

A morfologia gramatical contextual e sua relação com a proporção da morfologia livre

Grammatical contextual morphology and its relationship with the proportion of unbound morphology

Eudes Barletta Mattos¹, João Paulo Lazzarini-Cyrino²

Universidade Federal da Bahia (Brasil)

RESUMO

A expressão morfológica de significados gramaticais não só varia nas línguas do mundo quanto à forma – afixos, clíticos, partículas –, como também quanto a sua determinação pela sintaxe. Categorias gramaticais determinadas pela sintaxe, também chamadas de *morfossintáticas* ou *contextuais*, têm sido descritas sob a ótica de processos flexionais ligados ao interior da palavra; no entanto, muitas línguas usam morfologia livre para veicular significados gramaticais – existindo casos de expressão de categorias contextuais por partículas. Usando gramáticas descritivas de 19 línguas amostradas ao redor do mundo como fonte de dados, e quantificando frequências de morfemas livres ou presos, investigamos a correlação entre morfologia gramatical contextual e a proporção de morfologia gramatical livre. Há uma correlação negativa entre esses dois parâmetros, e uma tendência a maior expressão de categorias determinadas pela sintaxe em línguas com altos percentuais de morfologia presa.

PALAVRAS-CHAVE:

Morfossintaxe. Tipologia Linguística. Afixo. Partícula. Flexão Contextual.

Recebido em: 20/08/2024

Aceito em: 10/12/2024

ABSTRACT

The morphological expression of grammatical meaning not only varies across the world's languages in formal terms – affixes, clitics, particles –, but also in terms of its determination by syntax. Grammatical categories determined by syntax, also called morphosyntactic or contextual, have been described from the perspective of inflectional processes inside the word; however, many languages use unbound morphology to convey grammatical meanings – there are cases of contextual categories being expressed by particles. Using descriptive grammars from 19 languages sampled around the world as data source, and quantifying frequencies of unbound or bound morphemes, we investigated the correlation between contextual grammatical morphology and the proportion of unbound grammatical morphology. There is a negative correlation between these two parameters, and a tendency towards greater expression of categories determined by syntax in languages with high percentages of bound morphology.

KEYWORDS:

Morphosyntax. Linguistic Typology. Affix. Particle. Contextual inflection.

¹ E-mail: eudesbarlettam@gmail.com | ORCID: 0000-0002-7346-5154

² E-mail: jpcyrino@gmail.com | ORCID: 0000-0002-3462-7114

1. Introdução

A linguística tem tradicionalmente dividido os significados numa dicotomia lexical-gramatical, cabendo a estes últimos o desempenho de diversas funções no sistema de uma dada língua; daí receberem também a denominação de *funcionais* (Spencer, 2000; Croft, 2002). Edward Sapir, já em 1921, chama a atenção ao fato de que cada língua tem um ou mais métodos para exprimir significados gramaticais, ou, em suas palavras, para “indicar a relação entre um conceito secundário e um elemento radical”. Por elementos radicais, entendemos que Sapir se referia aos elementos veiculadores de significados lexicais, referentes a conceitos por ele ditos *básicos* ou *fundamentais* (Sapir, 1921). A veiculação de ao menos alguns desses significados funcionais, realizada por meio de *processos gramaticais*, se observaria em todas as línguas do mundo, servindo para *afetar os conceitos básicos*. Esses processos variam em sua natureza: o autor elenca ordem de palavras, composição, afixação, modificação interna do radical ou elemento gramatical – seja numa vogal, seja numa consoante –, e diferenças acentuais e tonais.

O exame e amostragem das línguas do mundo, bem como o surgimento de bancos de dados como o WALS (Dryer e Haspelmath, 2013) e o AUTOTYP (Bickel e Nichols, 2014) de fato atestam haver uma série de estratégias possíveis para exprimir informação gramatical por parte das diferentes línguas. Em (1), temos um exemplo de língua cujos processos gramaticais se realizam, em grande parte, por meio da morfologia; percebe-se, por exemplo, que as funções sintáticas são marcadas pela existência de casos morfológicos estruturais, como nominativo e acusativo codificando sujeito e objeto direto, respectivamente. Atesta-se também a inserção das palavras em paradigmas de declinação surgidos graças ao alto grau de fusão entre morfemas comum às línguas indoeuropeias (Dryer e Haspelmath, 2013). Além disso, vê-se que o verbo se conjuga na terceira pessoa do singular para concordar com o sujeito *ṭwōrāšta*, “formador”, e se flexiona no passado perfeito para veicular a informação de tempo e aspecto:

(1) Antigo avéstico[†] (Irã: indoeuropeia)

<i>aṭ</i>	<i>zī</i>	<i>ṭwā</i>	<i>fšuyantaē=cā</i>
e	de.fato	2SG:ACC	criador.de.gado:DAT =e
<i>vāstrāi=cā</i>	<i>ṭwōrāšta</i>		<i>tatašā</i>
pastor:DAT=e	formador:NOM		criar.PST:3SG ³

³ Abreviações utilizadas: 1 primeira pessoa; 2 segunda pessoa; 3 terceira pessoa; ABL ablativo; ABS absolutivo; ACC acusativo; AGR concordância; ART artigo; ASP aspecto; AUX auxiliar; CF característica constante; CFIN final de

“E de fato o formador te criou para o benefício do criador de gado e do pastor.” (Yasna 29,6; Haig, 2008)

Em (2), porém, vemos uma sentença de uma língua que se apoia mormente na ordem das palavras para veicular a maioria de seus significados gramaticais: tanto o sujeito *kau* “você” quanto os objetos direto *joi* “dinheiro” e indireto *’imu* “ele” são indicados por sua posição em relação ao verbo, não havendo qualquer outra alteração fonológica ou acréscimo de material morfológico. Não há morfologia de tempo ou aspecto, que são inferidos pelo contexto ou pelo uso de advérbios, sem exigir flexão concomitante do verbo:

(2) Kéo (Ilha de Flores: austronésia)

<i>kau</i>	<i>négha</i>	<i>ti’i</i>	<i>’imu</i>	<i>joi</i>	<i>o</i>
2SG	já	dar	3SG	dinheiro	PERG

“Você deu o dinheiro a ele?” (Baird, 2002)

Por fim, em (3), vê-se uma língua que lança mão de diferenças tonais para veicular significados gramaticais – no caso, uma distinção aspectual:

(3) Guébie (Costa do Marfim: kru)

<i>e⁴</i>	<i>li³</i>	<i>ja³¹</i>	<i>e⁴</i>	<i>li²</i>	<i>ja³¹</i>
1.SG	comer.IPFV	coco	1.SG	comer.IPFV	coco

“eu comi coco”

“eu como coco” (Sande, 2017)

Entende-se que algumas línguas explicitam necessariamente um maior número de categorias gramaticais em suas sentenças; fenômenos como manifestação morfológica da concordância e da regência, por exemplo, são considerados fenômenos *maduros* (Dahl, 2004) ou *complexificações* redundantes presentes em apenas alguns sistemas linguísticos (Trudgill, 2011), enquanto muitas outras línguas suprem as mesmas necessidades comunicativas humanas sem a necessidade de exprimir essa morfologia redundante, como se vê na sentença kéo em (2) (Dahl,

constituente; CL classificador nominal; DAT dativo; DECL declarativo; DEF definido; DEM demonstrativo; DO objeto direto; ERG ergativo; FOC foco; IND indicativo; INF infinitivo; INS instrumental; IPD impeditivo; IPFV imperfectivo; LOC locativo; MIN número mínimo; NEG negação; NMLZ nominalizador; NOM nominativo; NONFUT não-futuro; NS não-singular; PERG marcador de pergunta; PFV perfectivo; PL plural; PRED predicativo; PST passado; R1 relacional 1; REL relativo; SG singular; VENT ventivo; VP prefixo vocálico

2004; McWhorter, 2007; Trudgill, 2011). *Grosso modo*, essa morfologia gramatical redundante corresponde ao que Geert Booij chama de *flexão contextual*; em suas palavras:

