



Protótipos e Categorias na Classificação das Línguas Africanas: Uma Leitura Cognitivista da Proposta de Joseph Greenberg

Prototypes and Categories in the Classification of African Languages: A Cognitive Reading of Joseph Greenberg's Proposal

Rosana Menezes de Barros

Instituto Federal de Educação, Ciência e Tecnologia do Amazonas (IFAM).

Resumo: Este estudo investiga o diálogo entre a Linguística Cognitiva — em especial a teoria dos protótipos e a noção de categorização — e a classificação genética das línguas africanas proposta por Joseph Greenberg (1963). Parte-se do pressuposto de que toda taxonomia linguística constitui, antes de tudo, um ato cognitivo de categorização, e de que os efeitos prototípicos descritos por Rosch (1978) e sistematizados por Lakoff (1987) oferecem um instrumento produtivo para compreender como falantes e estudiosos representam mentalmente as grandes famílias linguísticas do continente. Tomando como objeto a macrofamília Níger-Congo — a mais numerosa da África —, o estudo discute de que modo línguas como o suaíli podem ser lidas como membros prototípicos da categoria, ao redor dos quais se organizam membros periféricos segundo um princípio de gradiência e de semelhança de família. Problematisa-se, ainda, a tensão teórica entre a natureza categórica do parentesco genético e o caráter gradiente da categorização prototípica, argumentando-se que a articulação entre os dois campos é de ordem heurística e representacional, e não genealógica. Conclui-se que a perspectiva cognitivista enriquece a leitura da obra de Greenberg e contribui para a valorização da diversidade linguística africana.

Palavras-chave: linguística cognitiva; protótipos; categorização; línguas africanas; Joseph Greenberg.

Abstract: This paper investigates the dialogue between Cognitive Linguistics — particularly prototype theory and the notion of categorization — and the genetic classification of African languages proposed by Joseph Greenberg (1963). It assumes that every linguistic taxonomy is, above all, a cognitive act of categorization, and that the prototype effects described by Rosch (1978) and systematized by Lakoff (1987) provide a productive tool for understanding how speakers and scholars mentally represent the continent's major language families. Taking the Niger-Congo macrofamily — the most populous in Africa — as its object, the study discusses how languages such as Swahili may be read as prototypical members of the category, around which peripheral members are organized according to a principle of gradience and family resemblance. It further problematizes the theoretical tension between the categorical nature of genetic kinship and the gradient character of prototypical categorization, arguing that the link between the two fields is heuristic and representational rather than genealogical. The paper concludes that the cognitive perspective enriches the reading of Greenberg's work and contributes to valuing African linguistic diversity.

Keywords: cognitive linguistics; prototypes; categorization; African languages; Joseph Greenberg.

INTRODUÇÃO

A África abriga uma das mais notáveis concentrações de diversidade linguística do planeta. Estima-se que no continente sejam faladas cerca de duas mil línguas — aproximadamente um terço de todos os idiomas vivos do mundo (Simons; Fennig, 2023). Diante dessa profusão, a tarefa de organizar, comparar e classificar tais línguas constituiu, ao longo dos séculos, um desafio que mobilizou viajantes, missionários e, mais tarde, linguistas profissionais. Foi apenas na metade do século XX, contudo, que um quadro classificatório coerente e amplamente aceito veio a emergir, sobretudo a partir dos trabalhos pioneiros de Joseph Harold Greenberg (1915-2001).

Ao propor, em *The Languages of Africa* (1963), a redução das línguas africanas a quatro grandes famílias — afro-asiática, níger-congo, nilo-saariana e khoisan —, Greenberg rompeu com taxonomias anteriores, muitas vezes superficiais e culturalmente enviesadas, fundadas em critérios raciais ou geográficos pouco rigorosos. Sua proposta abriu caminho para um entendimento mais profundo da evolução, das relações de parentesco e das interconexões entre os idiomas do continente (Petter, 2015).

Paralelamente, no âmbito dos estudos da linguagem, consolidava-se a Linguística Cognitiva, abordagem que concebe a língua não como um sistema autônomo, mas como faculdade profundamente conectada às demais capacidades cognitivas humanas — percepção, memória, raciocínio e categorização (Ibarretxe-Antuñano; Valenzuela, 2012; Ferrari, 2011). No seio dessa corrente, a teoria dos protótipos, desenvolvida pela psicóloga Eleanor Rosch (1978) e posteriormente incorporada à linguística por George Lakoff (1987), forneceu um modelo influente para compreender como a mente humana organiza o conhecimento em categorias.

