenberg, num

artigo teórico posterior, debateu o tema da classificação tipológica, apontando as múltiplas diretrizes — fonéticas, mórficas, sintáticas, semânticas — aí possíveis [14].

Foi também, em última análise, um critério tipológico que orientou W. Schmidt na aplicação dos seus "círculos lingüísticos" **(Sprachenkreise)** às línguas sul-americanas [15]. O seu trabalho, neste sentido, foi obscurecido por três preocupações concorrentes: a manutenção de uma classificação genética, cujas famílias recebem secundàriamente a caracterização de um "círculo" de traços tipológicos; a hipótese difusionista de que êsses traços passaram de um grupo inicial a outros, sustentada com vistas para a tese do monogenismo da linguagem; a correlação entre os traços lingüísticos e os traços culturais que constituem os "círculos culturais" **(Kulturkreise)** da sua escola antropológica.

Não obstante, para fins puramente descritivos, os traços lingüísticos — fonéticos, mórficos, sintáticos — por êle escolhidos são expressivos para a classificação tipológica de uma língua.

Nestas novas diretrizes é plenamente possível utilizar as listas vocabulares de que dispomos, para delas obter dados tipológicos precisos. Um levantamento geral dêsses dados poderá, por sua vez, fundamentar, em seguida, uma classificação tipológica das línguas indígenas sul-americanas , e particularmente do Brasil.

Aqui, limitar-me-ei a enumerar os dados dessa espécie que é capaz de nos fornecer o exame das listas vocabulares, mesmo as menos satisfatórias em referência ao registro fonético e à análise da composição dos vocábulos colhidos. Valer-me-ei na exemplificação, por isso, da lista de Martius para o Akroá, a qual na sua imperfeição é o único documento de que dispomos sôbre essa língua indígena desaparecida [16].

* *
*

No âmbito fonético, é importante — como ressaltou W. Schmidt — apurar a existência, ou não, das vogais que êle chama impròpriamente "anormais" e em fonética se costuma designar como "mistas" (vogais anteriores, ou palatais, arredondadas, isto é — emitidas com arredondamento dos lábios, à maneira do **u** francês, e vogais posteriores, ou velares, não arredondadas, isto é, emitidas com distensão dos lábios, à maneira do y, /ï/, "gutural" tupi). Os coletores de vocabulários assinalam, em regra, a presença dessas vogais, quando elas existem, embora nem sempre distingam a série palatal e a velar. A Martius, cuja língua nativa, o alemão, distingue em seu sistema vocálico uma série de vogais "mistas", anteriores arredondadas, não escapou a presença delas no Akroá, onde podemos registrar /ö/ "misto", oposto a /o/, e /ü/ "misto", oposto a /u/:

"farina — **cuüt** — **patschú**", "nox — **macraüi**" (com clara distinção entre /ü/ e /u/ e /i/), "frigidus, a, um — **itöde(i)**", etc.

E' não menos importante depreender a existência, ou não, de um /a/ francamente palatal, em face do /a/ de articulação central, porque com essas duas espécies de /a/ o sistema de vogais adquire a configuração típica que Trubetzkoy chamou "quadrangular" **(Vierecksystem)** [17]. Nas listas dos coletores alemães o /a/ palatal vem assinalado por **ä**, distinto de **a** e **e**, e consta da lista Akroá de Martius: "eo, ire — **ktoroguanä**", em face de — "edo, ere — **wassattá**", "os, ossis — **thahické**", por exemplo.

Para as consoantes, são dados de monta saber, pelo menos, se há na língua — 1) um contraste entre oclusivas surdas e oclusivas sonoras (cuja indistinção W. Schmidt assinalou em várias línguas sul-americanas) [18]; 2) consoantes geminadas, (as quais, quando aparecem, impressionam logo o coletor, fazendo-o dobrar a letra consoante); 3) a aspiração, e se é apenas inicial, ou também medial, ou ainda final. Com efeito, êsses três aspectos do consonantismo determinam a configuração tipológica do sistema de consoantes.

Nas lista Akroá de Martius, não é evidentemente possível estabelecer oposições bimembres por um método de comutação ortodoxo (como port. **pote, bote** etc.) para provar a existência de consoantes sonoras como fonemas. Entretanto, o registro ora de letra de consoante sonora, ora de letra de consoante surda, em qualquer posição, parece indicar a presença dêsses dois tipos de fonemas, destacando o Akroá da linha geral das línguas Jê, que não raro se enquadram na observação acima citada de W. Schmidt: "calceus — **baracutschá**", "longus, a, um — **packüséde**"; "filia — **dewakonó**", "maritus — **tamarorhä**"; "asso, are — **iguakrorú**", "dens — **aiquá**". Quanto à geminação consonântica, pode-se deduzi-la em face de diferenças de registro como: "brevis, e — **uttudü**", "frigidus, a, um — **itöde(i)**"; "sagitta — **tikkite**", "calidus, a, um — **rowacroké**; "puella — **tabba**", "soror — **aitiba**". Finalmente, a aspiração aparece em posição inicial ("coelum — **höuate(i)**"), final ("capillus — **assaih**"), medial intervocálica ("coquo, ere — **pisaühul(r)i**"), pós-vocálica ("volo, velle — **ihtsche**"), depois de /r/ ("maritus — **tamarorhä**").

