

A ALTERNÂNCIA DO DIMINUTIVO –INHO/-ZINHO NO PORTUGUÊS BRASILEIRO: UM ENFOQUE VARIACIONISTA

Myrian Azevedo de FREITAS*
Maria Fernanda M. BARBOSA**

- **RESUMO:** Neste trabalho, investiga-se a alternância do diminutivo -inho/-zinho, delimitando os fatores que determinam a escolha do falante por um ou outro formativo no Português Brasileiro. Mais especificamente, a partir dos pressupostos da Sociolinguística Variacionista, pretende-se analisar a formação produtiva do diminutivo a fim de constatarmos se temos um único sufixo diminutivo ou se se trata de dois processos distintos que dão origem, respectivamente, às formações X-inho e X-zinho. Neste estudo, utilizou-se a técnica de análise de variância multifatorial (ANOVA *Multi Way*) para medidas repetidas com intervalo de confiança de 95%, a fim de avaliar se o efeito dos fatores sociais (gênero, dialeto regional e faixa etária) é determinante na seleção do(s) sufixo(s) em análise e se há ou não convergência entre as formas diminutivas –inho e/ou -zinho, medidas ao nível de atitudes dos respondentes por um teste de aceitabilidade. Os resultados revelam que há evidência para admitirmos a existência de duas formas distintas no Português Brasileiro, -inho e –zinho, e indicam que há uma interação entre os fatores dialeto regional, gênero e faixa etária na seleção dos referidos sufixos diminutivos.
- **PALAVRAS-CHAVE:** Sufixos diminutivos no Português Brasileiro. Formação de palavras. Alternância. Variação linguística.

Introdução

Sendo um dos processos de construção de palavras extremamente produtivos no Português Brasileiro (doravante PB), as formações diminutivas X-inho e X-zinho já foram fonte de muitas observações por sua peculiaridade na língua portuguesa, seja pelos estudos de cunho linguístico, seja pela tradição gramatical. De fato, no processo de formação de diminutivos por sufixação, a(s) forma(s) –inho/ -zinho podem anexar-se a quase todas as classes de palavras do léxico, exceto os artigos.

Villalva (2000), ao tratar das formas avaliativas e Z-avaliativas, como prefere denominar as formações X-inho e X-zinho, afirma que esses sufixos apresentam

* UFRJ – Universidade Federal do Rio de Janeiro. Faculdade de Letras. Rio de Janeiro – RJ – Brasil. 21941-917 – myrfreitas@ufrj.br

** UFRJ – Universidade Federal do Rio de Janeiro. Faculdade de Letras. Rio de Janeiro – RJ – Brasil. 21941-917 – fernanda136@gmail.com

características específicas. Diferentemente da maioria dos morfemas derivativos e flexivos, tais formas não alteram a categoria sintática e nem as propriedades morfossintáticas ou morfossemânticas da base. Em outras palavras, o acréscimo de tais primitivos morfológicos não muda a classe da base, nem atribui gênero ([[bonito]_{Adj}inho]_{Adj}, [[cama]_sinha]_s), ao contrário do que ocorre com a maioria dos sufixos, como –ção e –mento, por exemplo ([[nacionaliza]_vção]_s, [[estaciona]_vmento]_s). A essas características, somam-se as peculiaridades de que sufixos avaliativos adjungem-se a radicais (sapatinho) enquanto que os Z-avaliativos anexam-se a palavras (lençoizinhos).

Com efeito, ao examinarmos minuciosamente as formas –inho / –zinho, é conveniente levantarmos as seguintes questões:

- a) Trata-se apenas de um único sufixo diminutivo ou devemos admitir a existência de dois sufixos distintos?
- b) A seleção de um ou outro elemento mórfico é condicionada somente por motivações estruturais ou é possível admitir que os fatores sociais exercem influência na escolha do falante? Em que medida se verifica ou não essa interação?

Note-se que não é tarefa fácil explicitar as regras que definem a preferência do falante por um ou outro elemento mórfico, dada a amplitude de uso e a vasta recorrência na língua. Assim, pretende-se analisar a formação produtiva do diminutivo no PB com vistas a elucidar as questões supracitadas.

O texto apresenta-se estruturado da seguinte maneira: na segunda seção, tem-se as principais descrições sobre a alternância do diminutivo –inho/-zinho na literatura. Na terceira seção, apresenta-se um panorama da Sociolinguística Variacionista. Na quarta seção, é especificada a metodologia empregada na elaboração do instrumento de medição utilizado para captação de dados. A quinta seção descreve os resultados obtidos no teste de aceitabilidade e sua significância estatística. Na sexta seção, identifica-se e avalia-se a modelagem estatística aplicada à massa de dados a partir da aceitabilidade das formas diminutivas –inho e/ou –zinho. Em seguida, apresentam-se as hipóteses que norteiam este estudo. Posteriormente, mostram-se os resultados acerca do fenômeno em estudo. Por fim, têm-se as considerações finais.

O diminutivo –inho/-zinho na literatura

Nos capítulos XXXIX – XLII, Fernão de Oliveira (1536) considera que os diminutivos fazem parte do que denomina de dições tiradas (palavras derivadas) em oposição à dições primeiras (palavras primitivas), referindo-se que “[...] é regra geral que [...] os diminutivos acabem em inho ou inha como mocinho, mocinha” (OLIVEIRA, 1536, p.61).

Na *Gramática da Língua Portuguesa*, Barros (1540) afirma que o diminutivo apresenta uma noção de pequenez, e não faz nenhuma menção direta às formas –inho e –zinho, somente as exemplifica:

Nome diminutivo é aquele que tem alguma diminuição do nome principal de onde se derivou como: de homem, homenzinho; de mulher, mulherzinha; de moço, mocinho; de criança, criancinha. E outros muitos que se formam e acabam em diferentes terminações, mais por vontade do povo que por regra de boa Gramática. (BARROS, 1540, p.304).

Também Barbosa (1822) menciona –inho e –zinho como diminutivos do português. No entanto, o autor acrescenta que há a inserção do segmento /z/ a fim de se evitar o hiato nos vocábulos terminados em ditongo:

Os *Diminutivos* são os que mudando a terminação de seus primitivos, lhes diminuem mais, ou menos a significação. Os que diminuem menos, acabam ordinariamente, os masculinos em *ete, ote, oto, (...)* e os femininos, em *eta, ota, agem, ilha, (...)*. Os que diminuem mais, acabam ou em *inho, inha*, quando os primitivos terminam em vogal ou consoante, como *Filhinho, Filhinha, Mulherinha, Rapazinho*; ou em *zinho, zinha*, quando os primitivos terminam em ditongo, como *Homemzinho, Leãozinho, Paizinho, Mãezinha*. O z eufônico faz-se necessário na derivação destes diminutivos, para evitar o hiato, nascido do concurso de três vogais. Porém, quando o mesmo se emprega sem esta necessidade nos que não acabam em ditongo; parece fazer sua diferença nos mesmos diminutivos, como se vê nestes dois *Mulherinha, Mulherzinha*. (BARBOSA, 1822, p.120-121, grifo do autor).

Observe-se que Barbosa (1822) parece referir-se à existência de uma só forma, -inho. No entanto, o autor revela que a variação entre os diminutivos ocorre em palavras terminadas em consoante como, por exemplo, mulherinha/mulherzinha. Mais adiante, o autor justifica a alternância entre os diminutivos, afirmando que “seja como for, o que é certo é, que a nossa Língua é muito rica neste gênero de derivação” (BARBOSA, 1822, p.120). Em seguida, menciona ainda, em sua *Grammatica Philosophica da Língua Portuguesa*, que a “significação de um primitivo [que] tome um argumento enorme, e dele vá descendo gradualmente até ao extremo contrário de pequenez, como se pode ver nos derivados (...) Mulherão, Mulherona, Mulherinha, Mulherzinha” (BARBOSA, 1822, p.121), revelando que há uma gradação de tamanho entre as formas e atribui a –zinho(a) um grau dimensional sutilmente menor que –inho(a).

Por outro lado, Freire (1842), em *Reflexões sobre a Língua Portuguesa*, também ressalta a noção dimensional dos diminutivos e apresenta informações sobre duas formas vigentes na língua portuguesa, –inho e –zinho. O autor afirma

que o emprego de -zinho só ocorre para evitar o hiato ou quando o vocábulo termina em consoante, considerando a existência de dois sufixos diminutivos distintos:

Os diminutivos são os que mudando a terminação de seus primitivos lhes diminuem mais ou menos a significação: acabam em *inho*, ou *inha*, como de *peixe*, *peixinho*; de *casa*, *casinha*; em *zinho* quando os primitivos rematam em ditongo, para se evitar o hiato pelo concurso de três vogais; v. gr., de *leão*, *leãozinho*, de *pai*, *paizinho*; igual terminação têm os nomes que acabam em consoante, posto que algumas exceções se notam em que há dois diminutivos da mesma palavra por diversa terminação: exemplo, de *casa* também há *casinhota*, de *peixe* também há *peixezinho*. Há osfindos em *ete*, como *pobrete*, de que temos igualmente *pobrezinho*: em *eta*, *ote*, e *ota*, exemplo, *ilheta*, *ilhote*, *ilhota*, que todos significam o mesmo, podendo alémdelos juntar-se *ilhana* mesma aceção: de *arca* se tem feito *arquinho*, *arqueta*, *arquilha*, e *arquetemasculino*. Vemos que os há em *itha*, como de *cama*, *camilha*, de que é mais vulgar *caminha*: raros são os emulo, exemplo *perdigoto*. (FREIRE, 1842, p.164, grifo do autor).

