

PRODUTIVIDADE CONSTRUCIONAL: O  
PAPEL DO *COVERAGE* MORFOFONOLÓGICO  
NA EXTENSIBILIDADE DE CATEGORIAS  
CONSTRUCIONAIS

CONSTRUCTIONAL PRODUCTIVITY: THE ROLE  
OF MORPHOPHONOLOGICAL COVERAGE IN THE  
EXTENSIBILITY OF CONSTRUCTIONAL CATEGORIES

Dennis de Oliveira Alves\*

Diogo Oliveira Ramires Pinheiro\*\*

**RESUMO:** Este artigo busca investigar a atuação do *coverage* morfofonológico sobre a extensibilidade de construções gramaticais para novos usos. Em poucas palavras, o *coverage* é entendido como o espaço categorial que inclui os usos já atestados para uma dada construção e os usos inovadores. Tomamos como ponto de partida o estudo de Suttle e Goldberg (2011), que atesta a relevância do *coverage* semântico pela combinação de efeitos de frequência de tipo e de variabilidade e similaridade semânticas. Neste estudo, porém, buscamos investigar se ocorre efeito de “coverage” associado às propriedades morfofonológicas dos itens. Para isso, desenvolvemos uma adaptação do experimento original, na qual foram incluídos verbos pertencentes às três conjugações verbais do português. Em síntese, o que os resultados demonstram é que, assim como o *coverage* semântico, também o *coverage* morfofonológico parece afetar a produtividade construcional.

**PALAVRAS-CHAVE:** Produtividade construcional. Extensibilidade. *Coverage* morfofonológico.

**ABSTRACT:** This paper aims to investigate the role of the morphophonological coverage in the extensibility of grammatical constructions to novel usages. Coverage is understood as the categorial space that includes attested usages of a particular grammatical construction as well as innovative ones. We take as a starting point the study by Suttle e Goldberg (2011), which demonstrates the relevance of coverage due to the

---

\* Doutorando e Mestre em Linguística pela Universidade Federal do Rio de Janeiro (UFRJ). E-mail: dennisalves@letras.ufrj.br, <https://orcid.org/0000-0003-3529-9623>.

\*\* Doutor em Linguística e Professor Adjunto na Universidade Federal do Rio de Janeiro (UFRJ). E-mail: diogopinheiro@letras.ufrj.br, <https://orcid.org/0000-0003-2403-504>.

combination of type frequency, semantic variability and semantic similarity effects. In this study, we aim to investigate if there a coverage effect can also be found when morphophonological features are taken into account. In order to achieve this goal, a new version of the original experiment was developed, in which verbs belonging to the three Portuguese verbal conjugations were included. In summary, the results show that, just like semantic coverage, morphophonological coverage also seems to affect constructional productivity.

**KEYWORDS:** Constructional productivity. Extensibility. Morphophonological coverage.

## INTRODUÇÃO

Frequentemente, construções gramaticais são usadas de modo inovador. Falantes produzem sentenças sem nunca as terem ouvido antes. Mas quais fatores são responsáveis por facilitar a extensão de uma construção gramatical para novos itens lexicais, resultando em enunciados inéditos? Por exemplo, o que leva o falante a usar um verbo intransitivo como “espirrar” em um padrão sintático “S V OD SPrep”, que veicula ideia de movimento causado, resultando em um uso inovador como “Ele espirrou o guardanapo para fora da mesa”?

Investigado na literatura construcionista a partir do conceito de *produtividade construcional*, esse problema tem atraído, em particular, a atenção de linguistas alinhados à Gramática de Construções Baseada no Uso (Goldberg, 2019; Suttle; Goldberg, 2011; Barðdal, 2008; Bybee; Eddington, 2006). Fundamentalmente, esse conjunto de trabalhos recorre a dois tipos de explicações para dar conta da possibilidade de extensão de uma construção gramatical para novos itens: aquelas baseadas no conhecimento estatístico, que se relaciona com a experiência concreta do falante com a língua, e aquelas baseadas no conhecimento gramatical, relacionado às propriedades gramaticais (fonológicas, morfológicas, sintáticas e semânticas) das construções.

No entanto, é possível observar que, entre os trabalhos voltados para o conhecimento gramatical, as generalizações de natureza semântica têm sido priorizadas, em detrimento de generalizações formais (Suttle; Goldberg, 2011; Boyd; Goldberg, 2011; Barðdal, 2008). Especificamente, esses trabalhos tendem a focalizar dois fatores: a coerência semântica verificável entre os itens que instanciam uma dada construção gramatical, isto é, a consistência semântica interna da categoria composta pelos itens que a instanciam, e a similaridade semântica entre esses mesmos itens e um novo item potencial para a construção, ou seja, o grau de semelhança entre o significado do item inovador na construção e o significado dos itens que já a instanciam. Os principais achados quanto ao papel do conhecimento gramatical no estudo da produtividade construcional estão descritos na seção 3 (Revisão da literatura).

Dada a solidez das descobertas referentes ao papel das generalizações semânticas sobre a produtividade construcional, este estudo parte do seguinte questionamento: seria

possível que as generalizações *morfofonológicas* implicitamente conhecidas pelo falante também atuassem, juntamente com o conhecimento semântico, na regulação da produtividade construcional? Em outras palavras, perguntamo-nos se a coerência morfofonológica de uma categoria de itens lexicais, bem como a similaridade morfofonológica desses itens com o item-alvo, também contribuiriam para a extensibilidade de uma construção gramatical para usos inéditos.

Para responder a essa pergunta, desenvolvemos, com base no paradigma elaborado por Suttle e Goldberg (2011), um experimento no qual os participantes deveriam avaliar sua confiança para estender uma construção gramatical para usos inéditos. Diferentemente de Suttle e Goldberg (2011), manipulamos os estímulos a fim de verificar o efeito de propriedades morfofonológicas dos itens (e não de seus atributos semânticos) sobre a produtividade construcional. Especificamente, tomamos como objeto gramatical verbos pertencentes às três conjugações do português, em sua forma infinitiva. Os detalhes dos procedimentos metodológicos serão explicitados na seção 4.

Este artigo está organizado da seguinte forma. Na próxima seção, dedicada aos pressupostos teóricos, apresentamos os princípios básicos da Gramática de Construções Baseada no Uso e discutimos o problema da produtividade construcional. Na seção 3, apresentamos uma revisão da literatura, descrevendo os resultados já obtidos em estudos sobre o fenômeno em foco. Em seguida, na seção 4, apresentamos a metodologia adotada para responder ao nosso questionamento. Na seção 5, reportamos os resultados obtidos e discutimos sua interpretação. Por fim, na seção 6, sintetizamos os principais achados e destacamos as contribuições deste trabalho.

## **PRESSUPOSTOS TEÓRICOS**

Nesta seção serão apresentados alguns princípios básicos da Gramática de Construções (GC) e, mais especificamente, da Gramática de Construções Baseada no Uso (GCBU), arcabouço teórico em que esta pesquisa se insere. Em seguida, será introduzido o conceito de produtividade construcional, objeto de estudo da pesquisa.

### **GRAMÁTICA DE CONSTRUÇÕES E GRAMÁTICA DE CONSTRUÇÕES BASEADA NO USO: PRINCÍPIOS BÁSICOS**

Em linhas gerais, a Gramática de Construções (GC), surgida na década de 1980, é um modelo de arquitetura do conhecimento linguístico que emerge de uma reação ao modelo consagrado pela tradição gerativa, a partir do trabalho de autores como Charles Fillmore, Paul Kay e George Lakoff (Fillmore, 1985; Lakoff, 1987; Fillmore; Kay; O'Connor, 1988).

O princípio básico da GC é o de que as construções gramaticais, definidas como um pareamento convencional de informações de forma (propriedades fonológicas, morfossintáticas

e prosódicas) e significado (propriedades semânticas, pragmáticas e discursivas), podem descrever, integralmente, o conhecimento linguístico do falante. Nesse sentido, a gramática do falante deixa de ser vista como um sistema de regras derivacionais e passa a ser percebida como um inventário de construções.