[...] há uma clara diferença em conteúdo semântico entre flexão inerente e flexão contextual. Flexão inerente expressa, assim como a derivação [expressa], uma certa quantidade de informação independente, enquanto a informação expressa por flexão contextual é redundante, e apenas reflete certos aspectos da estrutura sintática da sentença. Por exemplo, marcação de um substantivo como plural não é previsível com base na estrutura sintática, enquanto a marcação de plural do verbo finito o é, uma vez o sujeito identificado⁴. (Booij, 1993, p. 30, tradução nossa)

Na esteira dessas reflexões, em grande parte impulsionadas pelas conceituações de Booij, vemos subsequentes esforços para se classificarem os conteúdos gramaticais, com ênfase na veiculação destes por parte de material morfológico. Spencer (2000), ao chamar os significados gramaticais de *características-f*, afirma que cada língua tem seu inventário de características-f, com dois papéis principais: i. regular a estrutura sintática, como os fenômenos de concordância e regência; e ii. veicular certos significados *funcionais* como tempo, definitude e plural. Anna Kibort (2008) propõe uma tipologia das características gramaticais, dividindo-as em *morfossemânticas* – que codificam uma série de distinções semânticas, mas não são requeridas pela sintaxe, pelos mecanismos de concordância ou regência, a exemplo da maior parte das instâncias do tempo verbal; e *morfossintáticas* – isto é, determinadas pelo contexto sintático, como concordância de gênero ou classe nominal, ou de número. No entanto, a autora aduz que

palavras flexionadas mostram variação na forma. As diferentes formas estão correlacionadas com significados ou funções que rotulamos como “características”. No entanto, nem todas as características identificadas através da morfologia *flexional* são morfossintáticas⁵. (Kibort, 2008, tradução e grifos nossos)

As línguas, portanto, variam não só na estratégia escolhida para veicular significados gramaticais – ordem de constituintes, mudanças tonais, morfologia – como também, em as exprimindo morfológicamente, variam em quais categorias gramaticais serão necessariamente explicitadas – em especial, no grau de redundância sintagmática. Vê-se que a discussão, porém, muitas vezes se atrelou à *flexão* como expressão das categorias gramaticais, sejam inerentes ou

⁴ “[...] there is a clear difference in semantic content between inherent inflection and contextual inflection. Inherent inflection expresses, like derivation, a certain amount of independent information, whereas the information expressed by contextual inflection is redundant, and only reflects certain aspects of the syntactic structure of the sentence.’ For instance, the marking of a noun as plural is not predictable on the basis of syntactic structure, whereas the plural marking of the finite verb is, once the subject has been identified.”

⁵ “Inflected words show variation in form. The different forms are correlated with meanings or functions which we label as ‘features’. However, not all features that are identified through inflectional morphology are morphosyntactic.”

determinadas sintaticamente, com constante alusão ao conceito translinguisticamente problemático de *palavra* e suas consequências, como dificuldades de delimitação entre morfologia e sintaxe – a divisão entre os níveis de análise é feita com base na palavra, e a palavra é definida com base na divisão entre os níveis de análise, o que é um raciocínio circular (Mattos e Lazzarini-Cyrino, 2021). Há ainda entendimentos minoritários como o de Alain Kihm (2003, p. 335), que afirma que “[m]orfologia flexional é a realização explícita de elementos funcionais”; é dizer, *flexão* não se limitaria a manifestações no nível da *palavra*, não necessitando ser morfologia presa.

De toda sorte, é sabido que há línguas sem processos usualmente reconhecidos como *flexionais*, isto é, línguas nas quais os significados gramaticais são veiculados por *palavras de função* ou *partículas fonologicamente independentes*. A linguística, por tradição, classificou tais línguas holisticamente como *isolantes* ou *analíticas*, definidas com base numa razão baixa de morfemas por palavra, e ausência de processos flexionais (Schwegler, 1990; Mattos e Lazzarini-Cyrino, 2021), a exemplo do vietnamita:

(4) Vietnamita (Vietnã: austroasiática)

nó đã sửa xe rồi.

3SG PFV consertar carro já

“Ele já consertou o carro.” (Phan e Duffield, 2019)

O vietnamita realiza o perfectivo, um valor da categoria gramatical *aspecto*, por meio de um morfema fonologicamente independente, *đã*, comumente chamada de *partícula*. O aspecto, no entanto, é uma categoria gramatical inerente, não ditada pelo contexto sintático; da observação de sentenças de línguas como o vietnamita, o chinês ou o iorubá, vê-se que tais línguas, quando veiculam significados gramaticais, tendem a se limitar aos significados inerentes, não apresentando os contextuais (McWhorter, 2016). De fato, as mais conhecidas entre as línguas desse perfil, cuja principal estratégia para veicular categorias gramaticais é usar morfemas livres, são notáveis por não apresentar morfologia gramatical sintaticamente determinada ou redundante; mas essa observação se estende para as demais línguas com morfologia majoritariamente livre num geral? Isto é, caso uma língua possua alta porcentagem de morfemas gramaticais livres, ela necessariamente apresentará baixos níveis de morfologia gramatical contextual?

Diante de tudo quanto exposto, o presente trabalho, além de fornecer um excuroso

organizador a respeito das diferentes maneiras de expressão morfológica de significados gramaticais que as línguas apresentam, também examinou a correlação entre proporção de morfologia gramatical livre e a proporção de morfemas gramaticais veiculadores de informação contextual, de modo a investigar se há ou não uma relação entre o perfil morfossintático de uma língua e a expressão de categorias determinadas pela sintaxe. Para tanto, utilizamos uma amostra de 19 línguas de diferentes contextos geográficos e filogenéticos; cada língua teve um número de sentenças selecionadas, cujos morfemas foram inventariados, classificados por frequência e analisados quanto a serem formas presas ou livres e mais gramaticais ou mais lexicais, e, sendo gramaticais, quanto a veicularem informação determinada pelo contexto sintático ou não. Descobrimos haver uma correlação negativa entre a proporção de morfemas gramaticais livres e a proporção de morfemas veiculadores de informação gramatical determinada pelo contexto sintático. Apesar disso, ao lado de línguas com alta proporção de morfologia livre e sem morfologia gramatical contextual, como o mandarim e o káo, observamos algumas línguas com alta proporção de morfologia livre, mas com presença de morfemas contextuais, como o daai chin.

2. Expressão morfológica de categorias gramaticais

Não surpreendentemente, mesmo ao se limitar o estudo dos significados gramaticais à expressão morfológica, percebe-se como as línguas do mundo variam enormemente tanto no que devem exprimir para formar uma sentença bem feita, quanto no grau de especificação e detalhamento do que exprimem. Uma das propostas de como avaliar a forma pela qual as línguas utilizam – e se utilizam de – sua morfologia para veicular significados gramaticais foi proposta por Östen Dahl (2004): o conceito de *verbosidade* como a medida em que as construções das línguas apresentam acréscimo de material morfológico em relação à mera concatenação de elementos.

Em suas palavras:

[...] em inglês, palavras quantificadoras e substantivos incontáveis podem ser unidos por simples justaposição, como em *much snow* ou *little snow*. Em francês, por outro lado, uma palavra como *beaucoup* "muito" não pode simplesmente ser justaposta ao substantivo *neige* "neve"; em vez disso, deve-se inserir a preposição *de* entre os dois: *beaucoup de neige* "muita neve" [...] Do ponto de vista do inglês, porém, o elemento *de* parece redundante. Podemos dizer que as construções quantificadoras francesas são caracterizadas por verbosidade, definidas como tendo maior peso fonético (contendo mais material) do que seria minimamente necessário (logo, um caso especial de redundância)⁶. (2004, p. 53, tradução nossa)

⁶ “[...] in English, quantifier words and mass nouns can be joined by simple juxtaposition, as in *much snow* or *little snow*. In French, on the other hand, a word like *beaucoup* ‘much’ cannot simply be juxtaposed with the noun *neige*

Dahl define concisamente a verbosidade como sendo a “dispensabilidade através das línguas” (*cross-linguistic dispensability*). Usando o exemplo do sintagma em inglês *one dog*, podemos compará-lo a seu equivalente em mandarim. Temos:

(5) Mandarim (China: sino-tibetana)

— 只 狗

yī zhī gǒu

um CL cachorro

“Um cachorro”

Tanto *one dog* quanto *yī zhī gǒu* desviam-se da concatenação irrestrita (simples justaposição de seus elementos) por terem uma ordem fixa, não livre; o numeral precede o substantivo quantificado. No entanto, a expressão em mandarim possui um elemento a mais em relação à mesma expressão em inglês, o classificador nominal *zhī*. Isso faz com que a expressão em mandarim seja mais verbosa *sensu* Dahl que sua equivalente inglesa.