Note-se que Freire (1842), apesar de identificar os contextos morfológicos de ocorrência de -inho e -zinho, também admite que a alternância entre os diminutivos ocorre para uma mesma palavra sem motivação aparente, como nos pares *peixinho/peixezinho* e *pobrinho/pobrezinho*, por exemplo. Observe-se que, ao contrário de Barbosa (1822), seu antecessor, o autor considera que não há a inserção do segmento consonantal /z/ para se evitar o hiato, mas sim que estamos diante de dois sufixos diminutivos na língua, -inho e -zinho.

Numa perspectiva da linguística histórica, Skorge (1957) assinala que os diminutivos -inho e -ito são os sufixos mais empregados e mais expressivos do português. Para explicar a presença do segmento consonantal /-z/, a autora, que considera tratar-se de um infix, afirma que é mais “correto recorrer ao -c- do latim nos substantivos *navicella*, *avicella* e **domnicellus* que podiam ser também a proveniência do infix -c- do espanhol e do italiano” (SKORGE, 1957, p.69). Mais adiante, Skorge (1957) acentua ainda que as regras para determinar a ocorrência de -inho ou -zinho são escassas e não apresentam respostas satisfatórias. No entanto, em relação à alternância entre os diminutivos -inho e -zinho, a autora afirma que o emprego de um ou outro formativo parece estar ligado ao ritmo da frase e destaca que “entre o povo emprega-se muito -inho. Os portugueses cultos tendem a empregar diminutivos em -zinho” (SKORGE, 1957, p.52). Por último, ao delimitar a distribuição de -inho e -zinho, Skorge (1957) afirma que somente os substantivos monossílabos admitem a forma -zinho. Nas demais classes de palavras, ocorrem flutuações que admitem ora -inho ora -zinho, alternando-se de acordo com a ‘vontade do falante’.

Ao analisar a formação do diminutivo no PB, Câmara Jr. (1975) afirma que –zinho é variante alomórfica de –inho e argumenta que o uso de –zinho é obrigatório diante de palavras que terminam em vogal tônica, sendo considerado como um caso de derivação por justaposição.

Leite (1974) propõe uma abordagem gerativa para o acento em português, atribuindo o traço [+acento] às vogais. A autora afirma que, no léxico, o radical não porta marca de acento e formula diferentes condições para aplicação de regras intrinsecamente ordenadas. Leite (1974, p.112) considera que “–zinho é anexado a formas terminadas em vogal enquanto –inho é acrescido a formas acabadas em consoante.”

Na proposta de Lee (1995), numa abordagem pela Fonologia Lexical Prosódica (INKELAS, 1989, 1993), o léxico do PB apresenta dois estratos ordenados: nível I (α), que comporta a derivação, alguns processos de composição com acréscimo de sufixos derivacionais ([rádio-tax]ista), [puxa-saqu]ismo) e a flexão irregular; e nível II (β), que abarca a formação produtiva (sufixos –inho e –zinho, advérbio –mente e grau –íssimo) e a flexão regular.

Lee (1995) informa que, em geral, o morfema –inho é anexado a formas não-verbais (nomes e adjetivos/advérbios) contendo vogal temática, como em casa/casinha, bonita/bonitinha, perto/pertinho. No entanto, o morfema –zinho é:

- a) adjungido a um não-verbo com vogal temática inexistente (atemático) como, por exemplo, café/cafezinho, flor/florzinha;
- b) anexado a palavras proparoxítonas e vocábulos terminados em sílaba pesada, como em lâmpada/lampadazinha, judeu/judeuzinho, mar/marzinho.

Contudo, o autor reconhece que ocorre variação entre os sufixos na fala cotidiana como, por exemplo, em lampadazinha ~ lampadinha, facilzinho ~ facinho, xicarazinha ~ xicrinha, etc.

Ao analisar o discurso dos pescadores artesanais, Tavares Jr. (1999) menciona que podem ser atribuídos aos sufixos –inho e –zinho mais três valores semânticos além do diminutivo: afetividade (tem uma escolinha ali na ilha), pejoratividade (só tem um manguezinho muito pouco) e intensificação (põe no sol para ele ficar bem sequinho). O autor informa ainda que, em uma mesma palavra, pode coexistir mais de uma ocorrência dos sufixos como, por exemplo, nas sentenças ‘esse é um passarinhozinho que canta aí’, ‘é uma tábua purinhazinha’, ‘aquele limo fininhozinho que junta na pedra’, etc.

No Português Europeu, Villalva (2000) assume que –inho e –zinho são sufixos diminutivos distintos, denominando-os respectivamente de formas avaliativas e Z-avaliativas, dado que as formas de base a que se anexam apresentam diferentes propriedades.

Segundo Villava (2000), o principal contraste entre os diminutivos –inho e –zinho está associado a categoria morfológica da base: “esta análise permite distinguir os sufixos avaliativos dos Z-avaliativos – dado que os primeiros se associam a radicais e formam novos radicais, enquanto os segundos se associam a palavras e formam palavras” (VILLALVA, 2000, p.316). Deste modo, a sufixação avaliativa seleciona radicais, enquanto os Z-avaliativos anexam-se a palavras.

O segundo contraste reside no fato de que as formas avaliativas apresentam um único acento lexical (sapatínho, casinha) enquanto as Z-avaliativas têm dois acentos lexicais, um referente à base (acento tônico), e outro ao próprio sufixo (acento subtônico) como em cafézínho, chapéuzínho, pézínhoetc.

O terceiro contraste refere-se ao fato de a flexão de gênero e número da base e do sufixo Z-avaliativos ser independentemente estabelecida. Deste modo, as flexões de gênero e número apresentam-se duplicadas nas formas em -zinho, aparecendo antes e depois do sufixo (livrozinho, quadrozinho, anelzinho > aneijinhos, marzinho > marezinhos), ao passo que as construções X–inho apresentam uma única marca de gênero e de plural (lojinha, cervejinha, beijo > beijinhos, lata > latinhas).

O quarto contraste que distingue a aplicação entre as formas está relacionado ao número de sílabas da base: “Com efeito, a um maior número de sílabas corresponde uma preferência pela sufixação Z-avaliativa, enquanto que os dissílabos recorrem à sufixação avaliativa” (VILLALVA, 2000, p.332). O terceiro contraste relevante refere-se ao fato de que:

[...] contrariamente às formas avaliativas que preservam o índice temático da base, as formas Z-avaliativas determinam a classe temática onde se integram, por concordância com o valor de gênero da forma de base e de acordo com a sua realização não-marcada, ou seja, tema em –o para as formas masculinas e tema em –a para as formas femininas. (VILLALVA, 2000, p.321).

A fim de explicar se –inho e –zinho são ou não variantes foneticamente condicionadas, a autora argumenta que o uso das formas z-avaliativas é obrigatório diante de:

- a) vocábulos que terminam com vogal (oral ou nasal), como em pontapezinho / *pontapeinho, alibizinho / *alibiinho, atunzinho / *atuninho, rãzinha / *rãinha, etc.;
- b) palavras em que a forma de base termina em vogal (oral ou nasal) átona como, por exemplo, taxizinho / *taxinho, alunzinho / *aluninho, videozinho / *videoinho, etc.;
- c) palavras com ditongo (oral ou nasal) tônico, como nos exemplos cacauzinho / *cacauinho, veuzinho / *veuinho, cançãozinha / *cançãoinha, etc.;

- d) palavras com ditongo átono, como em ordenzinha / *ordeninha, bençãozinha / *bençãoinha, trolezinho / *troleinho, etc.;
- e) palavras acabadas em consoante, como em hifenzinho / *hifeninho, sloganzinho / *sloganinho, smokingzinho / *smokinginho, etc.

Do ponto de vista formal, Villalva (2000) conclui que a distribuição dos sufixos avaliativos e Z-avaliativos é condicionada pela categoria morfológica da base. No entanto, “a escolha entre uns e outros é condicionada por múltiplos fatores como a recuperabilidade da forma de base, o seu número de sílabas, variação dialectal, índices de ocorrência e lexicalização” (VILLALVA, 2000, p.337).

Em *Formação e Classes de Palavras no Português do Brasil*, Basílio (2004) afirma que –inho e –zinho parecem ser elementos complementares, uma vez que –zinho é utilizado em ambientes fonológicos em que –inho não é usado.

A autora enfatiza ainda que o sufixo –inho se incorpora integralmente à fonologia da palavra base, enquanto –zinho apresenta peculiaridades semelhantes aos advérbios terminados em –mente, ressaltando que “a adição de –zinho mantém a linha geral da acentuação tônica da palavra base, como vemos: pálida – palidazinha, pá – pazinha, tatu – tatuzinho, túnel – tunelzinho” (BASÍLIO, 2004, p.72), mas transforma-se em acentuação subtônica. Além disso, Basílio considera que “a característica mais desconcertante das formações em –zinho é o fato de que a adição de –zinho não impede a flexão de gênero e número na palavra base (indiozinho/indiazinha, balãozinho/balõezinhos)” (BASÍLIO, 2004, p.72). Por fim, a autora conclui que, na formação do diminutivo, temos “dois elementos formadores, -inho e –zinho, cuja ocorrência é parcialmente complementar, mas cujo estatuto morfológico é radicalmente diferente” (BASÍLIO, 2004, p.72).

