

OS ESTUDOS EMPÍRICOS NA PESQUISA LINGUÍSTICA

*Ana Regina Calindro*¹

*Marije Soto*²

Desde os tempos mais remotos, os estudiosos analisavam obras clássicas com o intuito de verificar como o pensamento humano era desenvolvido (cf. ALTMAN, 2008). Logo, a análise de dados, através de textos consagrados, data de muito antes dos estudos empíricos empreendidos pelos fundadores do pensamento linguístico moderno, entre eles, Saussure (1916), Sapir (1921) e Bloomfield (1933). Desde então, os estudos linguísticos baseados em dados se diversificaram amplamente.

Com o tempo, os campos de investigação foram se ampliando e, portanto, demandando um maior grau de interdisciplinaridade. A mescla de saberes aumentou a complexidade e diversidade das perguntas de pesquisa. Ademais, o avanço da tecnologia também permitiu que *novas* perguntas pudessem ser feitas. Deste modo, metodologias centenárias convivem hoje com o que há de mais moderno, apresentando um leque rico e diverso de pesquisas com base em *corpus* histórico, coletas etnográficas, construção e consulta de *corpora* constituídos digitalmente, experimentos psicolinguísticos, análises acústicas por softwares especializados, estudos com neuroimagem, entre outros. Toda essa grande gama de pesquisas linguísticas tem sua vez nesta edição da Revista Linguística. O objetivo deste número, portanto, foi o de reunir uma amostra de pesquisas recentes baseadas nas mais diversas vertentes teóricas de análise e coleta de dados.

Diversas áreas da Linguística usam dados como base ou verificação de hipóteses, tais como: linguística histórica, linguística experimental (psicolinguística e neurolinguística), aquisição da linguagem, sociolinguística, linguística cognitiva, linguística de corpus, entre outras.

1 Profª. Adjunta do Departamento de Linguística e Filologia, Faculdade de Letras, Universidade Federal do Rio de Janeiro; membro do Programa da Pós-graduação em Linguística, Universidade Federal do Rio de Janeiro; email: anacalindro@letras.ufrj.br.

2 Profª. Adjunta do Departamento de Estudos da Linguagem, Instituto de Letras, Universidade do Estado do Rio de Janeiro; membro do Programa da Pós-graduação em Linguística, Universidade Federal do Rio de Janeiro; email: marijesoto@hotmail.com.

Este volume se inicia com dois textos que discorrem sobre a importância da escolha e elaboração de *corpora* com o intuito de formar uma base de dados que sirva para pesquisas futuras. Lílian Teixeira de Souza, em “Sobre a constituição de *corpora* para línguas com poucos recursos (*less resourced languages*)”, traz uma reflexão sobre os critérios estabelecidos pela Linguística de *Corpus* na composição de *corpora* para análises linguísticas em geral. Como, por exemplo, a constituição de documentação de línguas ameaçadas. Com este foco, a autora apresenta considerações sobre os *corpora* eletrônicos disponíveis de línguas com maior presença digital para, então, abordar a importância e relevância em se constituírem *corpora* de línguas com menor presença digital e menos recursos, como línguas indígenas brasileiras.

Em seguida, Evani Viotti, em seu artigo “Avaliando a vitalidade linguística em contextos de multilinguismo: etnografias versus modelos computacionais”, aborda a importância em se considerar questões socioeconômicas e políticas que caracterizam certas ecologias multilíngues ao estudar línguas como da região do Alto do Rio Negro na Amazônia. Neste contexto, a autora traz um panorama interessante sobre as relações sociais dos povos que habitam essa região. Tais considerações são a base para que se defenda a língua como um fenômeno social, parte de um sistema complexo, aberto, dinâmico e auto-organizado. Portanto, segundo a autora, a teoria da complexidade (BECKNER *et al.*, 2009; ELLIS; LARSEN-FREEMAN, 2009; KRETZSCHMAR, 2015; MUFWENE, 2013; MUFWENE; COUPÉ; PELLEGRINO, 2017; entre muitos outros) seria o caminho mais adequado para dar conta das características dinâmicas das línguas.