Vale destacar que as construções gramaticais podem ser de naturezas distintas. Elas podem variar, no que diz respeito à forma, em grau de preenchimento fonológico: itens lexicais (como *mochila, casa*), expressões fixas (como *encher linguiça, bater as botas*) e esquemas morfológicos ou sintáticos semipreenchidos (como *X-dor, vê se X*), por exemplo, apresentam algum grau de preenchimento<sup>1</sup>, enquanto esquemas sintáticos abertos (como *S V OD OI*) e padrões entoacionais (como o *ascendente*) carecem deste tipo de especificação. Igualmente, no tocante ao significado, algumas construções podem trazer informações de natureza semântica (por exemplo, a ideia de transferência de posse na construção bitransitiva *S V OD OI*), *ao passo que outras apresentam especificações pragmáticas (caso da entoação ascendente)*.<sup>2</sup>

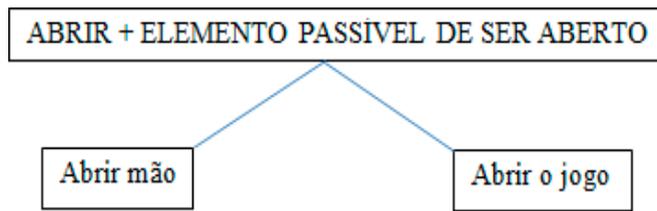
Embora possam se distinguir quanto ao nível de preenchimento fonológico, ou, ainda, quanto à natureza da informação de significado, todas têm em comum o fato de constituírem um pareamento entre forma e significado. Portanto, na GC assume-se que o léxico e a sintaxe não constituem módulos rigorosamente separados, mas que, em vez disso, compõem um *continuum de construções, que comporta desde* os elementos muito específicos até os padrões mais abstratos.

Quanto à organização, as construções se dispõem em uma rede de unidades inter-relacionadas, cujas relações de associação não se dão de forma aleatória, mas segundo um conjunto de critérios, entre os quais se destaca o grau de generalidade/especificidade. Desse modo, assume-se que são postuladas não só construções semipreenchidas, como ABRIR + ELEMENTO PASSÍVEL DE SER ABERTO, mas também construções mais concretas, como “abrir mão” e “abrir o jogo”.

Na prática, isso significa que as construções devem ser representadas de forma interconectada, de modo que as construções mais concretas, como “abrir mão” e “abrir o jogo”, se apresentam como uma especificação da construção mais abstrata ABRIR + ELEMENTO PASSÍVEL DE SER ABERTO. Como as construções mais concretas são subtipos da construção mais abstrata, podemos afirmar que esse tipo de relação é de natureza taxonômica. Esta representação pode ser observada a seguir.

<sup>1</sup> Aqui afirmamos que itens e expressões lexicais apresentam especificação fonológica, em vez de especificação lexical, porque: (i) afirmar que itens lexicais possuem especificação lexical seria redundante e, portanto, não informativo; e (ii) evidentemente, todo item lexical inclui informações de natureza fonológica, já que é composto por pelo menos um segmento sonoro.

<sup>2</sup> Reconhecemos que a distinção entre semântica e pragmática não é diretamente relevante, como evidencia o próprio fato de que esses dois tipos de informação se encontram no mesmo polo. Com efeito, a diferenciação entre esses dois componentes se presta, precisamente, a reforçar que tanto os fenômenos tradicionalmente associados à semântica quanto aqueles tradicionalmente associados à pragmática são igualmente abrangidos no polo funcional das construções gramaticais.

**Figura 1** – Rede de construções estruturadas por links taxonômicos

Fonte: O autor.

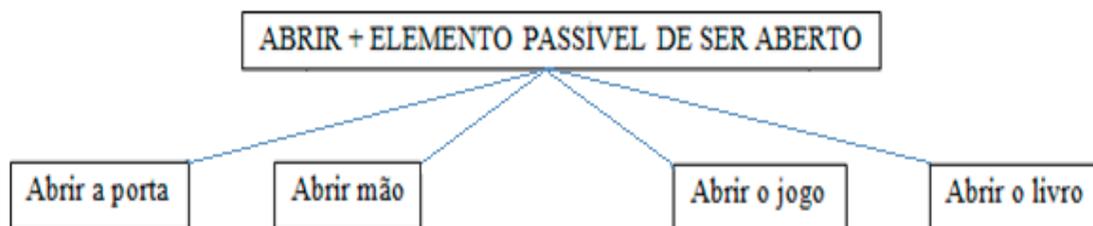
No que concerne à GCBU, uma vertente funcional-cognitiva da GC, podemos adicionar pelo menos dois princípios básicos que a diferem das variantes formalistas do modelo. O primeiro desses princípios, legado da tradição funcionalista norte-americana, diz respeito à influência decisiva do uso linguístico, ou seja, a experiência linguística do falante, no conhecimento subjacente. Essa ideia fundamental pode ser dividida em duas subpremissas: (i) não existe conhecimento linguístico inato; e (ii) a representação gramatical é constantemente moldada pela experiência.

Essa primeira perspectiva contrasta com a visão inatista da Linguística Gerativa, que postula que todo indivíduo já nasce com uma capacidade biológica para a linguagem. Sob esta perspectiva, a aquisição de uma língua é resultado direto do amadurecimento deste dispositivo inato. Em contrapartida, na concepção da Gramática de Construções Baseada no Uso (GCBU), a totalidade do conhecimento linguístico do falante é construída a partir do *input*.

A segunda subpremissa assume que a experiência com o *input linguístico* afeta continuamente o conhecimento internalizado do falante. Além de contradizer a ideia, fundamental para a tradição gerativa, de um “período crítico” para a aquisição de linguagem, essa premissa implica a possibilidade de representações redundantes na rede construcional, uma proposta que opõe a GCBU às vertentes formalistas da GC. Segundo essa hipótese, os falantes armazenam como construções independentes no *construction não* apenas construções idiomáticas (e, portanto, imprevisíveis), como “abrir mão” e “abrir o jogo”, mas ainda sequências regulares e previsíveis, como “abrir o livro” e “abrir a porta”, que poderiam ser derivadas de uma construção mais abstrata como ABRIR + ELEMENTO PASSÍVEL DE SER ABERTO, desde que sejam expostos a elas com frequência suficiente. Portanto, uma construção tenha alta frequência de ocorrência pode ser suscetível a uma representação independente. É por isso que se afirma que a GCBU assume a possibilidade de redundância representacional, uma vez que e admite a possibilidade de o falante armazenar mais do que é, de fato, necessário.

Retomando a representação do conhecimento linguístico, com o acréscimo do fator frequência como critério para a postulação de construções, enunciados como “abrir o livro” e “abrir a porta” passam a ser candidatos a figurar na rede construcional, desde que se demonstre que eles têm alta frequência de ocorrência. A figura 2, abaixo, ilustra esta representação.

**Figura 2** – Rede de construções estruturadas por links taxonômicos, levando-se em conta o critério de frequência



Fonte: O autor.

O segundo princípio específico da GCBU, este fortemente associado à tradição da Linguística Cognitiva, corresponde à ideia de que processos cognitivos gerais – e não especificamente linguísticos – devem ser evocados para explicar fenômenos linguísticos e, inclusive, a estrutura gramatical. Assim, mecanismos psicológicos como a analogia, a indução e a categorização são fundamentais, sob a ótica da GCBU, para explicar a estrutura da rede construcional armazenada na mente do falante.

#### PRODUTIVIDADE CONSTRUCIONAL

As construções podem ser frequentemente usadas de forma inovadora, permitindo ao falante produzir sentenças nunca sequer ouviu antes. A este fenômeno, damos o nome de criatividade linguística. Ao conceito de criatividade linguística, podemos relacionar o de produtividade construcional, que tem sido abordado na literatura linguística sob perspectivas diversas. Dentre os sentidos que os termos “produtivo” e “produtividade” podem assumir, Barðdal (2008) destaca três: *generalidade, regularidade e extensibilidade*.<sup>3</sup>

Para exemplificar esses conceitos, podemos pensar na categoria dos verbos regulares da língua inglesa, que formam o passado com o sufixo *-ed*: *esta categoria é altamente geral, na medida em que não há restrições fonológicas e/ou morfológicas para seus membros; é altamente regular, uma vez que é baseada em uma regra, um padrão; e altamente extensível, já que a maioria dos, se não todos, novos verbos da língua inglesa assumem o comportamento dos regulares. A construção V-ed pode, então, ser considerada produtiva, independentemente do sentido de produtividade que o analista esteja levando em conta – seja o de generalidade, o de regularidade ou o de extensibilidade.*

<sup>3</sup> Importante ressaltar que, na literatura construcionista, o termo *produtividade* também se relaciona com o que é conhecido como *frequência token*, ou *frequência de ocorrência*, que é uma medida da quantidade de vezes que determinada construção ocorre em um *corpus*. No entanto, tanto Barðdal (2008), como veremos na próxima seção, quanto outros autores como Traugott e Trousdale (2013), parecem favorecer o tratamento da produtividade como fenômeno atrelado à *frequência type* ou *frequência de tipo* – uma medida da quantidade de itens distintos que a construção comporta –, e não, necessariamente, à *frequência token*. Reconhecemos que, apesar de não gozar da mesma centralidade da *frequência type*, a *frequência token* tem sua importância no modelo. Nesta pesquisa, porém, seguimos Barðdal (2008) na opção por tratar a produtividade como *extensibilidade*, que, como ficará claro logo adiante, se relaciona diretamente com o fator *frequência de tipo*.