Outro conceito de interesse é o de *sobrespecificação* (*overspecification*, no original), proposto por John McWhorter (2007). Sobrespecificação se refere ao grau de especificação para além do estritamente necessário para a comunicação que uma língua usa para expressar uma categoria, como marcação de classe nominal, marcação numero-pessoal, marcação obviativa, posse inalienável, marcação de definitude, marcação mais minuciosa de tempo-aspecto-modo, marcação evidencial, presença de cópulas ou maior número de diferentes cópulas, distintas formas de negação, gradação mais minuciosa de demonstrativos, ergatividade, ajuste de valência, presença de marcação de foco, presença de partículas pragmáticas e relativização distinta de subordinação.

2.1 Morfemas gramaticais: afixos e partículas

Para o presente estudo, interessou como os diferentes morfemas veiculadores de significados gramaticais se apresentam. Dificulta bastante, porém, a confusão terminológica

‘snow’; rather, you must insert a preposition de in between: beaucoup de neige ‘much snow’ [...] From the point of view of English, however, the element de appears redundant. We may say that French quantifier constructions are characterized by verbosity, defined as having a larger phonetic weight (containing more material) than would be minimally necessary (thus a special case of redundancy).’

presente na literatura, sendo classificação dos morfemas muitas vezes feita com base em propriedades formais e semânticas ao mesmo tempo. Como bem colocado por William Croft (2000), em verdade, a divisão envolve quatro denominações muito utilizadas – raiz, afixo, morfema gramatical ou de função e morfema lexical ou de conteúdo –, e que não raro se confundem. Por exemplo:

[A] tradicional distinção entre palavras de conteúdo e palavras de função é uma em que palavras com o mesmo conteúdo geralmente têm formas livres não-marcadas através das línguas, enquanto palavras de função são frequentemente melhor traduzidas por morfemas presos em outras línguas. Também, diacronicamente, morfemas de conteúdo tendem a se manter livres, enquanto morfemas de função não raro migram entre formas presas (sufixais), semi-livres (clíticos) e livres⁷. (Kornai, 2020, p. 131. tradução nossa)

Igualmente:

Um morfema lexical é um morfema que pode ocorrer como uma palavra fonológica por si só, isso é, sem estar preso a um ou mais outros morfemas. *Morfemas gramaticais não podem ocorrer como palavras fonológicas por si só, mas são obrigatoriamente presos a outros morfemas*⁸. (Jagersma, 2010, p. 77, tradução e grifos nossos)

Nos exemplos trazidos acima, tanto András Kornai quanto Abraham Jagersma utilizam um conceito formal – de morfema preso ou livre – para definir um conceito semântico – o de morfema gramatical ou lexical. Já Haspelmath e Sims afirmam que “enquanto a distinção entre raízes por um lado e afixos por outro é em si bastante útil, esses conceitos acabam por ser mais complicados do que a imagem simples que vimos até então⁹” (2010, p. 22, tradução nossa), e que “um afixo se prende a uma palavra ou à parte principal de uma palavra. Geralmente tem um significado abstrato¹⁰” (ibid., p. 19). É frequente assumir que formas presas têm significados mais gramaticais ou “abstratos”, e formas livres, significados mais lexicais, como vemos em Velupillai:

Em inglês, raízes e radicais tendem a ser livres, mas há muitas línguas em que são presos [...] A raiz e o radical carregam informação lexêmica, isto é, a informação semântica básica

⁷ “That said, the traditional distinction between content words and function words is one where words with the same content generally have unmarked free forms across languages, while function words are often best translated by bound morphemes in other languages. Also, diachronically, content morphemes tend to stay free, while function morphemes often migrate between bound (suffixal), semi-free (clitic), and free forms.”

⁸ “A lexical morpheme is a morpheme that can occur as a phonological word in its own right, that is, without being bound to one or more other morphemes. Grammatical morphemes cannot occur as phonological words in their own right but are obligatorily bound to other morphemes.”

⁹ “While the distinction between roots on the one hand and affixes on the other is by itself quite useful, these concepts turn out to be more complicated than the simple picture that we have seen so far.”

¹⁰ “An affix attaches to a word or a main part of a word. It usually has an abstract meaning [...]”

da palavra [...] um afixo, por outro lado, é um morfema obrigatoriamente preso *que não carrega informação lexêmica alguma*.¹¹ (2012, p. 90, tradução e grifos nossos)

No capítulo 27, *Lexical and grammatical meaning*, do livro *Morphology: A handbook on inflection and word formation* (2000), Croft expõe de maneira clara os problemas de se confundirem tais conceitos, mostrando que a caracterização dos morfemas de uma língua pode se dar em dois diferentes planos, o formal e o semântico. A distinção raiz *versus* afixo, portanto, pertenceria a uma caracterização no plano formal, e a morfema gramatical *versus* lexical, a uma caracterização no plano semântico. Muitas vezes, há uma correlação direta entre os dois planos, e por isso é prontamente discernível se um morfema é lexical ou gramatical, se é uma raiz ou um afixo – além de haver frequente correspondência entre raízes e morfemas lexicais, e afixos e morfemas gramaticais. Veja-se por exemplo o caso de uma expressão mapudungun:

(6) Mapudungun (Chile e Argentina: araucana)

tripa-ke-rke-fu-y-ng-ün

sair-CF-REP-IPD-IND-3NS-PL

“(...) [eles] saíam (...)” (Smeets, 2008)

É relativamente descomplicado, nesse exemplo, apontar que a raiz é *tripa* “sair”, e os morfemas *-ke* (característica frequente), *-rke* (relatado), *-fu* (impeditivo), *-y* (indicativo), *-ng* (terceira pessoa, não singular) e *-ün* (plural), são afixos. Igualmente, pode-se afirmar sem maiores dificuldades que *tripa* é um morfema lexical, e os demais componentes, morfemas gramaticais. No entanto, a situação nem sempre é tão clara:

(7) a. Djabugay (Austrália: pama-nyungana) b. Mam (México/Guatemala: maia)

gulu-nda

t-ee

ele-DAT

3SG-para

“Para ele” (Croft, 2000)

“Para ele” (ibid.)

No exemplo do djabugay, a raiz é o morfema *gulu* “ele”, à qual se sufixa *-nda* DATIVO. No entanto, a mesma construção em mam apresenta a situação inversa: a raiz é *ee*, “para”, à qual se

¹¹ “In English roots and stems tend to be free, but there are many languages where they are bound [...] The root and stem carry lexemic information, i.e. the basic semantic information of the word [...] An affix, on the other hand, is an obligatorily bound morpheme which does not carry any lexemic information.”

prefixa *t-*, terceira pessoa do singular. Uma vez que os dois planos são ortogonais, a caracterização de um dado morfema se daria, sob a acepção de Croft, com base nas duas dimensões, a formal e a semântica. Dessa forma, assim como existem morfemas gramaticais afixais – a exemplo dos morfemas presos a *tripa* “sair” em (6) –, também existem morfemas gramaticais livres – frequentemente chamados na literatura de *partículas*, *palavras de função* ou *raízes vazias*, como o *le* no exemplo a seguir do mandarim:

(8) Mandarim

事情	已经	过去	了
<i>shìqíng</i>	<i>yǐjīng</i>	<i>guòqù</i>	<i>le</i>
assunto	já	passado	PFV

“(o assunto) são águas passadas.”