Em *Flexão & derivação em português*, Gonçalves (2005) menciona que “é possível adicionar –(z)inho a praticamente todos os nomes da língua (...), anexando-se a pronomes, advérbios, numerais e interjeições” (GONÇALVES, 2005, p.40). O autor ressalta ainda que “afixos que, em princípio, disputariam a mesma posição na cadeia sintagmática podem aparecer juntos na mesma palavra” (GONÇALVES, 2005, p.82). De fato, por vezes, um mesmo vocábulo aparece associado a dois afixos de aumentativo ou dois sufixos de diminutivo, constituindo-se em um processo recursivo: “a recursividade também pode ser usada com fins expressivos. São perfeitamente possíveis formas como ‘livrãozão’, com dois afixos aumentativos e ‘vidinhazinha’ com dois de diminutivo” (GONÇALVES, 2005, p.83-84).

Em *O diminutivo e suas demandas*, Bisol (2010) considera que há apenas um morfema e não dois, mencionando que “existe apenas um morfema, que é –inho, o qual se reveste de uma consoante epentética para satisfazer exigências estruturais, manifestando-se como –zinho” (BISOL, 2010, p.59).

No que tange a variação, Bisol (2010, p.72) admite que esta “circunda-se aos nominais temáticos em que se impõe a prevalência da forma original –inho,” considerando que “em nominais temáticos, há casos em que a epêntese é motivada” (BISOL, 2010, p.72). Assim, o primeiro caso listado por Bisol (2010) refere-se à flexão de gênero que se apresenta tanto na forma de base como no morfema diminutivo, ou seja, “haverá redundância de informação, o que, em nome da simplicidade tende a ser evitado. Não há erros, mas redundância de informação” (BISOL, 2010, p.72), citando exemplos como borboletinha ~ borboletazinha, patinho ~ patozinho, garotinha ~ garotozinha. No entanto, a autora ressalta que há exceções:

Todavia, há casos de disparidade entre VT [vogal temática] e gênero da forma de base e do diminutivo. Nesses a consoante /z/ é introduzida para preservar a vogal da base, manifestando-se, dessa forma, ambas as vogais, a da base internamente e a de DIM [diminutivo] que porta o gênero na posição que lhe é destinada. Isso acontece com nominais de gênero masculino com VT /a/ [o aroma → o aromazinho], com nominais do gênero feminino com VT /o/ [a contralto → a contraltozinha] e com nominais pertencentes à categoria comum de dois [o/a pianista → o pianistazinho, a pianistazinha]. (BISOL, 2010, p.72).

A autora considera que os casos de disparidade pertencem aos nominais que apresentam vogal temática em /e/, sendo destituídos “de qualquer relação com gênero, que abrem as portas para a variação analógica: verdinho ~ verdezinho, paredinha ~ paredezinha, correntinha ~ correntezinha” (BISOL, 2010, p.73). Assim, a autora conclui que “a variação em nominais temáticos está comprometida com o gênero gramatical” (BISOL, 2010, p.73).

O segundo caso ilustrado por Bisol (2010) refere-se ao princípio do contorno obrigatório (doravante OCP), que proíbe segmentos idênticos adjacentes. Deste modo, a autora postula que OCP é ativado e o segmento consonantal /z/ é inserido a fim de contornar a adjacência de segmentos que apresentam as mesmas propriedades fônicas como, por exemplo, em vinho → vinhozinho / *vinhinho, pinho → pinhozinho / *pinhinho, linho → linhozinho / *linhinho, etc. No entanto, a autora admite que a variação em nominais terminados em hiato constitui um dilema na análise do diminutivo, visto que “a parte final da base prosódica não oferece, como nos demais nominais temáticos, uma consoante para *onset*” (BISOL, 2010, p.73). Assim, de um lado, “via apagamento de VT [vogal temática], *ka.no.a>ka.no.i.ña*, o hiato da base é desfeito e o de DIM [diminutivo] permanece” (BISOL, 2010, p.73). Por outro lado, “via epêntese, *ka.no.a>ka.no.a.zi.ña*, o hiato da base é preservado para resolver o de DIM” (BISOL, 2010, p.73). Contudo, a autora constata que há exceções à regra e afirma que:

as ocorrências mais frequentes indicam que DIM tende a preservar o hiato do *input* e resolver o que lhe diz respeito, optando pela epêntese, mas a variação não fica de toda excluída, como em *atoa > atoazinha ~ atoinha*, *canoas > canoazinha ~ canoinha*, *garoa > garoazinha ~ garoinha*. (BISOL, 2010, p.73).

Por último, o terceiro caso mencionado por Bisol está relacionado ao fato de que, nos diminutivos, “a base que contém minimamente duas sílabas exibe pé binário de cabeça à esquerda” (BISOL, 2010, p.74). Deste modo, o acento secundário, quando herdado do principal, desloca-se para a sílaba imediatamente anterior a fim de evitar uma colisão acentual com retração de acento. No entanto, a autora constata que “a exceção está no grupo das proparoxítonas, que foge ao padrão geral [pé binário de cabeça à esquerda], mostrando um dátilo, o qual tende a ser preservado por DIM” (BISOL, 2010, p.74). Assim, a exceção encontra-se em dados como, por exemplo, (lâm.pa.da) → (lâm.pa.da.)(zí.nha) ~ (lâm.pa.)(dí.nha), (cá.te.dra) → (cà.te.dra.)(zí.nha) ~ (cà.te.)(drí.nha), etc.

Note-se que Bisol (2010) admite a variação existente em nominais temáticos, mas considera que “há casos que apontam para o controle de certos fatores, como redundância de gênero, hiato duplicado, acento marcado e OCP que motivam ou restringem a presença de consoante epentética” (BISOL, 2010, p.75).

Finalmente, Bisol (2010, p.82) chega à seguinte generalização: “O Diminutivo, cuja forma canônica é *-inho*, exige onset e preserva os elementos da base (*input*) e do *output* que são relevantes para a sua estruturação como palavra fonológica.”

Em suma, no que se refere às formas *-inho* e/ou *-zinho*, torna-se evidente a ausência de consenso entre os teóricos. Note-se que, de um lado, temos semelhanças semântica e fonética entre os formativos, fazendo com que muitos autores considerem a existência de uma única forma diminutiva, *-inho*. No entanto, haveria a inserção de um segmento consonantal epentético, consoante de ligação ou ainda um infixo que ocorre em determinados contextos, manifestando-se como *-zinho*. Para se considerar a possibilidade de *-z-* como infixo, seria necessário caracterizá-lo como um autêntico morfema, dotado de significação. Por outro lado, diante da flexão nominal, observa-se que a distribuição e o comportamento de tais formativos levam outros autores a interpretá-los como entidades morfológicas distintas.

A sociolinguística variacionista

A perspectiva teórica da Sociolinguística Variacionista ou Laboviana relaciona aspectos variáveis sistemáticos da linguagem humana a fatores sociais em uma dada comunidade. Neste sentido, língua e sociedade estão estreitamente

interligadas. Para Labov (1972), o objeto da pesquisa sociolinguística é a língua vernácula, sendo descrita como a língua falada na vida cotidiana pelos membros de uma comunidade.

Na década de 60, ao pesquisar o inglês falado na ilha de Martha's Vineyard, a respeito da variação fonológica da vogal núcleo dos ditongos /ay/ (ex. right) e /aw/ (ex. house), Labov (1972) notou fortes influências sociais decorrente da invasão de veranistas ingleses. Assim, o autor concluiu que os falantes que apresentavam uma atitude mais positiva em relação à ilha, os moradores locais, centralizavam mais os ditongos, variante não-padrão e estigmatizada, adotando uma pronúncia típica da ilha: um schwa [əu] e [əy], enquanto os veranistas adotavam uma pronúncia mais geral, considerada inovadora e de prestígio. Logo, Labov (1972) evidenciou o fato de a língua ser considerada um fator de identidade social. Com isso, o autor comprovou a possibilidade de a variação linguística ser objeto de sistematização, destacando o papel primordial dos fatores sociais na explicação da variação linguística.

Diante da heterogeneidade dos fenômenos linguísticos, Labov (1972) estabeleceu um modelo de descrição e análise linguística que considera a influência dos fatores sociais atuantes na língua, sendo conhecido como teoria da variação linguística.

A língua então passa a ser estudada como um sistema heterogêneo em constante processo de mudança, relacionando-se diretamente ao meio social em que está inserida. Em outras palavras, não existe uma língua homogênea, a linguagem humana muda com o tempo. Deste modo, Labov (1972) considera que, quando analisamos o contexto social em que a língua é utilizada, muitos elementos da estrutura linguística implicam na variação sistemática que reflete tanto a mudança no tempo quanto nos processos sociais:

[...] os procedimentos da linguística descritiva fundamentam-se na concepção de que a língua é um conjunto estruturado de normas sociais. No passado, foi útil considerar que tais normas eram invariáveis e compartilhadas por todos os membros da comunidade linguística. Entretanto, as análises mais detalhadas do contexto social em que a língua é utilizada vieram demonstrar que muitos elementos da estrutura linguística estão implicados na variação sistemática que reflete tanto a mudança no tempo quanto os processos sociais extralinguísticos. (LABOV, 1972, p.241).

No estudo da variação, as alternâncias de uso sofrem influências de fatores estruturais e sociais. Neste sentido, o uso de formas linguísticas é motivado pelas diferenças entre os falantes e pela própria heterogeneidade da língua. Portanto, sua alternância pode ser estatisticamente previsível, ou seja, torna-se passível de ser descrita e analisada cientificamente.

Partindo da concepção da relação entre língua e sociedade e de que a língua falada é variável e heterogênea, os objetivos principais da Sociolinguística Laboviana são:

- a) analisar e legitimar as variantes utilizadas numa comunidade de fala (LABOV, 1975);
- b) entender a relação entre variação e mudança linguística (WEINREICH; LABOV; HERZOG, 1968).