Por outro lado, a categoria composta pelo grupo dos verbos irregulares do inglês que alternam sistematicamente o fonema /i/ para /æ/ ou / / no passado (como *sing-sang*), por exemplo, diferentemente da construção *V-ed*, não é, concomitantemente, geral, regular e extensível. Apesar do nome, esta é uma categoria regular (visto que seu comportamento é baseado em um padrão), mas não é uma categoria geral (já que o seu padrão é altamente restrito quanto às especificações fonológicas dos verbos que seguem o padrão) nem plenamente extensível (pelo menos não tanto quanto a categoria dos regulares, pois, como já dito, a maioria ou todos os verbos novos formam o passado em *-ed*). Portanto, quando se fala em produtividade, é necessário especificar qual ou quais dos sentidos são adotados, tendo em vista a multiplicidade de sentidos do termo.

Neste trabalho, seguindo a proposta de Barðdal (2008), a produtividade é compreendida como *extensibilidade* – a possibilidade de se estender uma construção gramatical de modo que esta seja instanciada por novos itens lexicais. Dessa maneira, nesta pesquisa, em especial, busca-se contribuir para a discussão de quais fatores são responsáveis por favorecer a extensibilidade de uma construção gramatical, apontando a importância da analogia baseada na morfofonologia para a inovação linguística.<sup>4</sup>

## REVISÃO DA LITERATURA

Conforme pontuado na seção anterior, na literatura linguística acerca do tema, busca-se, em geral, entender quais são os fatores que podem afetar a produtividade de uma construção gramatical. Nesta seção, será apresentado um resumo das hipóteses já sugeridas.

### FREQUÊNCIA DE TIPO E COERÊNCIA SEMÂNTICA

Barðdal (2008) aponta a frequência de tipo (FT) e a coerência semântica (CS) como fatores responsáveis por aumentar ou reduzir a produtividade de uma construção. A primeira é entendida como o número total de tipos que podem instanciar uma dada construção abstrata ou semiabstrata; a segunda é definida como a consistência semântica interna entre os membros de uma categoria.

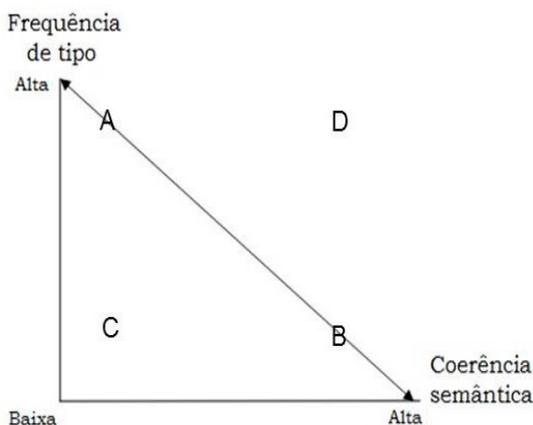
Para ilustrar o fator FT, podemos pensar nas construções morfológicas *X-th* e *X-ness* da língua inglesa. A construção *X-th* tem baixa frequência de tipo, uma vez que há poucos itens lexicais que podem instanciá-la. A construção *X-ness*, por outro lado, tem frequência de tipo mais alta, já que existe uma quantidade bem maior de itens que podem ser combinados com ela (Barðdal, 2008).

<sup>4</sup> Com isso, não estamos assumindo que, necessariamente, todo processo analógico conduz à mudança linguística. Neste estudo, no entanto, avaliamos o papel potencial da analogia para a motivação de inovações lexicais.

Para explicar o segundo fator, que diz respeito à coerência semântica, imaginemos duas categorias: a primeira sendo formada pelos verbos do inglês: *break, get e hold*; e a segunda, composta pelos verbos *break, crush e tear*. Ao pensarmos no significado dos verbos de ambas as categorias, chegamos à conclusão de que a coerência semântica interna do primeiro grupo é bem menor do que a do segundo, cujos membros são semanticamente mais próximos. Em outros termos, poderíamos dizer que a primeira categoria apresenta baixa coerência semântica, enquanto a segunda apresenta alta coerência semântica.

Segundo a autora, esses dois fatores se relacionam de forma inversamente proporcional. O gráfico a seguir ilustra a correlação por ela proposta:

**Figura 3** – Inversa correlação entre FT e CS



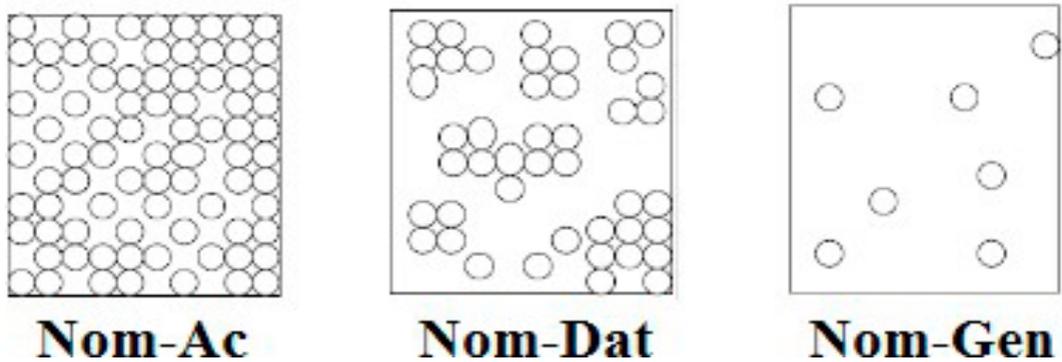
Fonte: Barðdal, J. **Productivity**: Evidence from case and argument structure in Icelandic. Amsterdam: John Benjamin Publishing Company, 2008. Tradução nossa.

De acordo com esta hipótese, para que dada construção tenha um grau alto de produtividade, é necessário que ela apresente alta FT e baixa CS, ou baixa FT e alta CS, o que equivaleria às construções representadas, nesta ordem, por A e B no esquema acima. Assim sendo, uma construção com FT e CS baixas – representada por C, no gráfico – seria também produtiva, mas em grau muito menor, entendendo-se a produtividade como um fenômeno gradiente. Já uma construção com FT e CS altas – representada por D – seria não só improdutiva, mas inexistente na língua, uma vez que a frequência de tipo leva, necessariamente, segundo Barðdal (2008), a uma menor coerência semântica entre os membros da categoria.

Para testar sua hipótese, a autora se vale de uma análise de *corpus*, em que são analisadas construções de caso do islandês. A autora investiga os empréstimos mais recentes na língua, com verbos do campo semântico da tecnologia, buscando verificar que construções seriam preferencialmente instanciadas por eles. Nesse estudo, três construções do islandês são analisadas: *Nominativo-Acusativo, Nominativo-Dativo e Nominativo-Genitivo*. Apresentam,

respectivamente, as condições alta FT / baixa CS, baixa FT / alta CS e baixa FT / baixa CS, e podem ser representadas pelos esquemas abaixo.

**Figura 4** – Representação esquemática das construções estudadas por Barðdal (2008)



Fonte: Barðdal (2008).

Partindo desse cenário, podemos esperar que as construções *Nominativo-Acusativo* e *Nominativo-Dativo* apresentem alto grau de produtividade, licenciando, assim, um maior número de empréstimos, enquanto a construção *Nominativo-Genitivo* teria um baixo grau de produtividade. Estas hipóteses foram, de fato, confirmadas. Dentre os 107 verbos analisados, 68 deles instanciaram a primeira construção, 39 instanciaram a segunda e nenhum instanciou a terceira construção. Os resultados sugerem, portanto, que a correlação inversa entre os dois fatores apontados por Barðdal parece, realmente, regular a produtividade das construções gramaticais.

#### COVERAGE SEMÂNTICO

Sob outra perspectiva, Suttle e Goldberg (2011) defendem que os falantes se mostram mais confiantes quanto à extensibilidade de uma construção para novos usos na medida em que as instâncias atestadas cobrem o espaço semântico que inclui o uso alvo. A relevância desta cobertura – que vamos chamar, seguindo as autoras, de *coverage* – é sustentada pela combinação de efeitos de frequência de tipo, variabilidade semântica e similaridade semântica, bem como pela interação entre eles.

Os efeitos frequência de tipo e variabilidade semântica correspondem aos dois fatores apontados por Barðdal (2008), com a diferença (pouco relevante) de que a variabilidade semântica é o exato inverso da coerência semântica. Assim, dizer que uma construção apresenta baixa coerência semântica equivale a dizer que ela apresenta alta variabilidade semântica e vice-versa.