Os tipos normais de sílaba, sôbre cuja importância é inútil insistir para a estrutura lingüística, podem ser deduzidos através de uma análise dos vocábulos, adotando-se o critério da silabação mais natural. Ter-se-á então em Akroá sílabas simples (V) e compostas (C V, C V C), bem como os ditongos (VV); ex.: "asso, are — **iguakrorú**" (i-); "infans — **aikutä**" (—ku—, —tá—); "sol — **putdöti**" (put-), "venter — **aintúbdü** (—tub—). Quanto ao tipo CCV, é preciso pôr de parte evidentemente os grupos de letras **tsch, tz, ng**, que devem corresponder a consoantes simples (os dois primeiros a africadas e o terceiro a uma nasal velar, assim representadas em alemão), mas o grupo **muta cum liquida** é indubitável (haja vista o

monossílabo: "domus — **kli**") e mesmo **muta cum muta** ("eo, ire — **ktoro guanä**"). Por outro lado, há indícios esporádicos de um tipo silábico CVCC, como em: "tonitrus — **thauawabscádi**" (—**wabs**—). Note-se, finalmente, o indício da existência de /r₀/ silábico, ou em função vocálica, no vocábulo "calcaneus — **aipāracrtade**, cuja silabação deve ser **ai** — **pa** — **ra** — **cr** — **ta** — **de**, com uma sílaba /kr₀/.

Não tem menor significação tipológica o aspecto normal do início e do fim de vocábulo, conforme é vocálico ou consonantal. Sendo o alemão uma língua eminentemente de final consonântico, é expressivo que Martius tenha sempre registrado uma vogal final; êle a põe entre parênteses em "multus, a, um — **uarorütet(i)**", sugerindo um som vocálico reduzido mas suficiente para não se interpretar o segundo **t** como final. Ao contrário, se separarmos as partículas de posse, prefixadas a muitos nomes, e de natureza vocálica, observaremos uma grande predominância do início vocabular consonantal.

A posição normal do acento tônico é outro índice tipológico, que em regra os coletores registram. No Akroá, há acentuação aguda, em regra, e, esporàdicamente, grave.

Já a distinção entre vogal longa e vogal breve, à maneira latina, por exemplo, não parece ter funcionado em Akroá. Martius, só uma vez usa sinais de quantidade breve e longa, respectivamente sôbre a vogal **a** de — "calcaneus — **aipāracrtade**" e "(digitus) pedis — **aipărăiki**", e tudo indica que se trata de um mesmo elemento mórfico radical **para** —, cujas vogais se abreviaram e alongaram mecânicamente na composição (talvez, aliás, o alongamento em **aipāracrtade** seja compensatório em face da silaba seguinte reduzida a — **cr** — /kr₀/).

No âmbito mórfico, é possível em regra, nas listas vocabulares de que dispomos, um estudo do sistema de pronomes pessoais **lato sensu**, para deduzir se as partículas usadas como sujeito de vocábulos verbais são as mesmas, ou não, que as partículas possessivas afixadas a nomes de parentesco ou de partes do corpo, pois isso evidencia, dentro da língua, a distinção nítida ou apagada, entre os verbos e os nomes. Também importa verificar se umas e outras são prefixadas ou sufixadas, como indicação das condições sintáticas gerais para a posição de um determinante em relação ao seu determinado, o que W. Schmidt considerou, com razão, sintàticamente típico [19].

Em Akroá, a partícula pronominal é prefixada (ao contrário, por exemplo, do Torá) [20], e o pronome pessoal absoluto (cf. "ego — **wa**") é que normalmente figura como sujeito verbal: "edo, ere — **wasattá**"; "oro, are — **wameremäúliohö**". Deve haver, não obstante, verbos de natureza nominal, que recebem a partícula possessiva para nomes, como sugere um item: "morior, i — **aicktol(o)ö**"[21].

Finalmente, é possível explorar certos aspectos de relação semântica para certas áreas vocabulares, como as partes do corpo, o parentesco, as côres, os nomes de números, no sentido de verificar homonímias, derivações e heteronímias, que podem ser muito expressivas, tanto do ponto de vista lingüístico como cultural.