Nem todos os linguistas, porém, aceitam que partículas sejam tidas por raízes, por considerarem *raiz* como algo a que se aplicam flexões (Croft, 2000) – um reflexo de uma tipologia balizada no padrão encontrado na maioria das línguas europeias. Mas mesmo sob essa acepção há raízes de conteúdo gramatical; é esse o caso, por exemplo, dos auxiliares do kanoê, aos quais se prefixam vários morfemas também gramaticais, como desinências de pessoa e modo declarativo:

(9) Kanoê (Brasil: isolada)

<i>nukũ</i>	<i>aj</i>	<i>ry</i>	<i>õ-e-re</i>
caba	1SG	ferroar	1-DECL-AUX

"A caba me ferroou." (Bacelar, 2004)

Já casos de morfemas lexicais livres e/ou com status de raiz, à semelhança de morfemas gramaticais afixais, são abundantes e intuitivos para boa parte dos ocidentais; é o exemplo das raízes em línguas indoeuropeias, ou as *palavras lexicais* (*palavras de conteúdo* ou *raízes cheias*) das línguas siníticas – no exemplo em (8), *shìqíng* “assunto”. A última combinação, porém, de morfemas lexicais presos, sem status de raiz, embora menos familiar para falantes de línguas indoeuropeias, mostra que a correspondência direta entre um extremo semântico e o formal não é regra. Veja-se:

(10) Äiwoo (Ilhas Salomão: oceânica)

pelivano sime de-ku-luwa-kä-ji

crianças.3MIN pessoa coisa-IPFV-tomar.O-DIR:3-1+2MIN

“As crianças que adotamos” (Næss, 2017)

Em äiwoo, por exemplo, existem os ditos *substantivos genéricos presos*, que, como a denominação indica, não ocorrem de forma independente, e sim presos a outras raízes; essas formas têm correspondentes livres com o mesmo significado¹². Assim, para *de-* “coisa”, presente em (10), tem-se a forma independente *jelâ* “coisa”. É o mesmo caso de *gi-* “homem”, forma presa, e *sigiläi*, “homem”, forma livre; *nye-/nyi-* “lugar”, forma presa, e *numa* “lugar”, forma livre; e *si-* “mulher”, forma presa, e *singedâ*, “mulher”, forma livre.

Ademais, não só a correspondência entre os dois planos nem sempre é direta, como também a classificação de um dado morfema em um só plano já oferece desafios. Em se tratando de uma caracterização no plano semântico, a existência de processos de gramaticalização dificulta qualquer classificação dicotômica entre gramatical ou lexical – uma vez que na prática essa divisão não possui contornos nítidos e abruptos. Em vez de uma distinção binária e absoluta, a proposta de Croft é de que cada morfema ocuparia uma posição num *continuum* entre puramente lexical e puramente gramatical, com muitos casos intermediários entre os dois polos. Diacronicamente, os morfemas se moveriam entre esses extremos, de acordo com processos de relexificação ou gramaticalização, tornando a classificação algo dinâmico e não tão simples. O mesmo poderia ser dito de caracterizações no plano formal, com formas livres e presas nos extremos coexistindo com formas intermediárias, como os clíticos¹³. Um caso de gramaticalização em português em que se podem apontar formas intermediárias no *continuum* lexical/gramatical é o do verbo *estar*. Originalmente um verbo pleno, com o significado claramente lexical de *estar de pé*, tal item figura em construções como verbos de ligação – “ele (es)tá bem” – ou auxiliares – “ele (es)tá escrevendo”, contextos gramaticalizados e com possível redução de material fonológico

¹² Para Næss, os substantivos genéricos presos são “particularmente difíceis de analisar em termos de categorias morfossintáticas estabelecidas, como prefixos nominalizadores, compostos, classificadores ou núcleos de orações relativas; parecem compartilhar propriedades de todas elas, embora não se enquadrem claramente em nenhuma delas”, “[...] particularly difficult to analyse in terms of established morphosyntactic categories such as nominalising prefixes, compounds, classifiers, or heads of relative clauses; it seems to share properties of all of these, while not falling clearly into any of them” no original (2017, p. 916).

¹³ Remetemos o leitor à leitura de um dos textos seminais de morfologia em português, escrito por Mattoso Câmara Jr. (1967), no qual ele também aborda a distinção entre afixos, clíticos e partículas sob a terminologia *formas presas*, *formas dependentes* e *formas livres*.

representada pela supressão da primeira sílaba. Para além, também se apontam instâncias ainda mais gramaticais, a exemplo do uso da forma reduzida “tá” como advérbio de afirmação ou marcador discursivo – oriundos respectivamente das construções *tá certo* e *tá entendendo?* (Santos, 2016; Silva, 2021).

3. Categorias gramaticais contextuais

No *Yearbook of Morphology* de 1993, num capítulo intitulado *Against split morphology*, Geert Booij aborda a distinção entre flexão e derivação para a morfologia – ponto de vista conhecido como teoria da morfologia cindida (*split morphology* em inglês). Inicialmente citando Stephen Anderson (1982), para quem a morfologia flexional seria aquela “relevante à sintaxe”, Booij argumenta contra a dicotomia flexão/derivação em duas categorias discretas e apartadas uma da outra; seu principal ponto é o fato de algumas formas de flexão serem capazes de alimentar processos formadores de palavras – logo, seriam mais próximas da morfologia tida por “derivacional” que os demais componentes flexionais¹⁴.

O próprio Anderson, em 1988, havia oferecido uma classificação das categorias flexionais em quatro tipos: (i) *categorias configuracionais* (como marcação de caso em substantivos), atribuídas com base na estrutura maior na qual a palavra ocorre; (ii) *categorias de concordância* (aspectos da forma de uma palavra determinados por outras palavras na mesma estrutura); (iii) *categorias sintagmáticas* (como marcação de caso num sintagma nominal realizada no núcleo nominal); e (iv) *categorias inerentes* (como número em substantivos). Booij (1993) alude a esta análise e elabora sua própria classificação; para o autor, as categorias (i-iii) são *flexão contextual* e a categoria (iv), *flexão inerente*. A morfologia flexional inerente seria a forma de flexão mais próxima do que, considerando a morfologia cindida, se chama de derivação; sob esta categoria, estão propriedades inerentes tais qual número, para substantivos, e tempo/aspecto, para verbos.

Muito embora Booij tenha se concentrado no holandês na primeira publicação sobre o tópico, outras línguas serviram de exemplo e argumentação em seus subsequentes trabalhos (1996), bem como de diversos outros autores que adotaram ou trabalharam com a distinção flexão inerente *versus* flexão contextual (Kihm, 2003; Luís, 2010; Haspelmath e Sims, 2013; McWhorter, 2016, 2019; Bisang, 2020; Gardani, 2014, 2018, 2020). A flexão inerente, portanto, compreende a realização de categorias que, embora possam ser relevantes à sintaxe, veiculam

¹⁴Gonçalves (2011) e Haspelmath (2024) apresentam discussões interessantes acerca da problemática *flexão x derivação*, o primeiro em português.

informação independente, como os já citados número em substantivos e tempo/aspecto verbal, mas também a exemplo de casos gramaticais portadores de informação semântica, não ditada pelo contexto sintático. Servem de exemplo o locativo (ex.: turco *ev-de* [casa-LOC] "na casa"), o ablativo (ex.: quéchua de Huallaga *mayu-pita* [rio-ABL] "[vindo] do rio") e o instrumental (ex.: russo *nož-om* [faca-INS] "com uma faca"). Já a flexão contextual consistiria em categorias dependentes do contexto sintático, a exemplo dos casos estruturais – como nominativo e acusativo –, e demais morfemas requeridos em regência ou concordância, exprimindo informação em grande parte redundante (Haspelmath e Sims, 2013).

3.1 Informação gramatical contextual para além da flexão

Vimos na seção 2 que *afixos*, *clíticos* e *partículas* são categorias formais, e *morfemas gramaticais* e *morfemas lexicais* são categorias semânticas, não se confundindo e sim se relacionando de maneira ortogonal. Também na seção supracitada, demonstramos que *afixos* podem ser gramaticais ou lexicais e que categorias gramaticais podem ser veiculadas tanto por morfologia presa, quanto por morfologia livre. Quanto à morfologia gramatical sintaticamente determinada ou redundante, porém, sua expressão em diferentes línguas não raro foi associada a processos de flexão.