Já a similaridade semântica é um efeito apontado exclusivamente por Suttle e Goldberg (2011) e diz respeito à possível similaridade entre os casos atestados e o uso alvo (isto é, o uso novo, potencial). Para ilustrar, pensemos em uma construção que pode ser instanciada pelos

verbos *quebrar*, *correr* e *dormir* (o que significa que eles pertencem a uma mesma categoria), e em outra instanciável por *quebrar*, *rasgar* e *repartir* (que formam, então, outra categoria). Pensemos também que o uso alvo é um enunciado com o verbo *cortar*. Se considerarmos a similaridade entre o significado dos verbos de ambas as categorias e o verbo do uso alvo, concluiremos que os membros da primeira apresentam menor proximidade semântica em relação ao do uso alvo em comparação com os da segunda, que pertencem ao mesmo campo semântico (prejuízo à integridade material). Em outras palavras, poderíamos dizer que, no primeiro caso, a similaridade semântica é baixa, enquanto, no segundo, é alta.

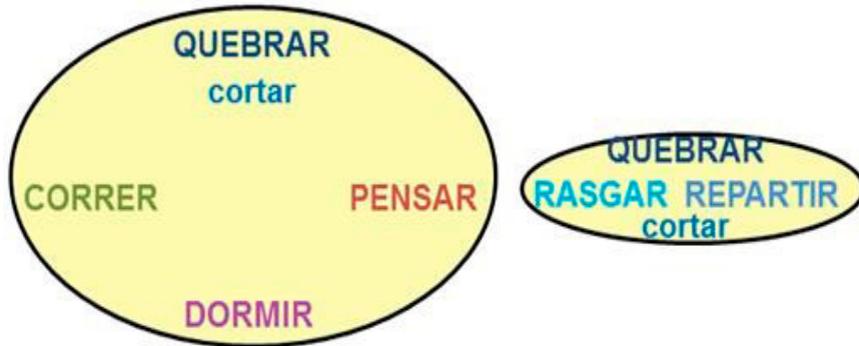
Para testar sua hipótese, as autoras desenvolveram um experimento com construções de estrutura argumental de uma língua fictícia que podiam ser instanciadas por verbos do inglês, levando em conta sua frequência de tipo e suas particularidades semânticas. A tarefa experimental consistia em analisar os enunciados apresentados como casos atestados nesta língua fictícia e, a partir disso, atribuir uma nota (que podia variar entre 0 e 100) que representasse o grau de confiança do participante quanto à possibilidade de produção de um dado enunciado potencial. Os três efeitos contemplados por elas foram testados de forma combinada. Primeiramente, foi feito um experimento a fim de testar os efeitos da variabilidade e da frequência de tipo. Em seguida, foram testados os efeitos da variabilidade e da similaridade semânticas.

Nos dois experimentos, foram verificados efeitos de interação, além de um efeito principal de variabilidade no primeiro e de similaridade no segundo. No primeiro experimento, os estímulos da condição de alta variabilidade foram mais confiantemente estendidos para usos novos do que aqueles pertencentes à condição de baixa variabilidade. O mesmo foi demonstrado para a condição de alta frequência de tipo em relação à baixa frequência, embora este fator tenha exibido um efeito mais marginal. Além disso, a variabilidade apresentou um efeito mais forte nos casos de alta frequência, o que sugere uma interação entre ambos.

No segundo experimento, para os casos de alta similaridade, quanto mais alta fosse a variabilidade, menor era a confiança do falante em estender a palavra para um uso novo. Nos casos de média similaridade, por outro lado, a confiança do falante aumentava quanto mais alta fosse a variabilidade entre os casos atestados. Por fim, para os casos de baixa similaridade, o fator variabilidade não foi significativo.

Os resultados são explicados em termos de *coverage*, que pode ser entendido como a medida em que os itens atestados recobrem o espaço categorial da categoria composta conjuntamente pela soma destes com o item alvo. Desse modo, o *coverage* é alto se a categoria resultante é densamente povoada e baixo se a categoria possui mais espaços vazios.

A figura abaixo ilustra este processo. Os itens atestados estão representados em maiúsculas e, os itens alvo, em minúsculas.

**Figura 5** – Baixo e alto *coverage*

Fonte: O autor.

Como podemos observar acima, a categoria à esquerda apresenta baixo *coverage*, na medida em que apresenta muitos espaços vazios, não sendo, portanto, amplamente coberta. O mesmo não ocorre, entretanto, com a categoria à direita, cuja cobertura categorial é alta. Em outras palavras, ela apresenta alto *coverage*. Em resumo, Suttle e Goldberg (2011) sugerem que os três fatores levantados previamente – frequência de tipo, variabilidade semântica e similaridade semântica – podem ser reduzidos ao conceito de *coverage*, de modo que o alto *coverage* aponta para um conseqüente aumento da produtividade construcional, ao passo que o baixo *coverage* reduz a possibilidade de se estender a construção para novos usos.

## METODOLOGIA

Tomando como base os resultados já alcançados no que concerne à temática da produtividade construcional, busca-se, nesta seção, apontar a metodologia adotada para responder à nossa pergunta de pesquisa. Conforme adiantado na introdução, perguntamo-nos se a coerência morfofonológica de uma categoria de itens lexicais, bem como a similaridade morfofonológica desses itens com o item-alvo, também contribuiriam para a extensibilidade de uma construção gramatical para usos inéditos. Para isso, foi desenvolvido um experimento off-line adaptado do estudo de Suttle e Goldberg (2011), o qual passamos a descrever.

### O MÉTODO EXPERIMENTAL: VISÃO GERAL

A tarefa experimental consistia em analisar os enunciados apresentados como casos atestados em uma língua fictícia e, a partir disso, atribuir uma nota de 1 a 5 que representasse o grau de confiança do participante – nossa variável dependente – quanto à possibilidade

de produção de outro enunciado na mesma língua. A seguir, segue um exemplo de estímulo recebido pelos participantes no experimento:

**Figura 6** – Exemplo de estímulo experimental

Suponha que você já tenha ouvido as seguintes frases sendo ditas por um falante nativo do zargotiano:

**As feg os telk padronizar-ju.**  
**As ritny os ilp interromper-ju.**

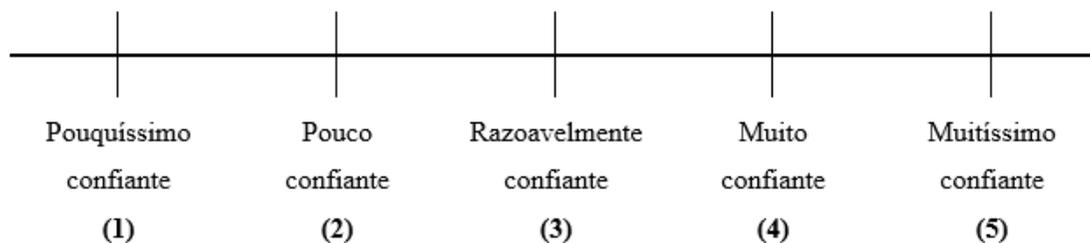
Levando em conta que as frases acima são possíveis no zargotiano, o quão confiante você se sente em produzir a frase abaixo?

**As siln os farg merecer-ju.**

Fonte: O autor.

Ao final de cada estímulo, os participantes tinham à disposição uma escala Likert de cinco pontos, como a representada em seguida, que servia de ferramenta para que pudessem atribuir a nota que melhor representava seu grau de confiança, após a análise dos casos atestados na língua, quanto à possibilidade de produção de cada um dos enunciados alvo.

**Figura 7** – Escala Likert



Fonte: O autor.

Como observado, os enunciados eram compostos por artigos e verbos do português, além de palavras inventadas para compor sintagmas da língua imaginária, organizados de modo a combinar nossas variáveis independentes. A fim de evitar que houvesse interferência da semântica dos verbos do português inseridos nas sentenças criadas, foi necessário incluir, nos estímulos experimentais, verbos de classes semânticas distintas. Assim, para este experimento, as variáveis independentes se restringiram, com exceção da frequência de tipo, a fatores de natureza morfofonológica, sendo estes a variabilidade entre os casos atestados – a qual chamaremos de variabilidade morfofonológica – e a similaridade entre os casos atestados

e o enunciado a ser julgado pelos participantes – a qual trataremos como similaridade morfofonológica. Especificamente, o controle dessas variáveis se deu pela vogal temática dos verbos que compunham as sentenças.<sup>5</sup>

## PROCEDIMENTOS

Os participantes recebiam um questionário impresso contendo algumas instruções iniciais acerca do experimento. Essas instruções eram lidas juntamente com os participantes e reforçadas oralmente. Nas páginas seguintes, havia dois estímulos-teste que não foram levados em consideração na análise dos resultados. Os estímulos eram lidos em voz alta, juntamente com os participantes, que tinham o tempo que fosse necessário para analisar os enunciados e marcar suas respostas. Isto foi feito com o intuito de nos certificarmos de que não restariam dúvidas quanto à tarefa experimental que deveria ser cumprida. Para os estímulos críticos, que vinham logo em seguida, os participantes foram deixados livres para ler, reler, analisar e responder. Não foi delimitado tempo mínimo ou máximo para a resposta ao questionário. Entretanto, os participantes recebiam uma recomendação oral para que não respondessem de forma aleatória, mas de maneira consciente, analisando e levando em conta os enunciados dados como casos atestados na língua para a marcação de suas respostas.