O presente estudo situa-se na intersecção desses dois campos. Seu objetivo é articular criticamente os fundamentos conceituais de protótipo e categorização, próprios da Linguística Cognitiva, com as propostas classificatórias de troncos linguísticos desenvolvidas por Greenberg, com especial atenção à macrofamília Níger-Congo. Defende-se a hipótese de que toda classificação linguística é, em última instância, um ato de categorização, e que, por isso mesmo, os efeitos prototípicos podem iluminar tanto o modo como concebemos as famílias linguísticas quanto os próprios procedimentos heurísticos que orientaram a obra greenberguiana. Ao mesmo tempo, mantém-se vigilância quanto aos limites dessa articulação, distinguindo o parentesco genético da representação cognitiva das categorias.

Cabe registrar, por fim, a dimensão formativa e ético-política do tema. Em um contexto de valorização da consciência negra e de combate ao apagamento histórico das culturas africanas, reconhecer a riqueza e a sistematicidade das línguas do continente é também um gesto de reparação epistemológica, que recoloca a África no centro da reflexão linguística contemporânea.

A LINGUÍSTICA COGNITIVA E A CATEGORIZAÇÃO

Linguagem, Cognição e Experiência

A Linguística Cognitiva é uma abordagem que estuda a relação entre a linguagem, o pensamento e a cognição humana. Diferentemente de teorias formais que tratam a linguagem como um módulo autônomo, dotado de regras independentes das demais funções mentais, a Linguística Cognitiva sustenta que a linguagem reflete e está profundamente conectada às estruturas cognitivas gerais, como percepção, memória e raciocínio (Ibarretxe-Antuñano; Valenzuela, 2012). Nessa perspectiva, o significado não é um conjunto fixo de traços, mas uma construção dinâmica, ancorada na experiência corporal e cultural dos falantes — princípio que Lakoff e Johnson (1980) denominaram experiencialismo ou cognição corporificada.

Desse pressuposto decorre a centralidade da categorização. Categorizar é a operação cognitiva fundamental pela qual organizamos a infinidade de estímulos do mundo em classes manejáveis, atribuindo-lhes significado. Toda língua, ao nomear, já categoriza; e toda ciência que classifica seu objeto — inclusive a própria linguística, ao agrupar idiomas em famílias — exerce, no plano metateórico, uma atividade categorizadora. Compreender como funcionam as categorias é, portanto, condição para refletir criticamente sobre qualquer taxonomia.

Da Categorização Clássica à Teoria dos Protótipos

Por mais de dois milênios, prevaleceu no pensamento ocidental a concepção clássica ou aristotélica de categoria. Segundo esse modelo, as categorias são definidas por um conjunto de condições necessárias e suficientes; seus limites são nítidos e rígidos (um elemento ou pertence ou não pertence); e todos os seus membros gozam do mesmo estatuto, sem que haja exemplares mais ou menos representativos. Embora elegante, essa concepção mostrou-se inadequada para descrever boa parte das categorias naturais e linguísticas.

A primeira fissura veio da filosofia. Ao analisar o conceito de jogo, Wittgenstein (1953) observou que não existe um traço único compartilhado por todos os jogos — de tabuleiro, de azar, de bola, de palavras —, mas, sim, uma rede de semelhanças de família (*Familienähnlichkeiten*): assim como os membros de uma família partilham traços que se cruzam e se sobrepõem sem que nenhum seja comum a todos, os jogos formam uma categoria por encadeamento de semelhanças parciais. Essa noção tornar-se-ia uma das pedras angulares da teoria dos protótipos.

A consolidação experimental dessa intuição deve-se a Eleanor Rosch. Em uma série de estudos da Psicologia Cognitiva, Rosch (1978) demonstrou que as categorias humanas exibem efeitos de prototipicidade: diante da categoria ave, por exemplo, falantes tendem a julgar o pardal ou o sabiá como exemplares mais representativos do que o pinguim ou o avestruz. Os membros de uma categoria, portanto, não são equivalentes entre si; organizam-se em torno de um protótipo — o melhor exemplar, mentalmente mais saliente — a partir do qual os demais se distribuem por graus de tipicidade, em fronteiras difusas e graduais.