Em Akroá temos os seguintes casos, neste particular: homonímia entre "tio materno" e "pai" ("avunculus — **jungamá**"; "pater — **injungama**"); homonímia para "azul" e "negro" ("caeruleus — **schikutzacráng**"; "niger — **schikutzacráng**); heteronímia para "sol", "lua" e "estrêla", que opõe o Akroá a outras línguas indígenas brasileiras, sendo que a palavra para "sol" é um composto do elemento **put** ("sol — **putdöti**"), que, em outras línguas classificadas como Jê, aparece isolado no mesmo sentido. Infelizmente, a ausência dos nomes de números, na lista Akroá de Martius, não permite deduzir o sistema numeral pela natureza simples ou composta dos nomes de número a partir de 2, interpretando-se a forma composta como indicação da base numeral ali contida (assim, a base dual no Apinayé se evidencia na lista de Snethlage: "1 — **puti**", "2 — **adkrute**", ao lado de "3 — **adkrunati**", "4 — **adkrunadkru**", "5 — **adkrunadkruputi**) [22]. Da mesma sorte, a associação mórfica entre "5" e "mão" ou "pé" sugeriria uma base numeral quinária.

* *
*

Vê-se, assim, como o aproveitamento das listas vocabulares, nas línguas indígenas sul-americanas, não só permite utilizar essas fontes de maneira mais consentânea com o que elas nos podem oferecer, mas também prepara elementos para uma classificação tipológica dessas línguas.

Foi uma exploração neste sentido que aqui se quis sugerir.

Não se trata de abandonar os esforços para uma classificação genética, mas apenas de aguardar para isso um material lingüístico mais adequado do que as precárias listas vocabulares, com que essencialmente se tem até aqui operado, e de aproveitar estas últimas, por sua vez, para as finalidades de uma classificação tipológica.

E com isto também se estará indiretamente trabalhando para a classificação genética, no pressuposto de que uma família lingüística, por mais diferenciadas que se apresentem as línguas componentes, sempre traz em si uma vaga marca tipológica, a qual a põe em contraste com as famílias completamente distintas.

NOTAS

1) Em "**Grundriss der Sprachwissenschaft**" (1876).
2) Em "**Die Haupttypen des Sprachbaues**" (1910).

3) Em "**Über die Sprache und Weisheit der Inder**" (Fr. Schlegel, 1908) e "**Observations sur la langue et la littérature provençale**" (A. Schlegel, 1918).

4) Em "**Compendium der vergleichenden Grammatik der indogermanischen Sprachen**" (1861).

5) "**Introduction**" em: Meillet — Cohen, **Les Langues du Monde**, Paris, 1924, p. 9. Na 2a. edição (1952), essa Introdução de Meillet, tôda dedicada à classificação genética, foi substituída por outra, da responsabilidade de M. Cohen, onde se insiste na importância da classificação tipológica, ao lado da genética (p. XXXIII ss.).

6) "**Expédition dans les parties centrales de l'Amérique du Sud**", tome V, Paris, 1851.

7) **Wörtersammlung brasilianischer Sprachen**, Leipzig, 1867.

8) Chestmir Loukotka, "Línguas Indígenas do Brasil". **Revista do Arquivo Municipal**, LIV, São Paulo, 1939, pp. 150, 151.

9) "The Languages of South American Indians". **Handbook of South American Indians**, vol. 6, Washington, 1950, p. 165.

10) R. Jakobson, "Typological Studies and their Contribution to Historical Comparative Linguistics". **Reports for the Eighth International Congress of Linguists**, Oslo, 1957.

11) Em "**Language, introduction to the study of speech**", New York, 1921.

12) R. Jakobson, cit., p. 4.

13) "A Quantitative Approach to the Norphological Typology of Language". **Method and Perspective in Anthropology, papers in honor of Wilson D. Wallis**, Minneapolis, 1954.

14) "The Nature and Use of Linguistic Typologies". **IJAL**, Bloomington, Ind., 1957, vol. 23, n. 2. Tr. port. em **Revista de Antropologia**, 1958, vol. VI, n. 1.

15) **Die Sprachfamilien und Sprachenkreise der Erde**, Heidelberg, 1926, p. 209 ss.

16) Martius, cit., pp. 145-146.

17) "Zur allgemeinen Theorie der phonologischen Vokalsysteme". **TCLP**, I, Prag, 1929, p. 88.

18) W. Schmidt, cit., p. 285.

19) W. Schmidt, cit., p. 381 ss.

20) Curt Nimuendajú, "As Tribos do Alto Madeira". **Journal de la Société des Américanistes**, XVII, Paris, 1925, p. 148 ss.

21) A indistinção entre /l/ e /r/ intervocálicos, assim assinalada por Martius, contrasta na lista com a nítida diferenciação entre as duas líquidas como segundo elemento de um grupo consonântico.

22) Snethlage, "Unter nordostbrasilianischen Indianern". **Zeitschrift für Ethnologie**, vol. 62, Berlin, 1931, p. 196.