## COMPOSIÇÃO DOS ESTÍMULOS

Foram postas em julgamento sentenças da língua fictícia que chamamos de *zargotiano*, as quais continham verbos do português e combinavam nossas três variáveis independentes – a saber, variabilidade morfofonológica, similaridade morfofonológica e frequência de tipo –, cada uma delas podendo variar entre baixa e alta, distribuídas da seguinte forma:

**Tabela 1** – Esquema de distribuição das variáveis independentes

	<b>BAIXA</b>	<b>ALTA</b>
<b>VARIABILIDADE</b>	Uma única vogal temática para todos os verbos atestados	2 ou mais vogais temáticas diferentes entre os verbos atestados
<b>SIMILARIDADE</b>	Vogal temática do verbo alvo diferente da vogal temática de todos os verbos atestados	Vogal temática do verbo alvo igual a pelo menos uma vogal temática dos verbos atestados
<b>FREQUÊNCIA</b>	2 itens atestados	6 itens atestados

Fonte: O autor.

<sup>5</sup> Embora, potencialmente, as palavras inventadas pudessem ser consideradas na avaliação dos participantes e, portanto, afetar os fatores de natureza morfofonológica, não se verificou interferência desses itens no experimento original. Com efeito, parece esperado, inclusive, que o participante se concentre nos itens conhecidos (neste caso, os verbos) em detrimento daqueles desconhecidos e sem significado, já que, como aponta Goldberg (2019), a memória e a atenção são fortemente orientadas para o significado.

Ao todo trabalhamos com 8 condições experimentais: baixa variabilidade/baixa similaridade/baixa frequência (BV/BS/BF), baixa variabilidade/baixa similaridade/alta frequência (BV/BS/AF), baixa variabilidade/alta similaridade/baixa frequência (BV/AS/BF), baixa variabilidade/alta similaridade/alta frequência (BV/AS/AF), alta variabilidade/baixa similaridade/baixa frequência (AV/BS/BF), alta variabilidade/baixa similaridade/alta frequência (AV/BS/AF), alta variabilidade/alta similaridade/baixa frequência (AV/AS/BF) e alta variabilidade/alta similaridade/alta frequência (AV/AS/AF). A seguir, são apresentados exemplos de estímulos em diferentes condições.

### Quadro 1

Exemplo 1:

Suponha que você já tenha ouvido as seguintes frases sendo ditas por um falante nativo do zargotiano:

**As alk os burd remover-yi.**  
**As kwap os tirp ler-yi.**

Levando em conta que as frases acima são possíveis no zargotiano, o quão confiante você se sente em produzir a frase abaixo?

**As knib os gorp restringir-yi.**

Fonte: O autor.

Como é possível observar há aqui uma combinação específica dos níveis das três variáveis independentes. O que resulta na condição BV/BS/BF, já que os estímulos apresentam uma única vogal temática (-E-) para os verbos dos itens atestados, o que corresponde a baixa variabilidade, e que esta é diferente da do verbo usado na instância alvo (-I-). Ou seja, há uma baixa similaridade. Além disso, há apenas 2 itens atestados, o que equivale a baixa frequência.

### Quadro 2

Exemplo 2:

Suponha que você já tenha ouvido as seguintes frases sendo ditas por um falante nativo do zargotiano:

**As planf os guft redimir-ky.**

**As gotdy os blub dispor-ky.**

**As vupt os claper inverter-ky.**

**As sift os cogt possuir-ky.**

**As blodp os zacy manter-ky.**

**As konf os bends decompor-ky.**

Levando em conta que as frases acima são possíveis no zargotiano, o quão confiante você se sente em produzir a frase abaixo?

**As skoos os gare engrandecer-ky.**

Fonte: O autor.

Neste segundo exemplo a combinação das variáveis independentes resulta na condição AV/AS/AF, uma vez que há 3 vogais temáticas diferentes (-E-, -I- e -O-) para os 6 verbos atestados. O que corresponde a alta variabilidade e alta frequência. A vogal temática do verbo alvo (-E, neste caso) é igual a pelo menos uma vogal temática dos verbos atestados o que, para fins deste estudo, corresponde a uma alta similaridade, conforme especificado na tabela 1.

Cada um dos participantes foi exposto, além dos estímulos-teste, a 16 estímulos críticos, sendo dois deles para cada condição experimental, distribuídos em sequência aleatória, sem que houvesse qualquer tipo de padrão que pudesse interferir nas respostas dadas pelos participantes.

### PREVISÕES EXPERIMENTAIS E ANÁLISE ESTATÍSTICA

Tendo como ponto de partida o estudo de Suttle e Goldberg (2011) sobre a atuação do *coverage semântico*, especificamos três previsões similares quanto ao resultado do experimento aplicado para testar a atuação do *coverage morfofonológico*. Primeiramente, previmos um efeito principal de similaridade, que se daria através de uma diferença significativa entre as condições

de alta e baixa similaridade. As condições de alta similaridade receberiam notas mais altas na escala do que as de baixa similaridade, que tenderiam a receber notas mais baixas. Este resultado confirmaria a previsão de que a alta similaridade favorece a produtividade.

Em segundo lugar, previmos um efeito de interação entre variabilidade e similaridade morfofonológicas, de forma que a variabilidade seria um fator relevante somente nas condições de alta similaridade. Tal como ocorre com a variabilidade e a similaridade semânticas, conforme apontado por Suttle e Goldberg (2011). Portanto, haveria uma diferença significativa entre as notas recebidas nas condições de baixa e alta variabilidade quando a similaridade fosse alta, confirmando a nossa hipótese de que o efeito da variabilidade sobre a produtividade depende da similaridade morfofonológica.

Por fim, previmos que haveria uma diferença também significativa entre as notas recebidas nas condições de baixa e alta variabilidade quando em interação com a frequência. Esta terceira previsão confirmaria a hipótese de que o efeito da variabilidade morfofonológica sobre a produtividade depende da frequência de tipo.

Para verificar o efeito dessas previsões, foram utilizadas duas diferentes técnicas de estatística qui-quadrado: o teste qui-quadrado de homogeneidade, que foi usado para verificar se houve diferença entre as duas amostras distintas (neste caso, o conjunto de graus de confiança associados a alta e baixa similaridade) ou se elas são homogêneas, equivalentes; e o qui-quadrado de independência, usado visando aferir se as variáveis consideradas para as duas últimas previsões experimentais (variabilidade x similaridade e variabilidade x frequência, respectivamente) produziram efeito independentemente ou se houve interação entre elas.

## **RESULTADOS E DISCUSSÃO**

Apresentamos nesta seção, de forma sintetizada, os resultados obtidos no experimento. Eles estão divididos em três seções, que remetem a cada uma das três previsões experimentais apontadas na seção anterior. A quarta seção apresenta a discussão destes resultados e suas implicações no desdobramento da pesquisa.

### **PRIMEIRA PREVISÃO**

Conforme mencionado anteriormente, para que a nossa primeira previsão sobre o efeito da similaridade morfofonológica sobre a produtividade fosse confirmada, seria necessário que os estímulos pertencentes às condições de alta similaridade fossem avaliados com notas mais altas na escala Likert, comparativamente àqueles pertencentes às condições de baixa similaridade. A tabela abaixo mostra a média dos graus atribuídos pelos participantes, sendo as quatro primeiras condições referentes a alta similaridade e as quatro últimas, a baixa similaridade.