Nas palavras de Anna Kibort:

A especificação de uma característica inerente pode ser pensada como exprimindo uma informação que logicamente pertence ao – ou surge a partir do – interior do elemento no qual ela se encontra, enquanto a especificação de uma característica contextual pode ser pensada como exprimindo uma informação que se origina logicamente fora do elemento na qual ela se encontra (na concordância, chamamos essa informação de “deslocada” [*displaced*], e na regência, de “marcação” [*brand mark*]). Então, características encontradas em controladores de concordância são características inerentes, enquanto características encontradas em alvos de concordância e em [elementos] regidos são contextuais.¹⁵ (2008, sem página, tradução nossa)

Como ilustração, veja-se:

¹⁵ “*Inherent feature specification can be thought of as expressing information that logically belongs to, or arises from within the element on which it is found, while contextual feature specification can be thought of as expressing information that logically originates outside the element on which it is found (in agreement, we call this information 'displaced', and in government, we can refer to this information as a 'brand mark'). Thus, features found on controllers of agreement are inherent features, while features found on agreement targets and on governees are contextual features.*”

(11) Yakkha (Nepal: sino-tibetana)

ka *kheʔ-ma* *mit-a-ŋ-na*
 1SG ir-INF pensar-PST-1SG-NMLZ.SG

“Eu quero ir.” (Schackow, 2015)

Em (11), o morfema *-ŋ*, realizador de concordância verbal de primeira pessoa do singular, é um exemplo de flexão contextual, por exprimir uma informação que se origina logicamente fora do elemento – o verbo – em que ele se encontra; no caso, a origem é o constituinte *ka*, “eu”. Esse é um caso de morfologia gramatical contextual presa, que tradicionalmente se chamou de *flexão contextual*. No entanto, vejamos (12):

(12) Daai chin (Myanmar: sino-tibetana)

mhnam pa: *noh* *aksük ksük* *a* *thi:ngj nu:* *ah*
 deus.criador ERG primeiro CFIN árvore grande S.AGR:3s
msom ma
 criar fazer.primeiro

“O deus criador antes de tudo criou uma grande árvore.” (So-Hartmann, 2009)

A sentença daai chin em (12) apresenta uma série de morfemas veiculadores de informação gramatical contextual, como *noh* (ergativo) e *ah* (concordância de sujeito de terceira pessoa do singular); no entanto, não são afixos, e sim partículas livres. Da mesma forma:

(13) Rapa Nui (Ilha de Páscoa: austronésia)

ina *ko* *kai* *i* *te* *kai* *mata*
 NEG NEG.IPFV comer ACC ART comida crua

“Não coma comida crua.” (Kieviet, 2017)

O morfema *i* (acusativo) do rapa nui, visto na sentença acima, veicula uma categoria gramatical determinada pela sintaxe – o caso nominal de natureza estrutural, frequentemente redundante por existir também como mecanismo de expressão a ordem dos constituintes. O morfema analisado, porém, também é uma partícula – um morfema livre, e não um afixo.

4. Medidas de morfologia gramatical determinada pelo contexto sintático

Para investigar a relação entre morfologia gramatical livre e morfologia gramatical determinada pelo contexto sintático, executamos um estudo quali-quantitativo. Gramáticas descritivas foram usadas como fontes de dados a serem inseridos manualmente num banco próprio, de acordo com o método desenvolvido e exposto em Lazzarini-Cyrino e Mattos (2020). O banco de dados preserva a estrutura das gramáticas descritivas: i. segmentação morfológica; ii. glosas; e iii. traduções, nos interessando sobretudo a segunda linha. Um programa foi feito especificamente para contabilizar os elementos das glosas, ordenando-os em listas de frequência de significados para cada uma das línguas – os morfemas mais frequentes no topo e os menos frequentes no fim. Dessas listas de morfemas, calculamos os valores correspondentes a 20% dos tipos de morfema para cada língua, amostrando esse valor a partir dos morfemas do topo da lista.

A separação de morfemas gramaticais dos lexicais seguiu a proposta de William Croft (2000), triando-os individualmente de acordo com *continua* em quatro dimensões – natureza enciclopédica, natureza generalista, conceptualização do conteúdo da experiência e domínio semântico –, nos quais um extremo é mais lexical e o outro, mais gramatical. Utilizamos cinco valores para cada dimensão: os dois polos – *L*, mais lexical, e *G*, mais gramatical –, bem como um valor intermediário *N*, e gradações entre o valor intermediário e o extremo gramatical – *Ng* – e entre o valor intermediário e o extremo lexical – *Nl*. Selecionamos apenas os morfemas que receberam *G* e *Ng*, descartando os demais; assim, obtivemos listas com os morfemas mais gramaticais para as 19 línguas escolhidas. As três primeiras colunas da tabela 1 ilustram o resultado do procedimento:

Tabela 1 – Três dos morfemas mais frequentes do araweté e suas respectivas classificações nas quatro dimensões

Morfema	Freq.	NE	NG	CCE	DS
1	259	G	G	Ng	G
R1	258	G	G	G	G
FOC	243	G	G	G	G

Cada morfema amostrado das 19 línguas utilizadas foi classificado de acordo com dois parâmetros: i. se livre ou preso; e ii. sua determinação ou não pelo contexto sintático – isto é, se

fazia parte da morfologia gramatical contextual ou não.

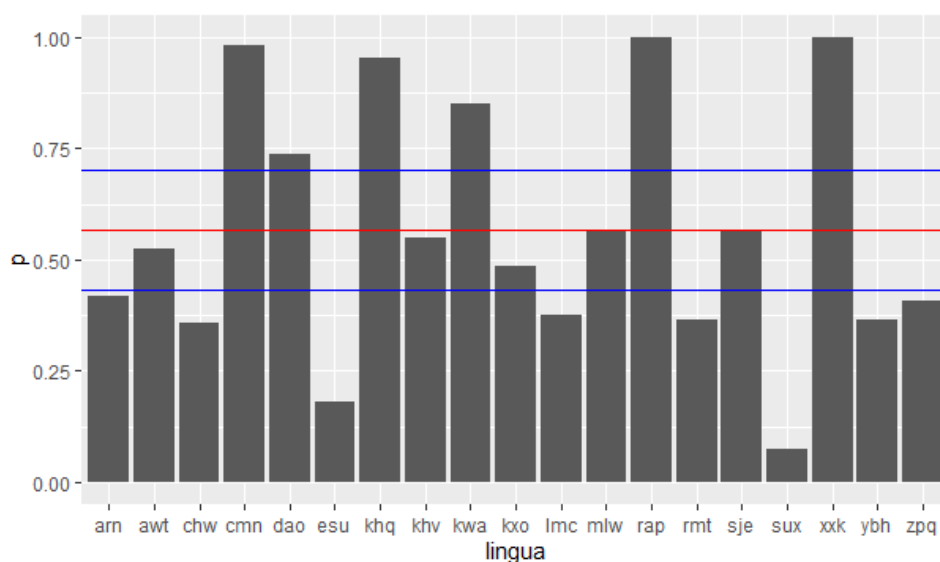
Tabela 2 – Línguas utilizadas, número aproximado de falantes, número de morfemas no nosso banco de dados, 20% dos morfemas -arredondados para cima quando necessário