Rosch identificou ainda a existência de um nível básico de categorização (por exemplo, cão, situado entre o nível superordenado animal e o subordinado poodle), no qual a cognição opera com máxima eficiência informativa. Esses achados foram corroborados, no campo do léxico, pelos estudos de Berlin e Kay (1969) sobre os termos básicos de cor, que revelaram regularidades translinguísticas na organização perceptual e categorial das cores.

Coube a George Lakoff, em *Women, Fire, and Dangerous Things* (1987), transpor sistematicamente esses princípios para a teoria linguística. Lakoff descreveu as categorias como estruturas radiais, organizadas a partir de um centro prototípico e de extensões motivadas — porém não previsíveis — em direção à periferia, e ancorou-as nos chamados Modelos Cognitivos Idealizados (MCIs), estruturas de conhecimento culturalmente compartilhadas que sustentam o sentido. Desse modo, a categorização deixa de ser uma operação lógica abstrata e passa a ser compreendida como um processo experiencial, gradiente e culturalmente situado (Ferrari, 2011).

A CLASSIFICAÇÃO DAS LÍNGUAS AFRICANAS POR GREENBERG

O Método da Comparação Multilateral

Joseph Harold Greenberg foi um linguista e antropólogo norte-americano, renomado por seu trabalho em tipologia e classificação de línguas. Para enfrentar a vastidão do panorama linguístico africano, Greenberg valeu-se de um método conhecido como comparação multilateral (ou *mass comparison*), que consiste na comparação simultânea do vocabulário básico de um grande número de línguas, em vez do cotejo par a par característico do método histórico-comparativo tradicional. A inspeção de amplos conjuntos de dados permitiria, segundo o autor, identificar padrões recorrentes de similaridade e inferir, a partir deles, relações de parentesco genético entre os grupos.

O procedimento foi inovador e, ao mesmo tempo, controverso. De um lado, sua amplitude possibilitou uma visão panorâmica e a proposição de agrupamentos abrangentes em tempo relativamente curto; de outro, a rapidez e a relativa informalidade do método suscitaram debates quanto ao seu rigor demonstrativo, como se discutirá adiante.

As Quatro Macrofamílias

Em *The Languages of Africa* (1963), Greenberg propôs a organização das línguas do continente em quatro grandes famílias ou macrofamílias:

a) Afro-asiática — abrange línguas semíticas (árabe, amárico, hebraico), berberes (no norte da África), cushíticas (somali, oromo, afar, no Chifre da África) e chádicas (como o hauçá), distribuídas pelo norte do continente, pelo Chifre da África e por partes do Sahel;

b) Níger-Congo (também denominada Níger-Cordofoniana) — a maior família linguística da África em número de línguas e de falantes, que engloba, entre outros, o vasto grupo banto (suaili, zulu, xhosa) e ocupa quase toda a África subsaariana;

c) Nilo-saariana — agrupamento mais heterogêneo e debatido, com línguas faladas sobretudo na região do Nilo e em partes do Sahel (núbio, dinka, nuer, massai);

d) Khoisan (ou coissã) — conjunto relativamente pequeno de línguas do sul da África, célebres por seus sons de clique consonantais, faladas sobretudo na Namíbia, em Botsuana e na África do Sul.

Essa arquitetura quadripartite, posteriormente refinada por diversos pesquisadores, tornou-se referência incontornável nos estudos de linguística africana.

Recepção e Crítica

O trabalho de Greenberg foi, a um só tempo, revolucionário e criticado. Parte dos linguistas argumenta que a comparação multilateral carece da precisão do método histórico-comparativo clássico, o qual exige a reconstrução sistemática de correspondências sonoras regulares e a análise detalhada de cada par de línguas para confirmar relações genéticas. Sem esse aparato, sustentam os críticos, semelhanças superficiais poderiam ser confundidas com parentesco efetivo.

Não obstante, as hipóteses de Greenberg revelaram-se extraordinariamente fecundas: forneceram um ponto de partida que permitiu a outros estudiosos — como Kay Williamson, Derek Nurse e Paul Newman — testar, refinar e, por vezes, refutar aspectos específicos da classificação, ampliando consideravelmente o conhecimento sobre as relações internas das famílias africanas (Heine; Nurse, 2000). Muitas das categorias que ele propôs continuam, até hoje, a orientar a tipologia e a classificação das línguas, o que o consagra como marco fundador desse campo.