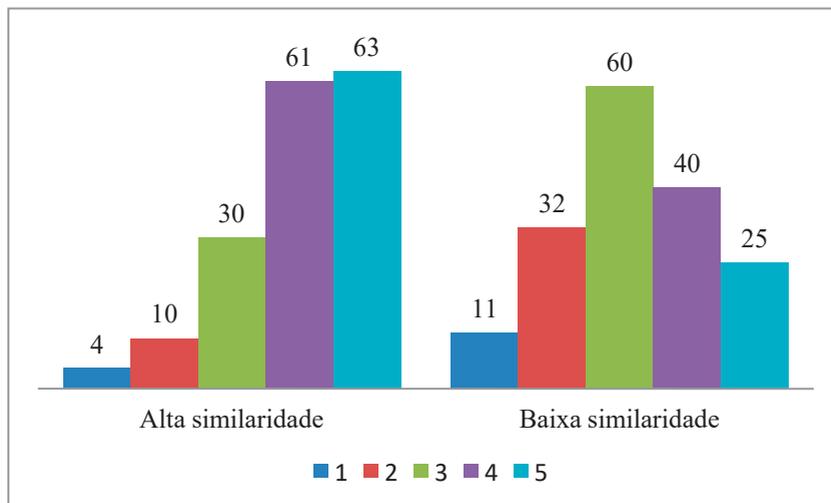
**Tabela 2** – Média geral dos graus conferidos pelos participantes aos estímulos críticos

CONDIÇÃO	<b>ALTA SIMILARIDADE</b> (BV/AS/BF, BV/AS/AF, AV/AS/BF, AV/AS/AF)	<b>BAIXA SIMILARIDADE</b> (BV/BS/BF, BV/BS/AF, AV/BS/BF, AV/BS/AF)
<b>MÉDIA</b>	4	3,21

Fonte: O autor.

A tabela acima aponta para uma diferença numérica entre as notas atribuídas aos grupos de estímulos. Podemos observar, conforme previsto, que a média dos graus atribuídos aos estímulos de baixa similaridade (3,21) é, de fato, numericamente menor que a média dos graus conferidos nas condições de alta similaridade (4).

Entretanto, para que se pudesse verificar se esta diferença era estatisticamente significativa, os resultados da amostra passaram pelo teste qui-quadrado de homogeneidade. Esse teste compara o conjunto de graus atribuídos nas duas condições focalizadas aqui, isto é, alta e baixa similaridade. O resultado dessa comparação pode ser visto no gráfico a seguir.

**Gráfico 1** – Teste qui-quadrado de homogeneidade (similaridade): P-valor < 0.001

Fonte: O autor.

Podemos observar, acima, que a quantidade de graus 4 e 5 atribuídos ao grupo de alta similaridade é 61 e 63, respectivamente. Enquanto apenas 40 e 25 notas 4 e 5, nesta ordem, foram atribuídas ao grupo de baixa similaridade. A existência dessa diferença, assim como dos demais graus disponíveis na escala Likert, sugere que o julgamento do falante é marcadamente diferente quando diante de estímulos com alta e baixa similaridade morfofonológica. Além disso, e mais importante, essa sugestão é confirmada pelo teste qui-quadrado de homogeneidade, cujo resultado mostra que a diferença na distribuição dos graus entre os dois

grupos é significativa ( $p < 0.001$ ). Ou seja, não se deve ao acaso, podendo ser atribuída de fato à existência de maior ou menor similaridade em cada condição. Isso sugere, que a similaridade morfofonológica tem real influência sobre a extensibilidade construcional. A tendência a notas mais altas nas condições de alta similaridade é, de fato, significativa, o que vai ao encontro de nossa primeira previsão.

## SEGUNDA PREVISÃO

Nossa segunda previsão era a de que haveria uma diferença significativa entre as notas recebidas nas condições de baixa e alta variabilidade quando a similaridade fosse alta, de modo que, na situação de alta similaridade, os estímulos das condições de baixa variabilidade receberiam notas mais altas, ao passo que as notas seriam mais baixas quando a variabilidade fosse alta. Além disso, previmos que a baixa similaridade não implicaria diferença significativa. As tabelas abaixo mostram as médias dos graus atribuídos às condições de alta similaridade (primeira tabela) e de baixa similaridade (segunda tabela).

**Tabela 3** – Graus conferidos pelos participantes aos estímulos críticos das condições de alta similaridade

CONDIÇÃO	BAIXA VARIABILIDADE (BV/AS/BF, BV/AS/AF)	ALTA VARIABILIDADE (AV/AS/BF, AV/AS/AF)
MÉDIA	4,19	3,81

Fonte: O autor.

**Tabela 4** – Graus conferidos pelos participantes aos estímulos críticos das condições de baixa similaridade

CONDIÇÃO	BAIXA VARIABILIDADE (BV/BS/BF, BV/BS/AF)	ALTA VARIABILIDADE (AV/BS/BF, AV/BS/AF)
MÉDIA	3,23	3,18

Fonte: O autor.

Na tabela 3 podemos observar uma diferença numérica entre a média das notas atribuídas aos estímulos das condições de baixa e alta variabilidade (4,19 x 3,81) nas condições de alta similaridade. Já na tabela 4, referente às condições de baixa similaridade, esta diferença é numericamente menor (3,23 x 3,18). A tabela 5, abaixo, apresenta a soma das notas por condição e nos ajuda a visualizar melhor estes resultados.

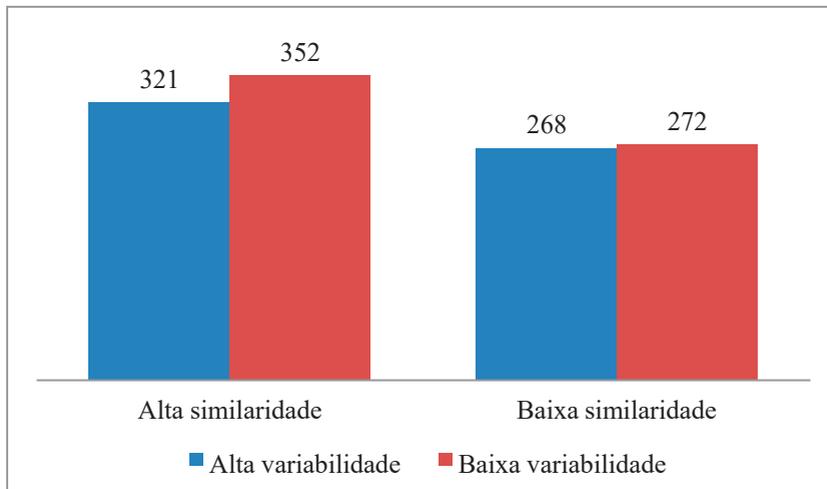
**Tabela 5** – Soma das notas por condição

	<b>ALTA SIMILARIDADE</b>	<b>BAIXA SIMILARIDADE</b>
<b>ALTA VARIABILIDADE</b>	321	268
<b>BAIXA VARIABILIDADE</b>	352	272
<b>DIFERENÇA</b>	31	4

Fonte: O autor.

Os números acima revelam uma diferença numérica aparentemente significativa entre as condições de alta similaridade (31), com a baixa variabilidade favorecendo a extensibilidade, exatamente como previsto. Os números revelam, ainda, uma diferença numérica muito pequena (4) entre as condições de baixa similaridade, o que sugere que a variabilidade não parece ser um fator relevante quando a similaridade é baixa.

Porém, para validar esses resultados era necessária uma comprovação estatística. Para isso, os resultados passaram pelo teste qui-quadrado de independência, também descrito anteriormente, com o objetivo de verificar se havia, de fato, um efeito de interação entre os fatores variabilidade e similaridade. O resultado do teste se encontra na sequência.

**Gráfico 2** – Teste qui-quadrado de independência (variabilidade x similaridade): P-valor = 0.5

Fonte: O autor.

Como se pode observar acima, o p-valor para este teste foi de 0.5, mostrando que a diferença não é estatisticamente significativa. Portanto, este resultado não aponta para existência de efeito de interação entre os dois fatores. O que contraria nossa previsão.

### TERCEIRA PREVISÃO

Segundo nossa terceira previsão, esperaríamos encontrar uma diferença significativa entre as notas recebidas nas condições de baixa e alta variabilidade quando a frequência fosse alta. Os estímulos das condições de baixa variabilidade receberiam notas mais altas, enquanto a variabilidade alta favoreceria notas mais baixas na escala. Previmos, ainda, que não haveria uma diferença expressiva quando a frequência fosse baixa. As tabelas abaixo mostram os graus atribuídos às condições de alta e baixa frequência.

**Tabela 6** – Graus conferidos pelos participantes aos estímulos críticos das condições de alta frequência

CONDIÇÃO	BAIXA VARIABILIDADE (BV/BS/AF, BV/AS/AF)	ALTA VARIABILIDADE (AV/BS/AF, AV/AS/AF)
MÉDIA	3,64	3,49

Fonte: O autor.

**Tabela 7** – Graus conferidos pelos participantes aos estímulos críticos das condições de baixa frequência

CONDIÇÃO	BAIXA VARIABILIDADE (BV/BS/BF, BV/AS/BF)	ALTA VARIABILIDADE (AV/BS/BF, AV/AS/BF)
MÉDIA	3,78	3,5

Fonte: O autor.