Contudo, cabe ainda à análise sociolinguística explicitar a inserção ou encaixamento de uma variável no sistema de relações sociais e linguísticas de uma comunidade. Para isso, dispõe-se de métodos estatísticos aplicados a um conjunto de dados a fim de se extrair regularidades altamente ordenadas que governam a variação na comunidade. Neste sentido, Guy e Zilles consideram que “antes do advento da metodologia de quantificação, a variação linguística era considerada secundária, aleatória ou mesmo impossível de ser cientificamente apreendida” (GUY; ZILLES, 2007, p.73). No entanto, “o uso de métodos estatísticos tem permitido demonstrar o quão central a variação pode ser para o entendimento de questões como identidade, solidariedade ao grupo local, comunidade de fala, prestígio e estigma, entre tantas outras” (GUY; ZILLES, 2007, p.73).

Materiais e métodos

A fim de compreendermos a percepção de falantes do PB diante da alternância do diminutivo –inho/ -zinho, foi realizado um teste de aceitabilidade. Este teste de julgamento foi formulado com o objetivo de se identificar prováveis reações atitudinais de falantes do português brasileiro diante da variação presente na escolha de uso dos formativos –inho e/ou -zinho.

Como parte preliminar do planejamento experimental do teste de aceitabilidade, procedeu-se à definição de escalas, em que se optou pela utilização de uma escala com pontuação. Adotou-se a escala de Likert (1932), de cinco pontos, modelo criado em 1932 para avaliar atitudes e hoje amplamente empregado para avaliações diversas, sendo atribuídos os valores: (1) para “perfeitamente possível”; (2) para “possível, mas não utilizaria”; (3) para “indiferente”; (4) para “improvável, raramente utilizaria”; e (5) para “não aplicável, nunca utilizaria”. A atribuição dos pontos à escala Likert (1932), em cada nível de resposta, procurou estabelecer uma ordenação das alternativas com intervalos equivalentes a fim de viabilizar a análise estatística. Deste modo,

admitimos que o questionário aplicado constitui-se em uma escala intervalar, visto que a distância entre as posições é a mesma e que tais posições, utilizadas para avaliar opiniões e atitudes, medem proporções que partem do mais favorável ao mais desfavorável. Dito de outro modo, além de os dados estarem dispostos numa ordem linear, do menor para o maior, o espaçamento igual dos níveis de resposta foi claramente indicado em uma escala visual apresentada no questionário aplicado.

Nosso instrumento de pesquisa foi composto por uma escala Likert (1932) com 10 sentenças referentes à variação detectada entre as formas diminutivas -inho, e/ou -zinho.

Em relação ao tamanho da amostra, visando evitar a discussão acerca do limite entre o que podemos considerar como uma amostra pequena ou grande, ou seja, buscando o limite do menor tamanho da amostra com melhor aderência à curva da distribuição normal, adotamos uma amostra composta por 40 informantes, sendo 17 homens e 23 mulheres, com faixa etária entre 18 e 26 anos, dialeto regional (carioca; fluminense; outro) e ensino superior incompleto.

Na análise das respostas, foi proposta uma escala de concordância em que o respondente indica seu grau de acordo ou desacordo para cada item. Neste caso, adotamos os seguintes valores para cada resposta:

+				-
[1] Concordo Plenamente	[2]	[3]	[4]	[5]Discordo Totalmente

Nesta escala, admitimos que o valor exatamente (3) seria considerado “indiferente” ou “sem opinião”, sendo o “ponto neutro”, equivalente aos casos em que os respondentes abstiveram-se de emitir uma opinião.

A partir da aplicação do teste de aceitabilidade, os dados obtidos foram analisados pelo programa EXCEL 14, que compõe o pacote *Office 2010* da *Microsoft Corporation*, com suporte ao pacote estatístico R, sendo caracterizado por REXCEL.

Resultados e análise do instrumento de medição

A partir das respostas fornecidas pelos entrevistados, apresentaremos, nas próximas seções, os resultados obtidos no teste de aceitabilidade e teceremos uma análise acerca do instrumento de medição aplicado neste estudo.

Faixa etária

Neste estudo, consideramos a faixa etária como uma variável qualitativa ordinal, dividida em intervalos mutuamente exclusivos e ordenados. Neste caso, a variável foi descategorizada e agrupada em intervalos fechados com 3 classes. Assim, os respondentes foram divididos, de acordo com sua idade, em três intervalos: {[18-20], [21-23], [24-26]}. Considerando a idade dos entrevistados, os dados revelam que 68% possuem de 18 a 20 anos, 15% encontram-se na faixa etária entre 21 e 23 anos e 17% estão entre 24 e 26 anos, conforme vemos na tabela 01.

Tabela 01 – Distribuição por faixa etária dos entrevistados

FAIXA ETÁRIA	Frequência Absoluta (f_i)	Frequência Relativa (fr_i) %
18 20 anos	28	70
21 23 anos	6	15
24 26 anos	6	15
TOTAL	40	100

Fonte: Elaboração própria.

Gênero

Neste trabalho, o gênero do respondente assume os atributos de uma variável qualitativa nominal. Deste modo, os entrevistados foram categorizados em função do gênero em: {Masculino, Feminino}.

Considerando o gênero dos entrevistados, temos 17 indivíduos do gênero masculino e 23 do gênero feminino, perfazendo um total de 40 respondentes.

Tabela 02 – Distribuição por gênero dos entrevistados

SEXO	Frequência Absoluta (f_i)	Frequência Relativa (fr_i) %
Masculino	17	42,5
Feminino	23	57,5
TOTAL	40	100

Fonte: Elaboração própria.

Dialeto regional

Para o teste de julgamento, definimos a variável dialeto regional em qualitativa nominal e foram considerados os seguintes valores: {Carioca, Fluminense, Outro}.

De acordo com o dialeto regional, os dados mostram que 83% dos entrevistados são cariocas, 13% são fluminenses e 5% situam-se em outras regiões do estado do Rio de Janeiro, como se observa na tabela 03.

Tabela 03 – Distribuição por dialeto regional dos entrevistados

DIALETO REGIONAL	Frequência Absoluta (f _i)	Frequência Relativa (fr _i) %
Carioca	33	83
Flumenense	5	13
Outro	2	5
TOTAL	40	100

Fonte: Elaboração própria.

Atitude dos respondentes

Considerando o julgamento efetuado pelos respondentes, avaliamos o grau de concordância ou discordância para cada questão a fim de medirmos a atitude do sujeito frente a cada item avaliado, conforme o exposto na tabela 04.

Com base nos resultados, pode-se notar que a maior tendência à concordância possui os escores mais altos concentrados nas respostas em que se atribuiu o grau 1 (posição altamente favorável) para os diminutivos –inho e –zinho respectivamente. Por outro lado, a forma –inhozinho apresenta escores mais elevados para as respostas de grau 5 (posição claramente desfavorável), revelando uma tendência à discordância quanto à utilização desse formativo. Neste caso, os dados mostram que os respondentes posicionaram-se positivamente em relação ao uso das formas –inho e –zinho, admitindo as duas possibilidades de uso no mesmo contexto, embora se observe uma pequena variação percentual entre os dois formativos em função de cada item avaliado. No entanto, os respondentes adotaram uma atitude mais conservadora diante da utilização da forma –inhozinho, revelando um alto índice de rejeição nos contextos em que se alternam as possibilidades de aplicação de –inho e -zinho.

Tabela 04 – Distribuição da atitude do respondente aos itens vinculados no uso dos sufixos diminutivos

ITENS / SUFIXOS	GRAUS DE ACEITABILIDADE														
	1			2			3			4			5		
	-Inho	-Zinho	-Inhozinho	-Inho	-Zinho	-Inhozinho	-Inho	-Zinho	-Inhozinho	-Inho	-Zinho	-Inhozinho	-Inho	-Zinho	-Inhozinho
Questão 1	35 = 88%	2 = 5%	0 = 0%	5 = 13%	15 = 38%	2 = 5%	0 = 0%	5 = 13%	2 = 5%	0 = 0%	10 = 25%	10 = 25%	0 = 0%	8 = 20%	26 = 65%
Questão 2	33 = 83%	6 = 15%	1 = 3%	4 = 10%	11 = 28%	3 = 8%	2 = 5%	2 = 5%	3 = 8%	1 = 3%	14 = 35%	3 = 8%	0 = 0%	7 = 18%	30 = 75%
Questão 3	13 = 33%	29 = 73%	1 = 3%	14 = 35%	5 = 13%	4 = 10%	5 = 13%	3 = 8%	3 = 8%	7 = 18%	3 = 8%	10 = 25%	1 = 3%	0 = 0%	22 = 55%
Questão 4	37 = 93%	2 = 5%	0 = 0%	2 = 5%	7 = 18%	1 = 3%	1 = 3%	8 = 20%	1 = 3%	0 = 0%	10 = 25%	6 = 15%	0 = 0%	13 = 33%	32 = 80%
Questão 5	25 = 63%	18 = 45%	1 = 3%	9 = 23%	12 = 30%	0 = 0%	6 = 15%	4 = 10%	1 = 3%	0 = 0%	5 = 13%	7 = 18%	0 = 0%	1 = 3%	31 = 78%
Questão 6	19 = 48%	27 = 68%	1 = 3%	10 = 25%	3 = 8%	1 = 3%	7 = 18%	4 = 10%	3 = 8%	1 = 3%	5 = 13%	5 = 13%	3 = 8%	1 = 3%	30 = 75%
Questão 7	14 = 35%	29 = 73%	0 = 0%	9 = 23%	6 = 15%	1 = 3%	6 = 15%	2 = 5%	1 = 3%	4 = 10%	3 = 8%	6 = 15%	7 = 18%	0 = 0%	32 = 80%
Questão 8	25 = 63%	20 = 50%	0 = 0%	8 = 20%	8 = 20%	2 = 5%	4 = 10%	5 = 13%	2 = 5%	0 = 0%	7 = 18%	5 = 13%	3 = 8%	0 = 0%	31 = 78%
Questão 9	4 = 10%	36 = 90%	0 = 0%	9 = 23%	2 = 5%	0 = 0%	8 = 20%	0 = 0%	1 = 3%	9 = 23%	2 = 5%	4 = 10%	10 = 25%	0 = 0%	35 = 88%
Questão 10	6 = 15%	30 = 75%	0 = 0%	10 = 25%	4 = 10%	1 = 3%	4 = 10%	4 = 10%	1 = 3%	6 = 15%	0 = 0%	6 = 15%	14 = 35%	2 = 5%	32 = 80%

Fonte: Elaboração própria.