UMA LEITURA PROTOTÍPICA DA MACROFAMÍLIA NÍGER-CONGO

A metáfora da “família linguística” e a semelhança de família

Há uma convergência conceitual notável — e até hoje pouco explorada — entre a metalinguagem da linguística histórica e a teoria dos protótipos. Ao falar em famílias de línguas, em parentesco, em línguas-irmãs e em línguas-mãe, a tradição comparativa mobiliza, de modo sistemático, a mesma metáfora do parentesco familiar que Wittgenstein (1953) tomou como modelo para descrever as categorias por semelhança. Essa coincidência não é fortuita: ela sugere que a própria razão classificatória da linguística se apoia, em seu nível metafórico, em um esquema cognitivo de organização por encadeamento de semelhanças.

A macrofamília Níger-Congo oferece um terreno privilegiado para examinar essa convergência. Conforme proposta por Greenberg e detalhada por autores posteriores (Williamson, 1989, *apud* Petter, 2015), trata-se de uma família extensa e internamente complexa, ramificada em troncos como o atlântico-congo, o mandê e o cordofaniano, e em subgrupos como o gur, o kwa, o benue-congo e, dentro deste, o numeroso grupo banto. Os membros dessa categoria não são homogêneos: distribuem-se de forma desigual quanto ao número de falantes, à documentação disponível e à projeção sociocultural — exatamente o tipo de assimetria que a teoria dos protótipos se propõe a descrever.

O Suaíli como Protótipo Categorical

Se aplicarmos a lógica prototípica à categoria línguas Níger-Congo, o suaíli (kiswahili) emerge como forte candidato a membro prototípico — o exemplar mentalmente mais saliente e representativo da família. Diversos fatores convergem para essa centralidade. Em primeiro lugar, o suaíli é, de longe, a língua banta com maior número de falantes, funcionando como língua franca em vasta porção da África Oriental e Central. Em segundo lugar, é a língua africana mais documentada, padronizada e ensinada fora do continente, o que reforça sua disponibilidade cognitiva. Em terceiro lugar, ocupa, na economia da representação, algo análogo a um nível básico roschiano: quando se evoca a noção de língua africana, o suaíli é, com frequência, o primeiro exemplar a vir à mente de falantes não especialistas.

Nessa leitura, o suaíli não é um membro “mais legítimo” da família em termos genéticos — todos os seus membros descendem igualmente do proto-Níger-Congo —, mas o exemplar mais típico do ponto de vista da saliência cognitiva e cultural. Trata-se, portanto, de prototipicidade representacional, e não de hierarquia filogenética.

Membros Periféricos e Gradiência

Em torno desse núcleo prototípico distribuem-se, por graus decrescentes de tipicidade, outros membros da categoria. Línguas como o iorubá, o igbo, o xhosa e o zulu, embora de enorme importância demográfica e cultural em suas respectivas regiões, tendem a ocupar posições mais periféricas no imaginário linguístico não especializado; e línguas como o lingala ou o quicongo, ainda menos projetadas internacionalmente, situam-se em faixas de tipicidade ainda mais distantes do centro. Não há, entre esses membros, uma fronteira nítida de pertencimento, mas uma gradiência — precisamente a estrutura que Rosch (1978) descreveu para as categorias naturais.

Importa frisar que essa gradiência é de natureza cognitivo-cultural, e não taxonômica: ela diz respeito ao grau em que cada língua é percebida como exemplar representativo da categoria, e não ao seu direito de pertencer a ela. Sob esse ângulo, a teoria dos protótipos não substitui a classificação genética de Greenberg, mas oferece uma chave para compreender como essa classificação é mentalmente representada, hierarquizada e ensinada.

Tensões Teóricas: Parentesco Genético *versus* Categorização Prototípica

A articulação proposta exige, no entanto, uma ressalva metodológica fundamental, sob pena de incorrer em equívoco categorial. A classificação genética e a categorização prototípica operam em planos distintos e obedecem a princípios diferentes. O pertencimento a uma família linguística é, em rigor, categórico: uma língua descende — ou não — de determinada protolíngua, e essa filiação não admite graus. A categorização prototípica, ao contrário, é gradiente: opera com melhores e piores exemplares, com fronteiras difusas e com efeitos de tipicidade.