Língua	Local: afiliação	N. falantes	N. morfemas	20%
Araweté	Brasil: tupiana	200	420	84
Cuwabo	Moçambique: bantu	800.000	595	119
Daai chin	Myanmar: sino-tibetana	252	1094	219
Dâw	Brasil: nadahup	174	253	51
Domari	Oriente Médio: indoeuropeia	161	437	87
Kanoê	Brasil: isolada	138	543	109
Khwarshi	Cáucaso: nakh- daguestaniana	130	1468	293
Koyra chiini	Mali: nilo-saariana	200.000	337	67
Kéo	Ilha de Flores: austronésia	40.000	397	79
Limilngan	Austrália: línguas da região de Darwin	Extinta	295	59
Mandarim	China: sino-tibetana	921.5 milhões	720	144
Mapudungun	Chile: araucana	260.000	430	86
Moloko	Camarões: afroasiática	8.500	299	60
Saami de Pite	Noruega/Suécia: urálica	20	568	114
Rapa nui	Ilha de Páscoa: austronésia	1.000	488	98
Sumério	Mesopotâmia: isolada	Extinta	668	134
Yakkha	Nepal: sino-tibetana	20.000	1027	205
Yupik do Alasca central	Alasca: esquimó-aleuta	19.750	543	109
Zapoteca de San Bartolomé Zogoocho	México: oto- mangueana	1.400	435	87

4.1 Proporção de morfemas gramaticais livres

Para o primeiro parâmetro, proporção de morfemas gramaticais livres, avaliamos se, nas gramáticas descritivas, o morfema em questão vinha glosado como livre (recebendo um S) ou preso a outros morfemas (recebendo um N). As 19 línguas amostradas tiveram proporções diferentes de morfemas gramaticais livres, conforme calculado por frequência relativa e demonstrado na forma gráfica na figura 1:

Figura 1 – Proporção de morfemas gramaticais livres nas diferentes línguas, com média das proporções (em vermelho) e o intervalo de confiança de 95% (em azul)



Fonte: autoria própria

As línguas que apresentaram proporção estatisticamente significativa (acima da linha azul superior) de morfemas gramaticais livres – mandarim (cmn, com 98,3%), daai chin (dao, com 73,7%), koyra chiini (khq, com 95,2%), dâw (kwa, com 85%), rapa nui (rap, com 100%) e kéó (xxk, com 100%), são descritas na literatura como sendo *isolantes* (So-Hartmann, 2009; Delmer, 2016; Kieviet, 2017; McWhorter 2016, 2019; Obert, 2020). Já o sumério (sum), o yupik do Alasca central (esu), o cuwabo (chw), o domari (rmt) e o yakkha (ybh) foram as línguas com as menores proporções de morfologia gramatical livre, respectivamente com 7%, 18%, 35%, 36%, 36% e 37%.

4.2 Morfologia gramatical contextual

Para o segundo parâmetro, determinação ou não pelo contexto sintático, duas colunas

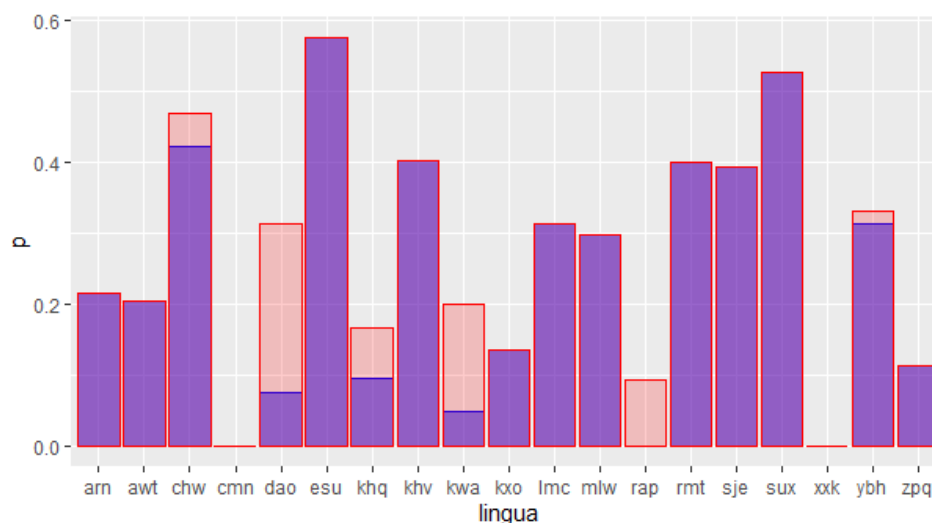
foram preenchidas; a primeira, FC, diz respeito à flexão contextual *sensu* Booij, mais restrita, logo, apenas os morfemas que são presos. A segunda coluna, CS, foi baseada num conceito mais amplo, selecionando quaisquer morfemas cuja informação gramatical veiculada era contextual – entrando nessa definição, portanto, partículas, clíticos, palavras independentes e também afixos. Como exemplo, mostramos alguns morfemas do daai chin na tabela 3:

Tabela 3 – Classificação de morfemas quanto a serem veiculadores de flexão contextual *sensu* Booij (FC) ou de informação gramatical contextual em geral (CS), exemplo do daai chin

Morfema	Freq.	FC	CS
NON.FUT	414	N	N
DEM	314	N	N
ERG	252	N	S
S.AGR:3S	174	N	S
PL	161	N	S
CF	138	N	S
LOC	130	N	N

Vemos que nenhum dos morfemas apresentados acima é contado como flexão contextual, uma vez que todos são trazidos pela gramática (So-Hartmann, 2009) como *partículas* ou *palavras independentes*. No entanto, os morfemas para marcação de caso ergativo ERG, concordância de sujeito de 3ª pessoa S.AGR:3S, plural verbal PL e final de constituinte CF, são veiculadores de informação gramatical determinada pelo contexto sintático, recebendo o valor S na coluna CS. Confrontamos, portanto, os valores mais estritos do que seria flexão contextual, doravante FC, e o parâmetro mais amplo que leva em conta simplesmente se o morfema veicula informação gramatical contextual, independentemente de ser um morfema preso (afixo, clítico) ou livre (partículas ou palavras de função), doravante CS.

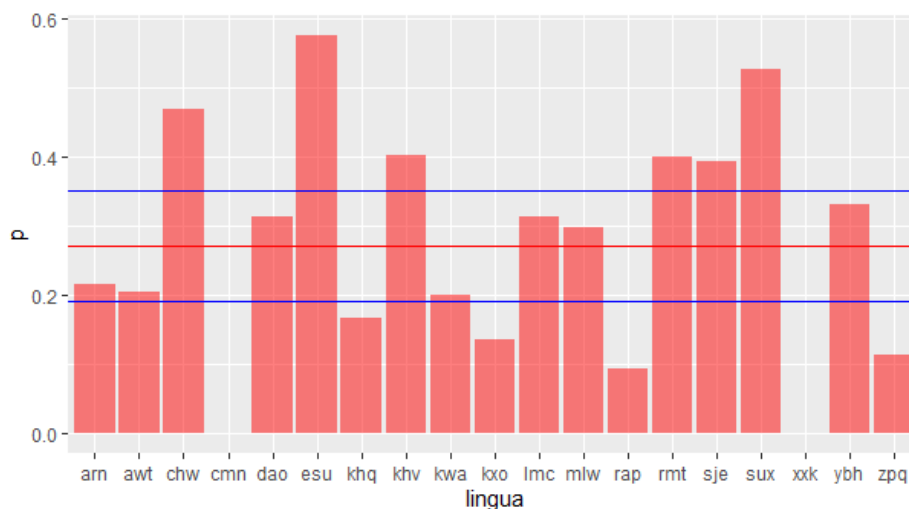
Figura 2 – Comparação entre flexão contextual *sensu stricto* (FC), em azul, e morfologia gramatical contextual (CS), em vermelho



Fonte: autoria própria

A figura 2 mostra a comparação gráfica dos dois parâmetros. CS revela-se, não surpreendentemente, mais inclusivo; todas as línguas que apresentam morfemas veiculadores de FC também apresentam CS, embora o contrário não seja verdadeiro – o rapa nui apenas apresenta morfemas gramaticais contextuais livres. Com a exceção do cuwabo (chw), que já apresentava uma proporção significativa de morfologia flexional contextual mesmo sob a acepção mais estrita (passou de 42% a 46,8%), todas as línguas que obtiveram aumento nas taxas de morfemas veiculadores de morfologia gramatical contextual ao se considerar o parâmetro mais amplo são línguas com altas taxas de morfologia gramatical livre – o daai chin (dao, de 7,5% a 31%), o koyra chiini (khq, de 9% a 16,6%), o dâw (kwa, de 5% a 20%) e o rapa nui (rap, de 0% a 9%).