Em relação às variáveis em estudo, relacionou-se a frequência das respostas dos informantes, que fizeram tal atribuição, a percentuais por linhas, refletindo a proximidade/afastamento de uso das formas diminutivas em questão, como podemos observar na tabela 05.

Tabela 05 – Distribuição da alternância –inho / -zinho conforme a frequência atribuída pelos informantes

FORMAS DIMINUTIVAS	FREQUÊNCIA DE ACEITABILIDADE					TOTAL
	1	2	3	4	5	
-INHO	211 / 400 = 52,8%	80 / 400 = 20,0%	43 / 400 = 10,8%	28 / 400 = 7,0%	38 / 400 = 9,5%	100%
-ZINHO	199 / 400 = 49,8%	73 / 400 = 18,3%	37 / 400 = 9,3%	59 / 400 = 14,8%	32 / 400 = 8,0%	100%
-INHOZINHO	4 / 400 = 1,0%	15 / 400 = 3,8%	18 / 400 = 4,5%	62 / 400 = 15,5%	301 / 400 = 75,3%	100%
TOTAL	414 / 1200 = 34,5%	168 / 1200 = 14,0%	98 / 1200 = 8,2%	149 / 1200 = 12,4%	371 / 1200 = 30,9%	100%

Fonte: Elaboração própria.

Na tabela 05, os dados mostram que 52,8% dos informantes afirmam utilizar amplamente o morfema –inho. No entanto, 20% reconhecem ser possível a utilização da variante –inho, embora não seja essa variante que utilizam com maior frequência. Embora presente um uso improvável na língua, 7% consideram utilizar ocasionalmente o diminutivo –inho e somente 9,5% declaram não fazer uso do morfema em questão. Contudo, para o formativo –zinho, observa-se que 49,8%

dos respondentes consideraram utilizá-lo plenamente, enquanto 18,3% admitem o uso da variante zinho, mas reconhecem não ser a variante a que recorrem num primeiro momento. Ainda que considerem improvável a possibilidade de uso na língua, 14,8% dos informantes afirmam utilizar esporadicamente o morfema –zinho e apenas 8,0% admitem não utilizar esse diminutivo. Finalmente, nota-se que somente 1% dos respondentes considera fazer uso da variante inhozinho. Entretanto, 3,8% admitem a possibilidade de uso da forma –inhozinho, embora não seja a variante à qual recorrem num primeiro momento. Contudo, 15,5% dos entrevistados consideram improvável o uso do diminutivo –inhozinho, mas admitem que de vez em quando o utilizam, enquanto 75,3% dos informantes declaram nunca utilizar esse formativo. Nesse teste, os respondentes que declaram não ter uma opinião formada ou que se abstiveram de emitir qualquer opinião correspondem a: 10,8% para o morfema –inho; 9,3% para o diminutivo –zinho; e 4,5% para a variante –inhozinho.

Posteriormente, calculou-se o *ranking* médio (doravante RM) da pontuação obtida em cada um dos cinco níveis. Neste caso, atribuímos os valores de 1 a 5 para cada resposta, a partir da qual foi calculada a média ponderada para cada item, baseando-se na frequência das respostas. Deste modo, os valores atribuídos foram: 1 – para o nível 5; 2 – para o nível 4; 3 – para o nível 3; e 4 – para o nível 2; e 5 – para o nível 1. Assim, tem-se a fórmula:

$$RM = \frac{\sum_{i=1}^n MG \times Vi}{\sum Vi} \tag{1}$$

Onde:

RM = *ranking* médio.

MG = média de pontos obtidos em cada grupo.

V_i = valor ou peso atribuído a cada resposta.

Em relação à avaliação do grau de concordância ou discordância de uso de cada uma das formas diminutivas, considerou-se o valor padrão 3 como o ponto neutro da escala de valores. Logo, a soma dos valores médios (RM Geral) acima de 3 é considerada como concordante, enquanto os menores que 3 são tidos como discordantes. Deste modo, realizou-se verificação quanto à concordância ou discordância das questões avaliadas através da obtenção do RM da pontuação atribuída às respostas, conforme se pode observar na tabela 06.

Tabela 06 – Ranking médio (RM)

SUFIXOS DIMINUTIVOS	Nível 1	Nível 2	Nível 3	Nível 4	Nível 5	RM Geral
	RM	RM	RM	RM	RM	
-INHO	2,64	0,80	0,32	0,14	0,10	4,00
-ZINHO	2,49	0,73	0,28	0,30	0,08	3,87
-INHOZINHO	0,05	0,15	0,14	0,31	0,75	1,40
Total	5,18	1,68	0,74	0,75	0,93	9,26

Fonte: Elaboração própria.

Na tabela acima, observa-se que, dos 40 respondentes, em média 4,00 declararam utilizar o formativo –inho, enquanto 3,87 dos entrevistados admitiram utilizar o diminutivo -zinho nos 10 itens do questionário aplicado. No entanto, o diminutivo –inhozinho foi utilizado em média por 1,40 dos respondentes. Note-se ainda que, das variáveis em análise, o RM geral encontra-se aproximadamente em torno de quatro, ou seja, evidencia que há um predomínio de nível satisfatório, indicando que as respostas tenderam para uma atitude positiva ou concordante quanto à possibilidade e uso dos referidos formativos, ficando somente a variável –inhozinho em um nível insatisfatório, pois atingiu uma média inferior ao ponto de corte ou ponto neutro estabelecido, o que revela uma atitude discordante em relação à possibilidade e uso do referido formativo.

A fim de medirmos a variabilidade dos dados em torno do RM, utilizou-se o cálculo do desvio padrão da média por meio da seguinte fórmula:

$$S = \sqrt{\frac{\sum_{i=1}^n (X_i - MG)^2}{n(n-1)}} \quad (2)$$

Onde:

S = desvio padrão.

X_i = variável usada para representar os valores individuais dos dados.

MG = média geral de pontos obtidos no RM.

Portanto, realizou-se verificação quanto à concordância ou discordância das questões avaliadas através da obtenção do RM da pontuação atribuída às respostas e analisou-se o grau de dispersão em relação a essa média, obtendo o desvio padrão do RM, conforme se pode observar na tabela 07.

Tabela 07 – Desvio padrão amostral (S)

SUFIXOS DIMINUTIVOS	Nível 1 (S)	Nível 2 (S)	Nível 3 (S)	Nível 4 (S)	Nível 5 (S)	S Geral
-INHO	1,64	0,41	0,20	0,28	0,37	2,90
-ZINHO	1,60	0,38	0,16	0,37	0,35	2,86
-INHOZINHO	0,09	0,10	0,08	0,42	2,31	3,00
Total	3,33	0,89	0,44	1,07	3,03	8,76

Fonte: Elaboração própria.

Na tabela 07, nota-se que os morfemas –inho e –zinho apresentam um desvio-padrão que se aproxima de suas respectivas médias, evidenciando que boa parte dos entrevistados consideraram de forma parecida o uso dos referidos diminutivos, revelando uma atitude positiva em relação aos diminutivos supracitados. Por outro lado, o formativo –inhozinho apresenta uma dispersão que ultrapassa sua respectiva média, indicando que a maioria dos respondentes revelaram-se contrários à aplicabilidade e uso da forma em questão, enquanto uma parcela mínima apresentou uma atitude concordante quanto à possibilidade e uso desse diminutivo, o que influenciou a média para baixo.

Validade, confiabilidade e capacidade discriminatória

Hernández, Fernández e Baptista (2003) ressalta que, ao utilizarmos um instrumento de medição, é necessário avaliarmos sua confiabilidade e validade para que a investigação aporte resultados reais e seja digna de consideração.

A qualidade do instrumento de medição é estimada pelos critérios de validade ou validade e por sua confiabilidade ou fidedignidade. O critério de validade está relacionado à capacidade de um instrumento medir com precisão o que de fato se propõe medir, ou seja, os dados da pesquisa devem ser precisos. Por outro lado, a confiabilidade refere-se à capacidade de reproduzir com consistência ou estabilidade os resultados obtidos quando o mesmo indivíduo ou objeto é avaliado, ou seja, qualquer pessoa que aplique um questionário, em condições idênticas, terá sempre os mesmos resultados. Segundo Guy e Zilles (2007), a confiabilidade equivale basicamente à possibilidade de reproduzir ou replicar um resultado. Os autores argumentam que, se repetirmos um estudo ou se outro pesquisador realizar um estudo equivalente com a utilização de critérios idênticos, se os resultados não forem iguais, não haverá confiabilidade, mas, se forem, podemos considerá-los confiáveis. Desse modo, Guy e Zilles (2007) consideram que “isto é uma questão de clareza dos procedimentos, do cuidado dos pesquisadores para evitar possíveis fontes de viés” (GUY; ZILLES, 2007, p.116) e mencionam ainda

que “alguns fenômenos do mundo são mais estáveis e outros menos; os mais estáveis dão sempre os mesmos resultados e são mais confiáveis” (GUY; ZILLES, 2007, p.116).