Seguindo essa vertente no âmbito da valorização e necessidade de ampliar o conhecimento de dialetos de regiões menos estudadas no Brasil - devido a questões sociais, históricas e econômicas (que vão além do escopo deste texto) - apresentamos dois artigos que trazem dados das regiões norte e nordeste do país.

Em primeiro lugar, as autoras Rosilene Lemos e Camila dos Santos Brito, integrantes do projeto AMPER-Amazônia, vinculado ao AMPER-POR, analisaram a variação geoprosódica do português falado na Amazônia Oriental no artigo “Estudo sociofonético do português falado na Amazônia: uma comparação interdialetoal entre Belém e Macapá”. Com base na metodologia do projeto e em outros estudos que abarcam diferentes regiões do Brasil (cf. MORAES, 1993; CRUZ; SEARA; MOUTINHO, 2015), as autoras analisaram o contorno prosódico nos vocábulos oxítonos presentes em sentenças declarativas neutras e interrogativas totais, a fim de observar a variação entoacional da frequência fundamental (F0) em posição nuclear dos enunciados. Os resultados desta pesquisa revelaram semelhanças prosódicas nos falares de Belém e Macapá.

O artigo “O papel da morfologia construcional na formação de antropônimos neológicos” de Letícia Santos Rodrigues, analisa dados históricos da Bahia, mostrando a relevância e importância de pesquisas que levam em consideração a ampla variação linguística existente no Brasil. Com base na Morfologia Construcional (cf. BOOIJ, 2010), a autora analisou antropônimos registrados em fichas de indivíduos associados a uma ordem religiosa de Salvador/Bahia nos séculos XIX e XX. A partir de antropônimos que apresentavam formativos germânicos (como *ald-* ~ *-ald/nald-* ~ *-nald/vald-* ~ *-vald*) foi confirmada a presença do modelo bitemático nos antropônimos neológicos. Além disso, verificou-se que quando os indivíduos eram expostos a um conjunto de antropônimos que se utilizam de dado formativo ou a um antropônimo muito produtivo, são capazes, por analogia, de depreender o modelo e reproduzi-lo ao criar novos antropônimos, seguindo esquemas construcionais pressupostos pela Linguística Cognitiva (cf. FERRARI, 2014).

Ainda no âmbito social e das análises linguísticas baseadas no uso, o artigo “Estereótipos na concordância de gênero em profissões: efeitos de frequência e saliência”, de Bruno Marques Pinheiro e Raquel Freitag, apresenta um estudo experimental. Os autores pediram o julgamento da masculinidade e da feminilidade de nomes comuns de dois gêneros de profissões no português brasileiro (como *o/a dentista*), com intuito de verificar se os falantes constroem representações mentais a partir dos estereótipos de gênero. Os autores hipotetizam que essas representações sofrem efeitos da frequência e de saliência - parâmetros que também afetam a gramática. Após a análise dos resultados, verificou-se que a frequência prototípica da profissão interferiu positivamente no julgamento, logo os parâmetros da estereotipia de gênero, medidos pela saliência e frequência, interferem no julgamento das profissões quanto ao gênero nos nomes comuns de dois gêneros. De acordo com os autores, o resultado reflete que o cérebro categoriza informações a partir das experiências linguísticas e sociais dos falantes, em que aspectos probabilísticos são fatores determinantes (BOD; HAY; JANNEDY, 2003).

Outra área da linguística com amplo uso de uma base empírica para as suas constatações é a psicolinguística. Este campo de investigação com foco nos aspectos mentais que possibilitam a aquisição, compreensão e produção de linguagem incorporou as metodologias da psicologia experimental. Nesta revista, há artigos sobre a aquisição de L1 e sobre a aprendizagem de L2, utilizando uma variedade de metodologias psicolinguísticas para verificação de hipóteses.