Figura 3 – Proporção de morfemas veiculadores de informação gramatical determinada sintaticamente (CS). Em vermelho, a média da proporção; em azul, o intervalo de confiança



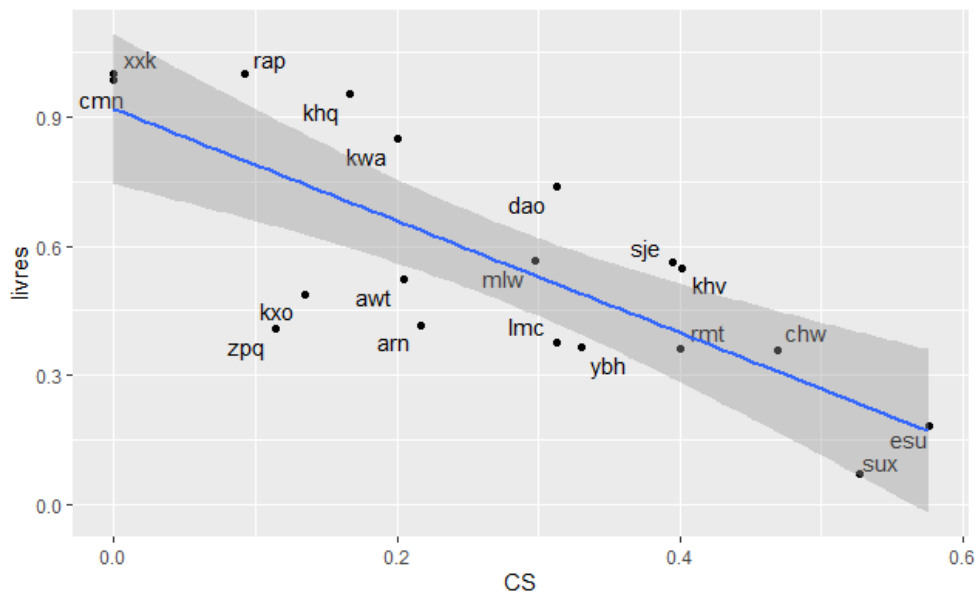
Fonte: autoria própria

Da análise da figura 3, depreende-se que seis das 19 línguas apresentaram proporção estatisticamente significativa (acima da linha azul superior) de morfologia flexional contextual – o cuwabo (chw), o yupik (esu), o khwarshi (khv), o domari (rmt), o saami de Pite (sje) e o sumério (sux). Das demais línguas, cinco são do grupo que possui alta proporção de morfemas gramaticais livres – koyra chiini (khq), mandarim (cmn), kéó (xxk) e dâw (kwa).

4.3 Correlação entre os dois parâmetros

A correlação entre morfologia gramatical livre e morfologia gramatical dependente de contexto sintático também pode ser visualizada na forma de um gráfico de reta (figura 4):

Figura 4 – Relação entre proporção de morfemas livres e proporção de morfemas veiculadores de flexão contextual (CS)



Fonte: autoria própria

Fica evidente que há uma correlação negativa entre a proporção de morfemas livres no inventário de morfemas gramaticais de uma língua e a proporção de morfemas gramaticais veiculadores de informação contextual. As duas línguas que mais apresentaram morfemas livres, mandarim (cmn) e kéio (xxk), foram as que menos apresentaram expressão de morfologia determinada pelo contexto sintático, por exemplo. Já as duas mais caracterizadas por terem alta proporção de morfologia contextual, yupik (esu, 57,5%) e sumério (sux, 52,7%) também apresentaram menor proporção de morfologia gramatical livre – isto é, maior proporção de morfologia gramatical presa. O yupik é descrito como sendo uma língua *polissintética* (Miyooka, 2012):

(14) Yupik do Alasca central (Alasca: esquimó-aleúta)

yu-urte-llr-an

allami-ku-ani

yu-urte-llru-unga

peessoa-virar-NMLZ-REL.3SG.SG

ano.passado-ASP-LOC.3SG.SG

peessoa-virar-PST-IND.1SG

“Eu nasci um ano depois de ele ter nascido.” (*ibid.*)

O sumério, tradicionalmente descrito como *aglutinante*, é uma língua extinta há cerca de 4.000 anos, falado no sul da mesopotâmia. Possui certa notoriedade por ter sido a primeira língua a ser registrada de que se tem conhecimento, e apresenta rica morfologia de casos estruturais sob

a forma de clíticos em posição final de sintagma:

(15) Sumério (Mesopotâmia: isolada)

ud inannak=e igi namtil=ak=ane=∅

dia Inanna=ERG olho vida=GEN=SEU=ABS

∅-mu-n-ši-n-bar-∅-ʔa=ʔa

VP-VENT-3SG-para-3SG.A-estar.fora-3NS/DO-NOM=LOC

“Quando Inanna o olhara com seus olhos provedores de vida.” (*ibid.*)

O *status* dos clíticos é um tópico controverso; embora necessariamente presos a algum hospedeiro, são menos seletivos quanto à natureza do elemento ao qual se combinam (Haspelmath e Sims, 2011), chegando a ser considerados semi-livres (Kornai, 2020), ou formas dependentes, na terminologia de Mattoso Câmara (1967). Na presente análise, consideramos clíticos como elementos presos. No entanto, a maior liberdade de combinação desses morfemas se comparados a afixos é algo a não ser desconsiderado, e o sumério representa uma língua com morfologia gramatical contextual grandemente expressa por meio de enclíticos, e não de palavras flexionadas em paradigmas.

Outras línguas que combinam alta proporção de morfemas veiculadores de flexão contextual com baixa proporção de morfemas livres são o cuwabo, língua bantu que, como os demais membros da família, é descrita como *aglutinante* e dispõe de um extenso sistema de classes nominais (Guérois, 2015), e o domari (40% de CS), língua indoeuropeia; diferentemente do sumério, em ambas, a morfologia gramatical contextual é basicamente expressa por meio de afixos, como no yupik.

5. Considerações finais

Após uma esquematização de como os polos *afixo* versus *palavra independente* e *gramatical* versus *lexical* são extremos de dois planos diferentes, um formal e um semântico, mostramos como a linguística ainda faz uma certa confusão terminológica e conceitual ao tratar da classificação dos morfemas. Ademais, nossa análise demonstrou haver uma correlação negativa entre a proporção de morfemas gramaticais livres – partículas, palavras de função, raízes vazias – e a proporção de morfologia determinada pelo contexto sintático, ou contextual; isto é, quanto

mais do inventário morfêmico de uma língua consistir em morfemas livres, menor será a tendência de esta língua exprimir morfologicamente categorias determinadas pela sintaxe. No entanto, a expressão morfológica dessas categorias gramaticais determinadas pela sintaxe não se limita aos processos ditos de *flexão*; atestamos aqui, mesmo com uma amostra não tão grande, a existência de línguas cujo inventário morfêmico gramatical é grandemente livre e que apresentam morfologia contextual, como é o caso do rapa nui e do daai chin, que possuem casos estruturais, e, esta última, concordância de número e pessoa, exprimidos por meio de partículas. Há, portanto, uma tendência à maior expressão morfológica de informação sintaticamente determinada por parte de línguas com morfologia gramatical presa, embora não seja uma regra absoluta.

Chama a atenção o caso do sumério, a segunda língua amostrada com maior proporção de morfologia gramatical contextual; diferentemente das demais línguas com alta proporção de morfologia contextual, o sumério a veicula em grande parte por meio de clíticos que se associam ao último elemento de um sintagma, não importando qual seja sua natureza. Vê-se, portanto, que a morfologia contextual não se restringe a expressões no nível da *palavra*, como os primeiros proponentes desse conceito poderiam fazer crer, uma vez que tradicionalmente associaram tal morfologia com processos de *flexão*. Independentemente do *status* dos clíticos, sua maior mobilidade e menor exigência de associação em relação a seu hospedeiro permitem aventar a possibilidade de, no mínimo, representarem uma posição de expressão da morfologia contextual intermediária em relação ao uso de afixos e ao uso de partículas. Mais pesquisas nesse caminho, com uma maior amostragem que inclua línguas cujas categorias gramaticais são grande ou majoritariamente expressas por clíticos, aparentam ser uma direção interessante.