Nas tabelas 6 e 7, observamos uma diferença numérica reduzida entre a média das notas atribuídas aos estímulos das condições de baixa e alta variabilidade, tanto quando a frequência é alta (3,64 x 3,49) quanto quando ela é baixa (3,78 x 3,5). A tabela 8 permite uma melhor visualização destes resultados.

**Tabela 8** – Soma das notas por condição

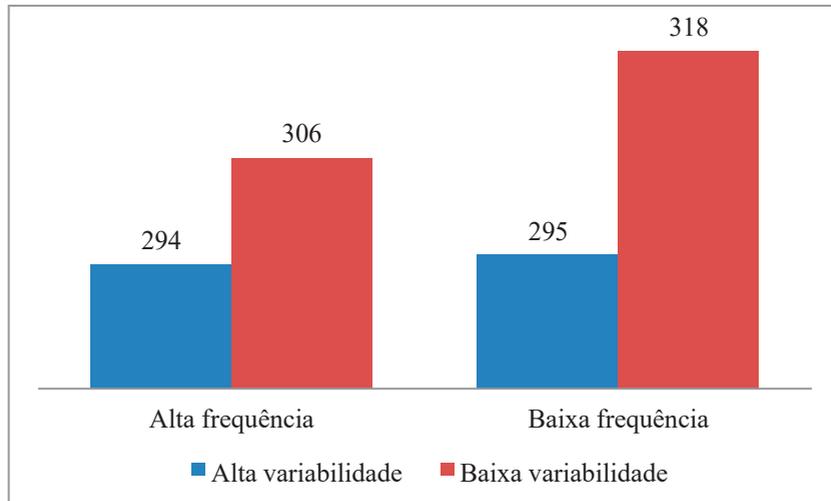
	ALTA FREQUÊNCIA	BAIXA FREQUÊNCIA
ALTA VARIABILIDADE	294	295
BAIXA VARIABILIDADE	306	318
DIFERENÇA	12	23

Fonte: O autor.

Ao contrário do esperando, os números revelam uma diferença numérica pequena entre as condições de alta e baixa frequência. De todo modo, para testar nossa hipótese, os resultados passaram pelo teste qui-quadrado de independência, com o objetivo de verificar

se, estatisticamente, havia um efeito de interação entre os fatores variabilidade e frequência. O gráfico abaixo exibe o resultado do teste.

**Gráfico 3** – Teste qui-quadrado de independência (frequência x variabilidade): P-valor = 0.8



Fonte: O autor.

Conforme o gráfico acima constata, o p-valor para este teste foi de 0.8, não sendo, pois, menor que 0.05, indicando que, estatisticamente, não há efeito de interação entre frequência e variabilidade, ao contrário do que esperávamos. Nossa terceira previsão não pôde ser confirmada.

## DISCUSSÃO

Os resultados reportados acima puderam confirmar apenas a primeira de nossas três previsões. Como vimos, nosso experimento aponta para a existência de um claro efeito de similaridade morfofonológica sobre a produtividade construcional, demonstrando que este pode ser um fator que encoraja o falante a estender uma construção para novos itens. Em outras palavras, a semelhança morfofonológica entre os itens atestados para dada construção e o item alvo parece favorecer sua extensibilidade para novos itens, o que confirma nossa primeira hipótese. Conforme apontado na seção *Revisão da literatura*, este foi o mesmo resultado encontrado por Suttle e Goldberg (2011) para a similaridade semântica.

Apesar de as outras duas previsões não terem sido confirmadas, os números permitem algumas reflexões. Refletimos que, embora estatisticamente não tenhamos verificado efeito de interação entre variabilidade e similaridade morfofonológicas, há dois indícios numéricos de que a alta similaridade pode ser necessária para que a variabilidade tenha impacto sobre

a produtividade construcional: as médias dos graus atribuídos às condições de baixa e alta variabilidade (4,19 x 3,81), e a diferença de 31 pontos entre a soma das notas conferidas às mesmas condições quando a similaridade era alta, em oposição à diferença de apenas 4 pontos quando a similaridade era baixa. Como essa diferença numérica se direciona de uma maneira esperada, ainda que não tenha se revelado estatisticamente significativa, é possível suspeitar que uma posterior ampliação da base de dados venha a resultar em significância estatística.

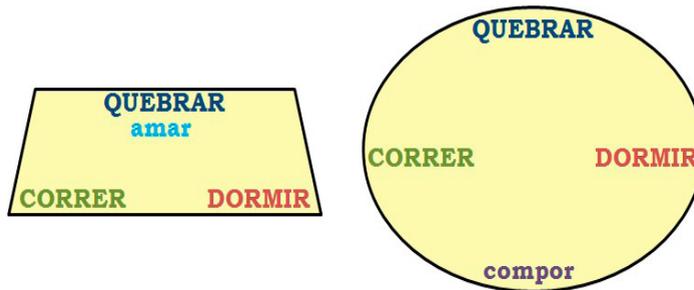
Desse modo, é possível supor que, para que a variabilidade seja um fator motivador da extensibilidade de uma construção gramatical para novos usos, talvez seja necessário que haja alta similaridade morfofonológica entre os itens atestados e o item alvo. Do contrário, a variabilidade morfofonológica deixaria de ser um fator significativo. Caso este resultado seja obtido em uma eventual ampliação deste estudo, então terá sido encontrado exatamente o mesmo padrão que Suttle e Goldberg (2011) identificaram para a variabilidade e a similaridade semânticas.

Além disso, nossa terceira previsão, embora também não tenha sido confirmada, nos leva a uma segunda reflexão. Como vimos, apesar de a diferença numérica observada entre o efeito da variabilidade nas condições de alta e baixa frequência ter sido muito reduzida (médias de 3,64 x 3,49 e 3,78 x 3,5, respectivamente) e estatisticamente não significativa, ela também vai em direção ao que se encontrou para a variabilidade semântica (Suttle; Goldberg, 2011). De acordo com as autoras, nos casos de alta frequência, a baixa variabilidade semântica aumenta a confiança do falante em estender a construção para um uso novo de forma mais acentuada do que quando há alta variabilidade. É exatamente o que se pode verificar a partir da diferença das médias nas condições de alta frequência (3,64 x 3,49), com a baixa variabilidade morfofonológica favorecendo a produtividade. Deste modo, os resultados referentes à nossa segunda e à nossa terceira previsão experimental encorajam a ampliação da base de dados, a fim de verificar se não estamos diante de falsos negativos.

Em suma, existem alguns indícios que estimulam a aposta na ideia de um *coverage morfofonológico como um mecanismo cognitivo capaz de explicar a produtividade construcional*. Vale lembrar, conforme apresentado na seção *Pressupostos teóricos, que o coverage é entendido como a medida em que os itens atestados recobrem o espaço categorial* da categoria composta conjuntamente pela soma destes com o item alvo.

Em primeiro lugar, o efeito estatisticamente significativo de similaridade morfofonológica mostra que, independentemente da variabilidade e da frequência, a alta similaridade fonológica encoraja a produtividade construcional, ao passo que a baixa similaridade a desencoraja. À luz da proposta de Suttle e Goldberg (2011) em termos de *coverage*, isso se daria justamente porque, nos casos de alta similaridade, a categoria composta pelos itens atestados e o item alvo apresenta alto *coverage, diferentemente do que ocorre quando a similaridade é baixa*. Para compreender esse ponto, observemos a figura abaixo:

Figura 8 – Efeitos da similaridade

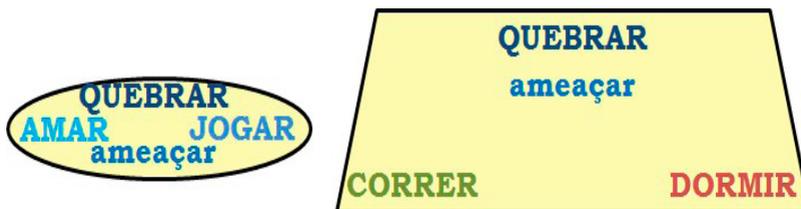


Fonte: O autor.

Como se pode ver, a primeira categoria, em que a similaridade morfofonológica entre o item alvo (*amar*, em minúsculas) e um dos itens atestados (*em maiúsculas*) é alta, é mais densamente povoada do que a segunda, em que a similaridade é baixa. Podemos observar, portanto, que a alta similaridade morfofonológica está associada a alto *coverage* e, conseqüentemente, alto grau de produtividade, enquanto a baixa similaridade está associada a baixo *coverage* e, como consequência, a um grau mais baixo de produtividade.