No artigo “Testing Children’s knowledge of generic null pronouns”, a autora Karina Bertolino aplica uma tarefa de julgamento de valor de verdade a um grupo de crianças brasileiras de quatro anos para observar se já fariam uma leitura genérica de estruturas impessoais, como em *Nessa escola não pode escovar os dentes depois de comer*. A partir de um entendimento do processo de aquisição desse

fenômeno dentro da teoria de Princípios e Parâmetros (HOLMBERG, 2016), ela mostra e explica que crianças de quatro anos já demonstram interpretações genéricas, o que distingue, por exemplo, o português europeu do português brasileiro. Ela ainda detalha um novo caminho metodológico para fazer estudos com crianças a partir dos dois anos, usando um paradigma de olhar preferencial, permitindo, assim, estudos futuros para verificar a interpretação impessoal ainda mais precocemente, pautado em evidências de produção de impessoais nulos de crianças adquirindo estoniano (TORN-LEESIK; VIJA, 2012).

Por um semelhante viés teórico, Fernanda Mendes discute a aquisição de estruturas de posse inalienável adjetivada em português brasileiro, tais como *O João tem os dedos quebrados* e *O João levantou os dedos quebrados*. A autora, que é especialista no assunto, apresenta os resultados de uma Tarefa de Seleção de Figura aplicada a duas faixas etárias (de quatro a cinco anos; de seis a oito), e a um grupo de controle com participantes adultos no artigo intitulado “Estruturas de posse inalienável adjetivada: apontamentos iniciais sobre a sua aquisição em português brasileiro”. Na tarefa, os participantes têm que escolher a figura compatível com sentenças como *Me mostra quem tem a juba laranja*. O objeto da discussão é a interpretação de posse inalienável, disponível ou não de acordo com a presença e/ou tipo de adjetivo que modifica o nome possuído. O motivo para essa questão são restrições dessa natureza em francês (KAYNE, 1975). Os dados, portanto, pretendem investigar se crianças estendem a interpretação inalienável, ou já mostram as restrições características da gramática alvo, como evidenciado pelo comportamento dos adultos.

O estudo com título “Interferência de verbos de concordância da Libras (L1) no uso de preposições na interlíngua de surdos aprendizes de português L2 (escrito)” das autoras Aline Mesquita e Heloisa Salles também parte de uma análise formal. As autoras lançam mão do conceito de interlíngua para investigar a influência de verbos que marcam a sua relação temática com argumentos com movimento, os chamados verbos de concordância em libras, comparando-os com dados sobre a aprendizagem do uso da preposição *para* no português brasileiro na modalidade escrita. Uma tarefa de produção controlada investiga as interferências positiva e negativa na interlíngua de um grupo de surdos adultos em escolas que promovem o bilinguismo (com a modalidade sinalizada para libras, e a escrita para PB).

Continuando a temática do bilinguismo, mas agora para aprendizes de inglês, os autores Ricardo Augusto de Souza e Jamila Viegas Rodrigues mostram como a influência de uso linguístico oral em L2 (inglês) afeta a capacidade discriminatória em pares mínimos (ex.: *them* /ðem/ ‘eles, elas’; *then* /ðen/ ‘então’) por falantes não nativos de inglês em um experimento de rastreamento ocular

nos moldes do paradigma mundo-visual (TANENHAUS *et al.*, 1995). A fundamentação teórica no seu artigo intitulado “O impacto do uso linguístico oral em L2 no reconhecimento de palavras por bilíngues tardios do português e do inglês” se baseia na Teoria de Exemplares (BYBEE, 2001), que explica a importância do uso na construção de representações sonoras do léxico em L2. Dessa forma, os autores não apenas apontam para a relevância de controlar a variável de uso linguístico mediante questionários de auto-declaração em pesquisas de bilinguismos, como também a relevância desse uso para o processo didático.