Referências

- ANDERSON, S. R. Inflection. In M. Hammond and M. Noonan (eds.), *Theoretical Morphology. Approaches in Modern Linguistics*. San Diego etc.: Academic Press, 23-44. 1988.
- BAIRD, L. et. al. *A grammar of Keo: an Austronesian language of East Nusantara*. 2002.
- BACELAR, L. N. *Gramática da língua Kanoê*. Ph.D. dissertation, Katholieke Universiteit Nijmegen, 2004.
- BICKEL, Balthasar; NICHOLS, Johanna. Autotypologizing databases and their use in fieldwork. In: *Proceedings of the international LREC workshop on resources and tools in field linguistics*, Las Palmas. Nijmegen: MPI for Psycholinguistics, 2002.
- BISANG, W. Radical analyticity and radical pro-drop scenarios of diachronic change in east and mainland Southeast Asia, West Africa and pidgins and creoles. *Asian Languages and Linguistics*, v. 1, n. 1, p. 34–70, 2020.
- BOOIJ, G. E. Against split morphology. In: *Yearbook of morphology 1993*. Springer, Dordrecht, p.

27-49, 1993.

CROFT, W. Lexical and grammatical meaning. In G. BOOIJ, C. LEHMANN; J. MUGDAN (eds.), *Morphology: A handbook on inflection and word formation* (257–63). Berlin/ New York: Mouton de Gruyter, 2000.

DELMER, J. F. *Analytic Languages in Mainland South-East Asia*. Term Papers, MA Linguistics Program, Payap University, Chiang-Mai, Thailand, 2016.

DRYER, M. S.; HASPELMATH, M. (eds.). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. 2013. (Available online at <http://wals.info>, Accessed on 2021-03-10.)

GARDANI, F. *Borrowing of inflectional morphemes in language contact*. Peter Lang GmbH, Internationaler Verlag der Wissenschaften, 2014.

GARDANI, F. On morphological borrowing. *Language and Linguistics Compass*, v. 12, n. 10, p. e12302, 2018.

GARDANI, F. Morphology and contact-induced language change. In: *The Oxford handbook of language contact*, p. 96-122, 2020.

GONÇALVES, C. A. *Iniciação aos estudos morfológicos: flexão e derivação em português*. São Paulo: Contexto, 2011.

GUÉROIS, R. *A grammar of Cuwabo (Mozambique, Bantu P34)*. 2015. Tese de Doutorado. Université Lyon 2.

HAIG, G. Ergativity in Iranian. 2015. Disponível em: <https://www.academia.edu/15321950/Ergativity_in_Iranian>. Acesso em 20 de janeiro de 2024.

HASPELMATH, M; SIMS, A. *Understanding morphology*. 2ª edição. London: Hodder Education, 2010.

HASPELMATH, M. Inflection and derivation as traditional comparative concepts. *Linguistics*, v. 62, n. 1, p. 43-77, 2024.

JAGERSMA, B. *A descriptive grammar of Sumerian*. 2010. Tese de Doutorado. Leiden University.

KIBORT, A. A typology of grammatical features. *Grammatical Features*. 2008. Disponível em: <<http://www.grammaticalfeatures.net/inventory.html>>. Acesso em: 18 de agosto de 2022.

KIEVIET, P. *A grammar of Rapa Nui*. Berlin: Language Science Press, 2017.

KIHM, A. Inflectional categories in creole languages. In PLAG, I. (ed.): *Phonology and morphology of creole languages*. Tübingen: Niemeyer, p. 333-363, 2003.

KORNAI, A. Phenogrammar. In: *Semantics*. Nova York: Springer, Cham, 2020. p. 127-175.

LAZZARINI-CYRINO, J. P.; MATTOS, E. B. Um estudo exploratório sobre a classificação de morfemas por agrupamento hierárquico para comparação tipológica. *Natal: Revista do GELNE*, v. 22, n. 2, Dossiê Morfologia: p. 395-407 2020.

LUÍS, A. R. The loss and survival of inflectional morphology: Contextual vs. inherent inflection in creoles. In: *Romance Linguistics 2009*. John Benjamins, 2010. p. 323-336.

MATTOS, E.; LAZZARINI-CYRINO, J. P. Classificação morfológica das línguas: da tipologia holística às definições circulares. *Revista Diadorim*, v. 23, p. 288-312, 2021.

MATTOSO CÂMARA Jr., J. *Princípios de Linguística Geral*. 4a. ed. Rio de Janeiro: Livraria Acadêmica.

1967.

MCWHORTER, J. *Language interrupted: Signs of non-native acquisition in standard language grammars*. Oxford University Press on Demand, 2007.

MCWHORTER, J. Is radical analyticity normal? In: VAN GELDEREN, E. (ed.). *Cyclical change continued*, Amsterdam: John Benjamins Publishing Company, p. 227-249, 2016.

MCWHORTER, J. The radically isolating languages of Flores: A challenge to diachronic theory. *Journal of Historical Linguistics*, v. 9, n. 2, p. 177-207, 2019.

MIYAOKA, Osahito. *A grammar of Central Alaskan Yupik (CAY)*. Berlin: De Gruyter, 2012.

NÆSS, Å. Beyond roots and affixes: Äiwoo deverbal nominals and the typology of bound lexical morphemes. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"*, v. 41, n. 4, p. 914-955, 2017.

OBERT, K. Complex predicates and space in Dâw (Naduhup language, AM). *STUF-Language Typology and Universals*, v. 73, n. 1, p. 27-55, 2020.

PHAN, T.; DUFFIELD, N. 'To be tensed or not to be tensed?' The case of Vietnamese. *Investigationes Linguisticae*, v. 41, p. 105-125, 2019.

SANDE, Hannah Leigh. *Distributing morphologically conditioned phonology: Three case studies from Guébie*. University of California, Berkeley, 2017.

SANTOS, A. L. *Usos dos verbos ser e estar no português brasileiro: uma abordagem funcional*. 2016. Dissertação (Mestrado em Língua Portuguesa) – Pontifícia Universidade Católica de São Paulo, São Paulo, 2016.

SAPIR, E. *An introduction to the study of speech*. Nova York: Harcourt, Brace, 1921.

SCHACKOW, Diana. *A grammar of Yakkha*. Berlin: Language Science Press, 2015.

SCHWEGLER, A. *Analyticity and syntheticity: A diachronic perspective with special reference to Romance languages*. Berlin: Walter de Gruyter, 1990.

SILVA, V. M. *Usos de está/tá na fala popular soteropolitana: variação e gramaticalização*. 2021. Dissertação (Mestrado em Estudo de Linguagens) – Universidade do Estado da Bahia, Salvador, 2021.

SMEETS, I. *A grammar of Mapuche*. Berlin: De Gruyter Mouton, 2008.

SO-HARTMANN, H. *A descriptive grammar of Daai Chin*. Berkeley: Sino-Tibetan Etymological Dictionary and Thesaurus Project, 2009.

SOLANO, E. J. B. *Descrição gramatical da língua araweté*. Ph.D. dissertation, Universidade de Brasília, 2009.

SONNENSCHNEIDER, A. H. *A descriptive grammar of San Bartolomé Zoogocho Zapotec*. University of Southern California, 2004.

SPENCER, A. Inflection and the lexeme. *Acta Linguistica Hungarica*, v. 47, n. 1, p. 335-344, 2000.

TRUDGILL, P. Social structure, language contact and language change. In WODAK, R.; JOHNSTONE, B.; KERSWILL, P. E. *The SAGE handbook of sociolinguistics*. Sage Publications, pp. 236-248, 2011.

VELUPILLAI, V. *An Introduction to Linguistic Typology*. Amsterdam: John Benjamins, 2012.