Em segundo lugar, a possível interação entre variabilidade e similaridade morfofonológicas sugere que, a depender da similaridade, a variabilidade morfofonológica pode ou não favorecer a extensibilidade de uma construção. Isso se dá da seguinte forma: quando a similaridade é alta, a baixa variabilidade favorece a extensibilidade em um grau maior do que a alta variabilidade o faz, já que esta produz baixo *coverage morfofonológico*, ao passo que aquela produz um alto *coverage*, como é possível observar na figura 9 a seguir:

Figura 9 – Efeitos da variabilidade nas condições de alta similaridade



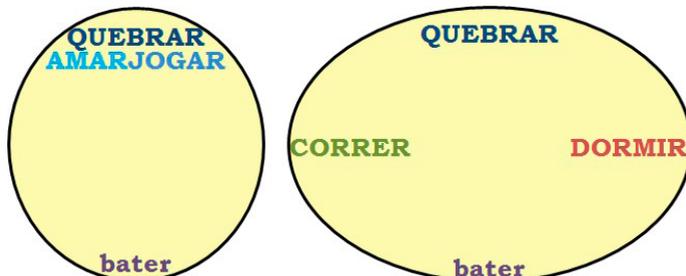
Fonte: O autor.

A categoria à esquerda, em que há alta similaridade morfofonológica entre o item alvo (*ameaçar*) e os itens atestados (*QUEBRAR*, *AMAR*, *JOGAR*) é mais densamente povoada do que a categoria à direita. Embora também apresente alta similaridade entre o item alvo (*ameaçar*) e um dos itens atestados (*QUEBRAR*), exibe também alta variabilidade entre os itens atestados (*QUEBRAR* x *CORRER* x *DORMIR*).

Por outro lado, quando a similaridade é baixa, a variabilidade não parece ser um fator relevante para a produtividade construcional. Em termos de *coverage*, isso pode ser interpretado como

o resultado do fato de que essa combinação produz, em ambos os casos, baixa cobertura categorial, como se pode ver na figura 10:

**Figura 10** – Efeitos da variabilidade nas condições de baixa similaridade

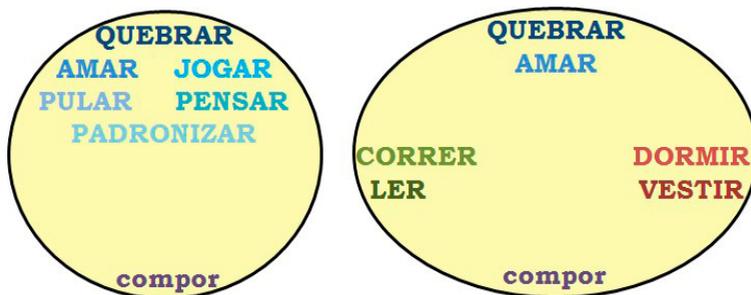


Fonte: O autor.

Nas duas categorias acima, independentemente do fato de a variabilidade morfofonológica entre os itens atestados ser baixa (categoria à esquerda) ou alta (categoria à direita), há baixa *coverage* e, portanto, baixo grau de produtividade construcional, uma vez que a similaridade morfofonológica entre o item alvo (*ameaçar*, nos dois casos) e os itens atestados (QUEBRAR, AMAR, JOGAR e QUEBRAR x CORRER x DORMIR) é baixa, o que confirma o possível efeito de interação entre a variabilidade e a similaridade morfofonológicas, descrito acima.

Finalmente, a também possível interação entre variabilidade morfofonológica e frequência de tipo revela que, de maneira análoga ao que ocorre na interação entre variabilidade e similaridade, o efeito da variabilidade sobre a produtividade construcional depende de a frequência ser alta, com a baixa variabilidade (à esquerda) produzindo, neste caso, maior *coverage* se comparado à alta variabilidade (à direita) e favorecendo, portanto, a extensibilidade da construção para novos itens, como se pode observar a seguir:

**Figura 11** – Efeitos da variabilidade nas condições de alta frequência



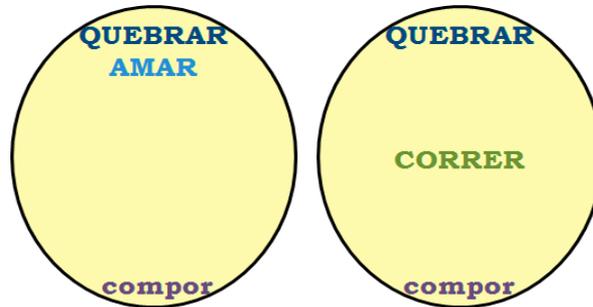
Fonte: O autor.

Obseva-se acima que a categoria à esquerda apresenta maior *coverage* que a categoria à direita. Isso se explica justamente pelo fato de que a primeira apresenta baixa variabilidade morfofonológica entre os itens atestados (QUEBRAR, AMAR, JOGAR, PULAR, PENSAR,

*PADRONIZAR*), diferentemente da segunda, em que a variabilidade é alta (*QUEBRAR, AMAR x CORRER, LER x DORMIR, VESTIR*).

No entanto, quando a frequência é baixa a variabilidade não parece ser um fator relevante para a produtividade construcional, já que produz, seja ela alta ou baixa, uma baixa cobertura categorial. Veja na figura 12:

**Figura 12** – Efeitos da variabilidade nas condições de baixa frequência



Fonte: O autor.

Na figura acima podemos observar que, quer a variabilidade morfofonológica entre os itens atestados seja baixa (*QUEBRAR, AMAR*), *quer seja alta* (*QUEBRAR x CORRER*), o grau de cobertura categorial é baixo nos casos de baixa frequência. O que, portanto, não favorece a produtividade construcional.

Em conformidade com Suttle e Goldberg (2011), sugerimos que os três fatores controlados – frequência de tipo, variabilidade e similaridade, com a diferença de que, para o nosso estudo, a variabilidade e a similaridade são morfofonológicas e não semânticas – podem ser reduzidos ao (ou interpretadas em termos do conceito de *coverage*, de modo que o alto *coverage* aponta para um conseqüente aumento da produtividade construcional, ao passo que o baixo *coverage* reduz a possibilidade de se estender a construção para novos usos.

## CONSIDERAÇÕES FINAIS

Neste trabalho buscamos contribuir com a discussão acerca da produtividade construcional. Este tema já bastante presente na literatura linguística em geral e na literatura em GCBU em particular. Verificamos, através deste estudo, um efeito claro da similaridade morfofonológica sobre a produtividade. Além disso, verificamos, nos resultados do experimento, algumas indicações de efeitos de interação entre variabilidade e similaridade morfofonológicas e entre frequência de tipo e variabilidade morfofonológica, ainda que essas interações não tenham alcançado significância estatística. Em resumo, o resultado mais robusto deste trabalho é o de que a existência de similaridade morfofonológica entre o uso alvo e usos

pré-existentes aumenta a confiança do falante em relação à extensibilidade de uma dada construção gramatical.

Apesar de a produtividade ser um fenômeno bastante estudado, não havia, ainda, resultados que apontassem para o impacto de fatores morfofonológicos sobre a extensibilidade de uma construção. Com efeito, a maior parte dos trabalhos construcionistas que abarca este tema tem dado preferência aos critérios semânticos. Nesse sentido, ao nos debruçarmos sobre questões que envolvem, para além do significado, a relação entre estrutura formal e som, nossa pesquisa contribui para o conhecimento de questões pouco ou ainda não investigadas, trazendo resultados inéditos.

## REFERÊNCIAS

- BARÐDAL, J. **Productivity**: Evidence from case and argument structure in Icelandic. Amsterdam: John Benjamin Publishing Company, 2008.
- BOYD, J.; GOLDBERG, A. E. **Learning what NOT to say**: The role of statistical preemption and categorization in a-adjective production. *In: Language*, 87, 1, pp. 55–83.
- BYBEE, J.; EDDINGTON, D. A usage-based approach to Spanish verbs of ‘becoming’. *In: Language*, 82, 2006, pp. 323–355.
- FILLMORE, C. J. Syntactic intrusions and the notion of grammatical construction. *In: Proceedings of the 11th Annual Meeting of the Berkeley Linguistics Society*, 1985, pp. 73-86.
- FILLMORE, C. J.; KAY, P.; O’CONNOR, C. **Regularity and idiomacity in grammatical constructions**: the case of let alone. *In: Language*, 63, 3, 1988, pp. 501-538.
- GOLDBERG, A. **Explain me this: Creativity, competition, and the partial productivity of constructions**. Princeton: Princeton University Press, 2019.
- LAKOFF, G. **Women, fire and dangerous things: what categories reveal about the mind**. Chicago: University Press, 1987.
- SUTTLE, L.; GOLDBERG, A. E. The partial productivity of constructions as induction. *Linguistics*, 49, 6, 2011. pp. 1237-1269.
- TRAUGOTT, E.; TROUSDALE, G. **Constructionalization and constructional changes**. New York: Oxford University Press, 2013.

Recebido para publicação em: 29 set. 2023.

Aceito para publicação em: 4 abr. 2024.