Disso decorre que os efeitos prototípicos não incidem sobre a filiação genética das línguas, mas sobre a representação que falantes e pesquisadores fazem dessa filiação. A teoria dos protótipos esclarece, ainda, dois outros aspectos da empreitada greenberguiana: de um lado, o próprio método da comparação multilateral apoia-se na percepção de padrões de semelhança — operação cognitiva por excelência — para inferir parentesco; de outro, os limites controversos de famílias como a nilo-saariana podem ser relidos como zonas de fronteira difusa entre categorias, isto é, como casos periféricos cuja classificação é cognitivamente menos estável.

A articulação entre os dois campos é, portanto, de ordem heurística e representacional, e não genealógica. Reconhecer esse limite, longe de enfraquecer a proposta, confere-lhe precisão: a Linguística Cognitiva não corrige Greenberg no terreno da reconstrução histórica, mas ilumina a dimensão cognitiva subjacente tanto ao ato de classificar quanto ao modo como as classificações são apropriadas.

CONSIDERAÇÕES FINAIS

Este estudo procurou demonstrar que a teoria dos protótipos e a noção de categorização, oriundas da Linguística Cognitiva, oferecem uma chave de leitura fecunda para a clássica classificação das línguas africanas proposta por Joseph Greenberg. Partindo do princípio de que toda taxonomia é, antes de tudo, um ato de categorização, mostrou-se como a macrofamília Níger-Congo pode ser representada segundo uma estrutura prototípica, na qual o suaíli ocupa a posição de exemplar central e as demais línguas se distribuem por graus de tipicidade.

Ao mesmo tempo, buscou-se preservar o rigor conceitual, distinguindo o parentesco genético — categórico — da categorização prototípica — gradiente —, e situando a articulação entre os dois campos no plano heurístico e representacional. As classificações de Greenberg, frequentemente debatidas e refinadas, permanecem um marco fundador da linguística africana; a perspectiva cognitivista, por sua vez, não as substitui, mas enriquece-lhes a compreensão, ao explicitar os mecanismos mentais que sustentam o gesto de classificar.

Para além do interesse teórico, o estudo reafirma o valor formativo e ético de se debruçar sobre as línguas africanas. Reconhecer sua sistematicidade, sua diversidade e sua centralidade nos estudos linguísticos é também um modo de

combater apagamentos históricos e de afirmar, no terreno da ciência, a riqueza das culturas do continente. Investigações futuras poderão estender a análise prototípica às demais macrofamílias e examinar, em dados empíricos de fala e de ensino, os efeitos de tipicidade aqui sugeridos.

REFERÊNCIAS

- BERLIN, B.; KAY, P. **Basic Color Terms: their universality and evolution**. Berkeley: University of California Press, 1969.
- FERRARI, L. C. **Introdução à Linguística Cognitiva**. 2. ed. São Paulo: Contexto, 2011.
- GREENBERG, J. H. **The Languages of Africa**. Bloomington: Indiana University Press, 1963.
- HEINE, B.; NURSE, D. (Eds.). **African Languages: An Introduction**. Cambridge: Cambridge University Press, 2000.
- IBARRETXE-ANTUÑANO, I.; VALENZUELA, J. (Dir.). **Lingüística Cognitiva**. Barcelona: Anthropos Editorial, 2012.
- LAKOFF, G. **Women, Fire, and Dangerous Things: what categories reveal about the mind**. Chicago: University of Chicago Press, 1987.
- LAKOFF, G.; JOHNSON, M. **Metaphors We Live By**. Chicago: University of Chicago Press, 1980.
- PETTER, M. **Introdução à Linguística Africana**. São Paulo: Contexto, 2015.
- ROSCH, E. **Principles of Categorization**. In: ROSCH, E.; LLOYD, B. (Eds.). *Cognition and Categorization*. Hillsdale: Lawrence Erlbaum Associates, 1978. p. 27-48.
- SIMONS, G. F.; FENNIG, C. D. (Eds.). **Ethnologue: languages of the world**. 26. ed. Dallas: SIL International, 2023. Disponível em: www.ethnologue.com. Acesso em: 18 nov. 2024.
- WITTGENSTEIN, L. **Investigações Filosóficas**. Tradução de José Carlos Bruni. São Paulo: Nova Cultural, 1999 [1953].