A psicolinguística e as suas incursões na neurociência também são centrais na resenha e na entrevista. Erica dos Santos Rodrigues assina a resenha do livro *Psicolinguística e Metacognição na Escola* organizado por Marcus Maia. Ela dá o título “O laboratório na escola e o desenvolvimento da metacognição linguística: uma experiência translacional com leitura”, o que remete à aplicação de modelos e metodologias da psicolinguística para diagnosticar e investigar estratégias cognitivas e metacognitivas dos alunos em relação à leitura, trazendo o laboratório de LER - Laboratório de Eletrofisiologia e Rastreamento Ocular (<http://www.ler.letras.ufrj.br/>) literalmente para a sala da aula das turmas de 8o e 9o ano de uma escola estadual. Em grande detalhe, a autora descreve os oito capítulos que apresentam os projetos de pesquisa de uma equipe multidisciplinar de três instituições do Rio de Janeiro – UFRJ, UERJ e CEFET-RJ. Centrais à proposta estão a coleta de dados e a oferta de oficinas aos alunos com intuito de desenvolver uma reflexão crítica acerca da construção cognitiva na qual estão inseridos a partir de um raciocínio científico. Sobre essa ambição, a autora conclui que é “uma ferramenta de metacognição que leva o leitor a examinar o seu próprio modo de pensar o ensino de língua na escola”.

A entrevista com a pesquisadora Katrien Segearts revela igual caráter ambicioso e inédito. A coordenadora do laboratório de Neurociências da Linguagem da Escola de Psicologia da Universidade de Birmingham (Reino Unido) desenvolve estudos sobre o processamento sintático e semântico associado a contextos comunicativos usando paradigmas experimentais inovadores. A equipe trabalha com a interação com avatares, entre interlocutores no scanner da ressonância magnética, e, na mais nova proposta, por jogos comunicativos. Nessa última proposta, os participantes irão jogar enquanto há coleta de EEG em tempo real. Ademais, os seus estudos com idosos geraram atenção na mídia, pois evidenciam a relação entre manutenção de processamento linguístico e a boa forma física. Na entrevista, ela fala sobre interdisciplinaridade, cooperação entre colegas, e experimentação com novos modos de investigar linguagem, mente e cérebro.

Em um estudo com forte controle metodológico, Gustavo Lopez Estivalet discute as “Variáveis

lexicais e ortográficas no acesso lexical das palavras do português brasileiro”. A partir de um experimento de decisão lexical na modalidade visual, ele objetiva investigar a frequência lexical, vizinhança ortográfica, categoria gramatical, número de letras e sílabas como possíveis influências no acesso lexical. O controle dessas variáveis é essencial na verificação de modelos da organização do léxico mental vigentes na literatura. Deste modo, Estivalet segue na sua trajetória de uma série de artigos com intuito e contribuir para um maior rigor na validação e seleção de estímulos lexicais com o uso de corpora na pesquisa em psicolinguística experimental no português brasileiro.

Por fim, as autoras Késsia Henrique e Cristina Name apresentam resultados de um estudo clássico psicolinguístico no artigo “O processamento de múltiplas dependências não adjacentes no PB”. Nele, as autoras investigam o processamento de estruturas com elementos codependentes, marcados morfossintática e/ou fonologicamente, com material interveniente. O possível custo cognitivo para essas dependências não adjacentes do tipo aninhado (ex. O_{A1} carro $A1$ que os $A2$ mecânicos $A2$ consertaram $B2$ bateu $B1$) já tem sido bastante explorado em estudos de aquisição. As autoras contribuem neste estudo, que conta com dois experimentos do tipo tarefa de labirinto (*maze task*), com dados importantes de processamento em adultos. Tempo de leitura maior na condição aninhada sugere que dependências nessa configuração causam maior demanda também no processamento linguístico adulto.

A grande variabilidade de temáticas, metodologias e perspectivas teóricas refletem a grande diversidade e complementaridade dos estudos linguísticos existentes atualmente no Brasil. Há um fio condutor no conjunto apresentado que é a preocupação em gerir, gerar e analisar dados com rigor metodológico, com as mais variadas possibilidades: documentar línguas, falsear hipóteses e investigar fenômenos empíricos no âmbito linguístico. Nós, editoras desta edição, esperamos que vocês, os leitores, encontrem inspiração científica a partir da riqueza dessa pluralidade.