Para demonstrar a validade do instrumento de medição, realizou-se um teste piloto com 05 (cinco) pessoas que preencheram o questionário para o teste final do instrumento. Nenhuma modificação foi necessária, pois não se registraram problemas relacionados à clareza, ao entendimento, ao preenchimento e à adequação ao *layout* do questionário. Finda essa etapa, o questionário foi aplicado na amostra completa.

Para determinação da confiabilidade do questionário, usamos o coeficiente *alpha* de Cronbach (1951), que consiste em uma análise da covariância e/ou correlações entre os itens do questionário a fim de medir a homogeneidade dos componentes da escala. Neste trabalho, nos dedicamos a uma análise de consistência interna do questionário aplicado utilizando o pacote estatístico R *project*, Versão 2.12.2. Assim, os resultados obtidos foram:

Quadro 1 – Teste *Alpha* de Cronbach

Alpha reliability	Standardized alpha	N itens
0.8133	0.8183	10

Fonte: Elaboração própria.

De acordo com Hair et al. (1998), rejeitam-se os fatores com *alpha* de coeficiente menores que 0,6, uma vez que valores abaixo deste indicam uma intensidade de associação muito baixa. Valores entre 0,6 e 0,7 indicam uma intensidade de associação moderada; entre 0,7 e 0,8, boa; entre 0,8 e 0,9, muito boa e 0,9, excelente. Assim, analisando-se a estrutura do instrumento utilizado para obtenção dos dados através do modelo de *Alpha* de Cronbach (1951), percebe-se que o constructo e a escala utilizados apresentam uma confiabilidade interna muito boa e, portanto, são adequados para o propósito para o qual foram designados (QUADRO 01).

Poder discriminatório

Entende-se por poder discriminatório a capacidade que os resultados de um teste tem de quantificar as diferenças entre os respondentes.

O poder discriminante de uma escala pode ser expresso pelo coeficiente δ de Ferguson (1949), que varia de 0 (todas as pontuações são iguais) a 1 (cada

indivíduo tem uma pontuação diferente). Assim, quanto mais próximo de 1 o índice de Ferguson estiver, maior também será a capacidade discriminante do teste. Como regra geral, um teste é considerado discriminatório se apresentar $\delta \geq 0,90$. Nesse teste, o valor estimado do delta de Ferguson (1949) para a amostra foi de $\delta = 0,93$, sendo considerado um valor que apresenta uma boa capacidade discriminatória.

Modelagem estatística

A fim de identificarmos a modelagem estatística mais adequada ao fenômeno em estudo, aferimos os pressupostos de Gaussianidade¹ e de a homocedasticidade das variâncias. Deste modo, entende-se que os pressupostos básicos assumidos neste trabalho delineiam os contornos de uma pesquisa honesta e não viesada, o que valer dizer cientificamente aceita.

Em relação ao diagnóstico de normalidade, utilizou-se o teste Shapiro-Wilk (SHAPIRO; WILK, 1965). Esse teste, proposto em 1965, calcula uma estatística W que testa se uma amostra aleatória (x_1, x_2, \dots, x_n) é proveniente de uma distribuição normal. Assim, verificam-se as seguintes hipóteses:

H_0 : A amostra provém de uma distribuição normal $N(\mu, \sigma^2)$

H_1 : A amostra não provém de uma distribuição normal $N(\mu, \sigma^2)$

Neste caso, rejeitamos H_0 se o p-valor² é menor que 5% (p-valor < 0,05) e não rejeitamos H_0 caso contrário. Logo, observou-se que, da aplicação do teste de normalidade de Shapiro-Wilk (SHAPIRO; WILK, 1965) resultou um p-valor = 0,8948, portanto, a um nível de significância de 5%, não podemos rejeitar a hipótese da normalidade. Assim, torna-se plausível que a amostra provenha de uma população normal.

Muitas técnicas estatísticas requerem a suposição de igualdade de variâncias das variáveis de interesse e, portanto, considerou-se crucial verificar essa premissa. Foi utilizado o teste de Levene centrado na média para averiguar o pressuposto de homogeneidade da variância. Esse teste é usado para verificar se k médias têm a mesma variância, ou seja, a variabilidade dos resultados, para cada situação experimental, deve ser aproximadamente igual. Assim,

¹ O termo gaussianidade faz referência ao matemático e astrônomo Gauss, que utilizou essa distribuição para modelar a variação no resultado de medições do posicionamento de planetas.

² O p-valor refere-se à menor escolha que teríamos feito para determinado nível de significância (α), de forma a rejeitarmos H_0 .

as variáveis a serem testadas devem pertencer a grupos populacionais com igual variância. Deste modo, para avaliarmos a homogeneidade da variância, admitiu-se um nível de significância menor que 0,05 ($\alpha < 0,05$). Deste modo, foram estabelecidas as hipóteses:

- H_0 : As variâncias para os diferentes grupos são iguais (existe homocedasticidade)
 H_1 : As variâncias para os diferentes grupos são diferentes (não existe homocedasticidade)

Neste caso, rejeitamos H_0 se o p-valor for menor que 5% e não rejeitamos H_0 caso contrário. Para as variáveis avaliadas no teste de aceitabilidade, foram obtidos os seguintes valores:

Quadro 2 – Teste de homocedasticidade

Levene's Test for Homogeneity of Variance (center = mean)	
Group	Pr(>F)
Dialeto Regional	0.0891
FaixaEtária	0.9850
Gênero	0.2832
Diminutivo	0.4345

Fonte: Elaboração própria.

Ao analisarmos o resultado do teste de Levene, constatamos que, como $p > 0,05$, então não rejeitamos a hipótese nula, ou seja, não existem evidências de uma diferença significativa entre as variâncias dos grupos analisados.

Cabe ressaltar que a independência entre os grupos é automática, pois são grupos vindos de populações mutuamente excludentes (homens ou mulheres; cariocas ou fluminense ou outro; 18 a 20 anos ou 21 a 23 anos ou 24 a 26 anos). Observe-se que os dados coletados seguem uma distribuição normal e foi verificada a igualdade de variâncias das atitudes entre as variáveis independentes (Teste de Levene). Assim, dentre os vários testes paramétricos existentes, realizamos uma análise de variância *multifatorial* (ANOVA *Multi Way*) por considerarmos, no caso específico deste estudo, ser o teste mais adequado a explicar a alternância do diminutivo –inho/ -zinho no PB.

Hipóteses

Entende-se que, nos casos de variação, diante de uma só forma é de se esperar que haja convergência entre os níveis, tanto para -inho, considerado como morfema diminutivo, quanto para /z/ em -zinho, tido como uma consoante epentética que emerge para satisfazer exigências estruturais. Assim, a hipótese conservadora formulada (H_0) é de que haja convergência entre as médias ($\mu_1 = \mu_2 = \dots = \mu_i$). No entanto, se houver diferenças significativas entre os níveis, temos evidências para admitir a existência de formas distintas na língua portuguesa, visto que teríamos formas que se afastam progressivamente entre si. Logo, a hipótese inovadora (H_1) é de que existe pelo menos uma das médias diferentes ($\mu_i \neq \mu_j$ para algum $i \neq j$). Entretanto, a hipótese secundária a ser investigada é verificar se há ou não interação entre os fatores sociais e a atitude do falante. Assim, a hipótese nula (H_0) é de que haja interação entre os fatores sociais (dialeto regional, gênero e faixa etária) e a seleção da(s) forma(s) diminutiva(s), dada a proposta sociolinguística adotada neste estudo. Por outro lado, a hipótese alternativa (H_1) é de que a(s) forma(s) diminutiva(s) ocorrem por motivações estruturais, sem interferência de fatores sociais.

Discussão dos resultados

Os dados foram submetidos à análise de variância **multifatorial** (ANOVA *Multi Way*), com nível de significância de 0,05, para avaliar se o efeito do gênero, do dialeto regional e da faixa etária é determinante na seleção da(s) forma(s) em análise e se há ou não convergência entre as formas. A seguir, na tabela 08, apresentamos os resultados obtidos na aplicação do modelo Π^3 de ANOVA **multifatorial**.

³ Este modelo considera que todos os fatores têm efeitos aleatórios.