REFERÊNCIAS

ALTMAN, Maria Cristina Fernandes Salles. Os eternos gramáticos: algumas considerações sobre norma e prescrição na descrição linguística. *In Homenagem: 80 anos de Evanildo Bechara*, 2008.

BECKNER, Clay; BLYTHE, Richard; BYBEE, Joan; CHRISTIANSEN, Morten H.; CROFT, William; ELLIS, Nick C.; HOLLAND, John; KE, Jinyun; LARSEN-FREEMAN, Diane; SCHOENEMANN, Tom. Language is a complex adaptive system. *Language Learning*, Ann Arbor, vol. 59, n.1, p. 1-26, dezembro 2009.

BOD, Rens; HAY, Jennifer; JANNEDY, Stefanie. *Probabilistic linguistics*. Mit Press, 2003.

BLOOMFIELD, Leonard. *Language*. New York: Holt, Rinehart & Winston, 1933.

BOOIJ, Geert. Construction morphology. *Language and Linguistics Compass*, United Kingdom, v. 4, n. 7, p. 543-555, 2010.

BYBEE, Joan. *Phonology and Language Use*. Cambridge: Cambridge University Press, 2001.

CRUZ, Regina; SEARA, Isabel; MOUTINHO, Lurdes. Intonation of Yes/No question in Portuguese: analysis of the relationship between intonation and lexical stress in the AMPER-POR corpus. In Nebot, Adrián Cabedo (ed.). *Perspectivas actuales en el análisis fónico del habla. Tradición y avances en la fonética experimental*, Valência: Universidade de Valência, p. 35-44, 2015.

ELLIS, Nick C.; LARSEN-FREEMAN, Diane. *Language as a complex adaptive system*. Oxford: Wiley Blackwell, 2009.

FERRARI, Lilian. *Introdução à linguística cognitiva*. São Paulo: Contexto, 2014.

HOLMBERG, Anders. Linguistic Typology. In ROBERTS, Ian. (org.), *The Oxford Handbook of Universal Grammar*, 354–76. Oxford University Press, 2016.

KAYNE, Richard. *French syntax*. Cambridge, Massachusetts: The MIT Press, 1975.

KRETZSCHMAR JR., William A. *Linguistics and complex systems*. Cambridge: Cambridge University Press, 2015.

MORAES, João Antônio. A Entoação Modal Brasileira: Fonética e Fonologia. *Caderno de Estudos Linguísticos: Campinas*, v. 25, p. 101-111, 1993.

MUFWENE, Salikoko S. The emergence of complexity in language. An evolutionary perspective. In: MASSIP-BONET, Àngels; BASTARDAS-BOADA, Albert (eds.), *Complexity perspectives on language, communication and society*. Dordrecht: Springer, 2013. p.197-218.

MUFWENE, Salikoko S.; COUPÉ, Christophe; PELLEGRINO, François. *Complexity in Language. Developmental and evolutionary perspectives*. Cambridge: Cambridge University Press, 2017.

SAPIR, Edward *Language: An Introduction to the Study of Speech*. Roxy Harris, Ben Rampton, 1921

SAUSSURE, Ferdinand de. *Curso de linguística geral*. São Paulo: Cultrix, 1975[1916].

TANENHAUS, Michael K.; SPIVEY-KNOWLTON Michael J.; EBERHARD, Kathleen; SEDIVY, Julie C. Integration of visual and linguistic information during spoken language comprehension. *Science*, 268, 1632–1634, jun 1995.

TORN-LEESIK, Reeli; VIJA, Maigi. Acquisition of the Impersonal Voice by an Estonian Child. *Journal of Baltic Studies*. v. 43, n. 2, 251-271, 2012. Disponível em <https://doi.org/10.1080/01629778.2012>.