Tabela 08 – Resultados da ANOVA multifatorial para a pontuação na escala de atitudes em relação ao uso da(s) forma(s) diminutiva(s) – inho e/ou –zinho, em função do dialeto regional, do sexo e da faixa etária

Efeito	Soma de quadrados (SQ)	Graus de liberdade (GL)	Média de quadrados (MQ)	F	p
DIALETO	0,8	2	0,4	0,0117	0,98833
FAIXA ETÁRIA	29,0	2	14,5	0,4383	0,64653
GÊNERO	0,0	1	0,0	0,0003	0,98557
DIMINUTIVO	17167,9	2	8583,9	259,5613	< 2e-16
DIALETO*FAIXA ETÁRIA	0,0	1	0,0	0,0007	0,97959
DIALETO*GÊNERO	218,0	2	109,0	3,2960	0,04171
FAIXA ETÁRIA*GÊNERO	6,5	2	3,3	0,0988	0,90606
DIALETO*DIMINUTIVO	204,1	4	51,0	1,5429	0,19691
FAIXA ETÁRIA*DIMINUTIVO	205,1	4	51,3	1,5503	0,19484
GÊNERO*DIMINUTIVO	41,3	2	20,6	0,6243	0,53802
DIALETO*FAIXA ETÁRIA*DIMINUTIVO	34,8	2	17,4	0,5268	0,59238
DIALETO*GÊNERO*DIMINUTIVO	83,2	4	20,8	0,6288	0,64324
FAIXA ETÁRIA*GÊNERO*DIMINUTIVO	25,3	4	6,3	0,1911	0,94244
Resíduos	2877,2	87	33,1		

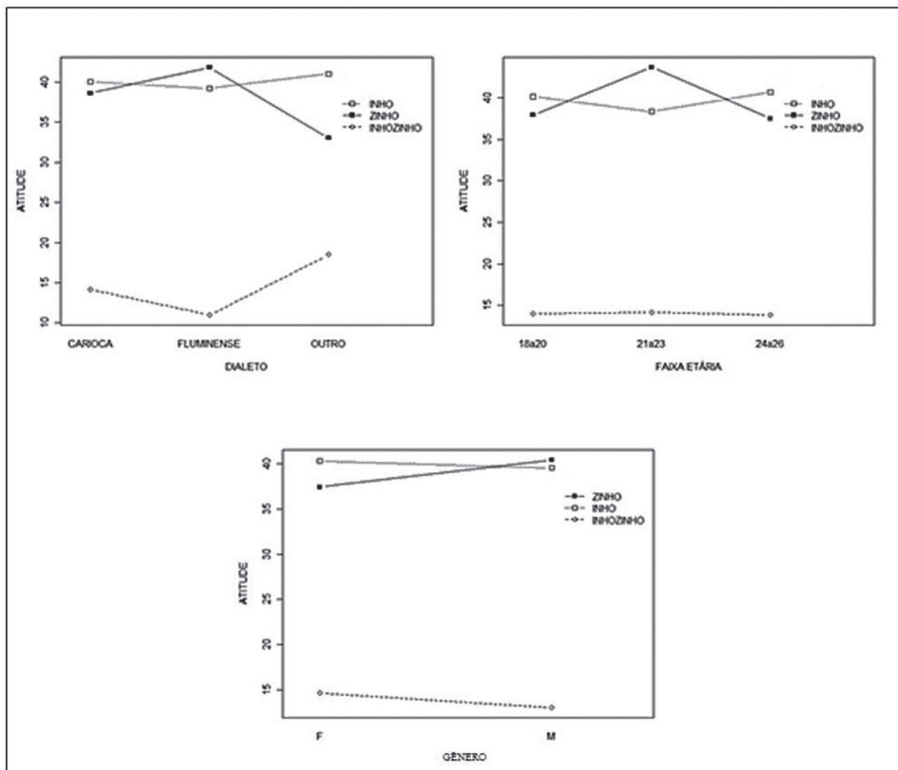
Fonte: Elaboração própria.

Tomados individualmente os efeitos principais, observa-se que o p valor para o fator diminutivo, aproximadamente zero, é muito menor que o p valor para a faixa etária (0,64653), o gênero (0,98557) e o dialeto regional (0,98833). Logo, para $\alpha = 0,05$, rejeitamos a hipótese conservadora (H_0), ou seja, o fator “diminutivo” tem efeito significativo sobre a variabilidade da variável de resposta “atitude”, apresentando potenciais diferenças entre os níveis, enquanto que não há evidência estatística para afirmar que, individualmente, os fatores “faixa etária”, “gênero” e “dialeto” exerçam efeito distinto sobre a percepção do falante, visto que, para estes fatores, $p > 0,05$. Em outras palavras, os fatores analisados apontam para a existência de diferenças significativas entre os formativos (-inho, -zinho e -inhozinho), o que nos leva a admitir a existência de mais de uma forma diminutiva na língua.

Quando observamos os cruzamentos entre os fatores analisados, vemos que todas as combinações entre os fatores (A*B, A*C, A*D, A*B*C, A*B*D, A*C*D) indicam que há possíveis interações, conforme o exposto nos gráficos de interação⁴ (GRÁFICO 01).

⁴ Neste estudo, adotou-se o método gráfico de interação para ANOVA sugerido por Levine, Berenson e Stephan (2000). A interação entre os fatores corresponde à diferença de comportamento de um fator nos diferentes níveis do outro fator (eixo) com respeito à característica de interesse.

Gráfico 01 – Interação entre a atitude do falante e os fatores “dialeto”, “faixa etária” e “gênero” na seleção das formas diminutivas



Fonte: Elaboração própria.

Como se pode observar no gráfico 01, a interação entre os formativos -inho, -zinho e -inhozinho é verificada para todos os fatores analisados, visto que as retas são congruentes ou estão inclinadas, ou seja, não são sensivelmente paralelas. Logo, para $p > 0,05$, não rejeitamos a hipótese nula (H_0), ou seja, concluímos que há evidências de um efeito de interação entre os fatores dialeto regional, gênero e faixa etária na seleção das formas diminutivas -inho e -zinho, o que nos leva a afirmar que as propriedades estruturais (número de sílabas, acentuação, segmento final da forma primitiva e flexão nominal) por si só não são determinantes na seleção das formas diminutivas. Em outras palavras, os fatores sociais (dialeto regional, gênero e faixa etária) revelam-se expressivos, condicionando a escolha do falante por um ou outro formativo.

Note-se que a probabilidade de ocorrer uma das variantes cresce na medida em que se conjugam os vários fatores sociais. Vejamos como as interações entre

as médias para os grupos analisados em cada fator são determinantes na escolha dos sufixos diminutivos –inho e –zinho (QUADRO 03).

Quadro 03 – Interação entre as médias para os fatores sociais em relação à escolha dos formativos –inho e –zinho

DIMINUTIVO -INHO				DIMINUTIVO -ZINHO			
FAIXA ETÁRIA / GÊNERO	DIALETO REGIONAL			FAIXA ETÁRIA / GÊNERO	DIALETO REGIONAL		
	CARIOCA	FLUMINENSE	OUTRO		CARIOCA	FLUMINENSE	OUTRO
18 a 20 anos				18 a 20 anos			
MASCULINO	39,43	39	NA	MASCULINO	37,86	42	NA
FEMININO	40,47	41	NA	FEMININO	37,12	40,50	NA
21 a 23 anos				21 a 23 anos			
MASCULINO	38	NA	NA	MASCULINO	44	NA	NA
FEMININO	39	NA	NA	FEMININO	43	NA	NA
24 a 26 anos				24 a 26 anos			
MASCULINO	44	36	42	MASCULINO	38,5	44	41
FEMININO	38	NA	40	FEMININO	38	NA	25

DIMINUTIVO -INHOZINHO			
FAIXA ETÁRIA / GÊNERO	DIALETO REGIONAL		
	CARIOCA	FLUMINENSE	OUTRO
18 a 20 anos			
MASCULINO	12,43	11	NA
FEMININO	15,35	10,5	NA
21 a 23 anos			
MASCULINO	13,5	NA	NA
FEMININO	15,5	NA	NA
24 a 26 anos			
MASCULINO	10	12	27
FEMININO	14	NA	10

Fonte: Elaboração própria.

Em média, no quadro 03, vemos que o sufixo –inho revela maior incidência de uso entre as mulheres do dialeto fluminense que compreendem a faixa etária de 18 a 20 anos. No entanto, entre os homens, essa incidência ocorre no dialeto carioca para os falantes que se encontram na faixa de 24 a 26 anos. Por outro lado, o diminutivo –zinho apresenta maior aplicação entre as mulheres cariocas de 21 a 23 anos. Entretanto, os homens cariocas e fluminenses também empregam majoritariamente essa variante. Contudo, são os cariocas de 21 a 23 anos e os fluminenses com idades entre 24 e 26 anos que a utilizam com maior incidência. Por último, a variante –inhozinho ocorre majoritariamente entre as mulheres cariocas de 21 a 23 anos, competindo com o sufixo –zinho para as mesmas condições de uso, ou seja, entre as cariocas com idades que variam de 21 a 23 anos. No entanto, os homens não cariocas e não fluminenses, aqueles pertencentes a outros dialetos regionais, revelam maior incidência de uso na faixa etária de 24 a 26 anos.

Teste Post-hoc

É importante salientar que o teste ANOVA apresenta informações se há uma diferença global entre os grupos, mas não nos permite identificar quais médias diferem entre si. Para isso, temos que utilizar uma análise de comparações múltiplas de médias (par a par) a fim de examinarmos quais dos pares apresentam diferenças significativas entre si. Neste caso, investigamos quais dos grupos diferem dos demais para o fator diminutivo, visto que a hipótese conservadora (H_0) de igualdade das médias foi rejeitada. Assim, utilizamos o teste *post-hoc* (*a posteriori*) de Bonferroni para confirmar onde as diferenças ocorreram entre os níveis. Mais especificamente, visamos avaliar a diferença entre a atitude do falante e a escolha da forma diminutiva. O nível de significância adotado foi de 0,05

O teste *post-hoc* de Bonferroni apresentou um índice de significância de aproximadamente zero ($< 2e-16$) para os pares inhozinho-inho e zinho-inhozinho, revelando que estes pares são significativamente diferentes ($p < 0,05$). Deste modo, pode-se evidenciar ainda que o produto -inhozinho é resultado da recursividade das formas -inho e -zinho, que se estabelece à direita do vocábulo, sendo o sufixo -zinho sempre posterior a -inho (vidinhazinha, passarinhozinho, bondinhozinho).

Considerações finais

Este estudo teve como objetivo principal avaliar, por meio de juízos de aceitabilidade, a alternância do diminutivo -inho/-zinho no PB. Deste modo, foram apresentadas duas hipóteses a fim de responder à questão central desta investigação: trata-se de uma única forma ou temos dois diminutivos distintos?

Partindo das diferentes interpretações conferidas ao fenômeno em estudo, esperar-se-ia que, diante de uma só forma diminutiva, houvesse convergência entre os níveis nos casos de variação. Logo, na hipótese conservadora (H_0) formulada, o formativo -inho seria considerado como o único morfema diminutivo e emergiria o segmento consonantal /z/ a fim de satisfazer as condições de boa formação estrutural (a evitação do hiato, a preservação da estrutura silábica da base e a preservação do acento marcado). Por outro lado, na hipótese inovadora (H_1), o afastamento progressivo entre os níveis levar-nos-ia a admitir a existência de duas formas distintas na língua portuguesa, -inho e -zinho. Considerou-se ainda uma hipótese secundária sobre a motivação da distribuição do diminutivo -inho/-zinho no PB. Assim, na hipótese nula (H_0), supomos que há interação entre os fatores sociais (dialeto regional, gênero e faixa etária) e a seleção do(s) sufixo(s) diminutivo(s) -inho/-zinho enquanto que, na hipótese alternativa (H_1), admitimos que a(s) forma(s) diminutiva(s) em questão ocorreria(m) apenas por motivações estruturais, sem interferência de fatores sociais.

Tomando como ponto de partida as hipóteses levantadas, os dados então foram submetidos à análise de variância **multifatorial** (ANOVA *Multi Way*), com nível de significância de 0,05. Deste modo, os resultados obtidos levam-nos a considerar a existência de duas formas distintas na língua portuguesa, visto que, para $p > 0,05$, a hipótese conservadora (H_0) foi rejeitada. Nesse sentido, tal resultado corrobora, de modo semelhante, como já foi referido na literatura por diferentes autores (LEITE, 1974; LEE, 1995; VILLALVA, 2000; BASÍLIO, 2004), a existência de duas formas diminutivas distintas na língua portuguesa, -inho(a) e -zinho(a). No entanto, embora esses autores refiram-se a algum grau de variação, suas análises restringem-se a aspectos estruturais, morfológicos e fonológicos, na distribuição e no comportamento desses formativos.

A fim de explicar a relação dos fatores sociais e a distribuição dos formativos -inho e -zinho, a segunda hipótese formulada neste estudo revelou que, para $p > 0,05$, não há evidência para rejeitarmos a hipótese nula (H_0), ou seja, concluímos que há efeito de interação entre os fatores dialeto regional, gênero e faixa etária na seleção das formas diminutivas -inho e -zinho. Assim, se -inho e -zinho são sufixos distintos e ocorre interação entre os fatores sociais na distribuição das duas formas diminutivas, a questão a ser respondida é: como os fatores sociais condicionam a escolha do falante por uma ou outra forma?

Em resposta à questão formulada, constatou-se que o formativo -inho apresentou maior incidência de uso entre as mulheres do dialeto fluminense que compreendem a faixa etária de 18 a 20 anos. Contudo, entre os homens, essa incidência foi verificada no dialeto carioca para os falantes que se encontram na faixa de 24 a 26 anos. Por outro lado, o diminutivo -zinho revelou maior aplicação entre as mulheres cariocas de 21 a 23 anos. No entanto, são os cariocas de 21 a 23 anos e os fluminenses com idades entre 24 e 26 anos que a utilizam com maior incidência. Por último, a variante -inhozinho ocorreu majoritariamente entre as mulheres cariocas de 21 a 23 anos, competindo com -zinho para as mesmas condições de uso, ou seja, entre as cariocas com idades que variam de 21 a 23 anos. Contudo, os homens não cariocas e não fluminenses, aqueles pertencentes a outros dialetos regionais, revelaram maior incidência de uso na faixa etária de 24 a 26 anos.

Neste caso, os resultados revelam tendências de uso que explicam a escolha do falante dentre dois sufixos diferentes que podem expressar o diminutivo. Em outras palavras, é a alternância de escolha do sufixo diminutivo a ser usado conforme a situação de uso.

FREITAS, M. A. de; BARBOSA, M. F. M. The alternation of the diminutive forms -inho/-zinho in Brazilian Portuguese: a variational approach. **Alfa**, São Paulo, v.57, n.2, p.577-605, 2013.

- **ABSTRACT:** *In this work, we will investigate the alternation of the diminutive forms -inho/-zinho, delineating the factor that determines the choice of the speaker towards one or another formative in Brazilian Portuguese. More specifically, based on the assumptions of the Variationist Sociolinguistics, we will analyze the productive aspect of diminutive formation in order to verify if we have one or more than one diminutive suffix, giving rise to two distinct processes, which are, respectively, X-inho e X-zinho formations. In this paper, we used multi-factor analysis of variance (ANOVA Multi Way) for repeated measures with 95% confidence limits, in order to evaluate if the choice of either suffix is determined by the effect of social factors (sex, age and regional dialect) and if there is convergence between the diminutive forms -inho and/or -zinho, measured by judgment test. The results reveal that there is evidence to admit the existence of two distinct diminutive forms in Brazilian Portuguese, -inho and -zinho. Besides, the results indicate that there is an interaction among regional dialect, sex and age in the both choice of diminutive suffixes.*
- **KEYWORDS:** *Diminutive suffixes in Brazilian Portuguese. Word formation. Alternation. Linguistic variation.*

REFERÊNCIAS

BARBOSA, J. S. *Gramática Filosófica da Língua Portuguesa*. Lisboa: Typographia da Academia Geral de Ciências de Lisboa, 1822.

BARROS, J. de. *Grammatica da Língua Portuguesa*. Lisboa: Imprensa Nacional – Casa da moeda, 1540.

BASILIO, M.M.P. *Formação e classes de palavras no português do Brasil*. São Paulo: Contexto, 2004.

BISOL, L. O diminutivo e suas demandas. *D.E.L.T.A.*, São Paulo, v.26, n.1, p.59-83, 2010.

CÂMARA JÚNIOR, M. *História e estrutura da Língua Portuguesa*. Rio de Janeiro: Padrão, 1975.

CRONBACH, J. L. Coefficient alpha and the internal structure of tests. *Psychometrika*, Springer New York, v.16, n.3, p.297-334, set. 1951.

FERGUSON, G. A. On the theory of test discrimination. *Psychometrika*, Springer New York, v.14, n.1, p.61-68, 1949.

FREIRE, J. F. *Reflexões sobre a Língua Portuguesa*. Lisboa: Typographia da Academia Geral de Ciências de Lisboa, 1842.

GONÇALVES, C. A. V. *Flexão e derivação em português*. Rio de Janeiro: Setor de Publicações da Faculdade de Letras da UFRJ, 2005.

GUY, G. R.; ZILLES, A. *Sociolinguística quantitativa: instrumental de análise*. São Paulo: Parábola Editorial, 2007.

HAIR, J. F. et al. *Multivariate data analysis*. 5.ed. New Jersey: Prentice Hall, 1998.

HERNÁNDEZ, R; FERNÁNDEZ, C; BAPTISTA, P. *Metodología de la investigación*. 3.ed. México: Mc. Graw-Hill, 2003.

INKELAS, S. Deriving cyclicity. In: HARGUS, S.; KAISSE, E. M. (Ed.). *The studies in Lexical Phonology*. San Diego: Academic Press, 1993.

_____. *Prosodic constituency in the lexicon*. Stanford: Stanford University, 1989.

LABOV, W. On the use of present to explain the past. In: INTERNATIONAL CONGRESS OF LINGUISTS, 11, 1975, Bologna. *Proceedings...* Bologna: Heilmann, 1975. p.825-851.

_____. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press, 1972.

LEE, S. *Morfologia e Fonologia Lexical do Português*. 1995. 201f. Tese (Doutorado em linguística) – Faculdade de Letras, Universidade Estadual de Campinas, 1995.

LEITE, Y. *Portuguese Stress and Related Rules*. 1974. 304f. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade de Texas, Austin, 1974.

LEVINE, D. M.; BERENSON, M. L.; STEPHAN, D. *Estatística: teoria e aplicações*. Rio de Janeiro: LTC Editora, 2000.

LIKERT, R. A. Technique for the Measurement of Attitudes. *Psychology*, New York, n.140, p.1-55, 1932.

OLIVEIRA, F de. *Gramática da Linguagem Portuguesa*. Vila Real: Centro de Estudos em Letras da Universidade de Trás-os-Montes e Alto Douro, 1536.

SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples). *Biometrika*, Londres, v.52, p.591-611, 1965.

SKORGE, S. Os sufixos diminutivos em português. *Boletim de Filologia*, Lisboa, tomos XVI e XVII, fascículos 3 e 4, p.40-90, 1957.

TAVARES, J.N.C. O uso dos sufixos –inho e –zinho no discurso dos pescadores artesanais. In: III CONGRESSO NACIONAL DE LINGUÍSTICA E FILOGOIA, 3, 1999, Rio de Janeiro. *Anais...* Rio de Janeiro: CiFEFiL: IL-UERJ: ABF: CEFIL, 1999.

VILLALVA, A. *Estruturas morfológicas: unidades e hierarquias nas palavras do português*. Braga: Fundação Calouste Gulbenkian, 2000.

WEINREICH, U.; LABOV, W.; HERZOG, M. I. Empirical foundations for a theory of language change. In: LEHMANN, W. P. (Ed.). *Directions for historical linguistics: a symposium*. Austin: University of Texas Press, 1968. p.95-195.

Recebido em fevereiro de 2012

Aprovado em novembro de 2012