

MÉTODOS OFF-LINE EM PSICOLINGUÍSTICA: JULGAMENTO DE ACEITABILIDADE

Cândido Samuel Fonseca de Oliveira

Thaís Maíra Machado de Sá¹

RESUMO

Os estudos psicolinguísticos procuram entender quais são os processos mentais envolvidos na faculdade da linguagem. Neste artigo, procuramos dar enfoque ao método *off-line* Julgamento de Aceitabilidade. Os métodos *off-line* são aqueles que investigam o processamento a partir de dados obtidos pós-processamento. Neles, os sujeitos são expostos às questões da pesquisa e após a exposição devem executar alguma tarefa. Ou seja, não se observam os processos mentais do sujeito ao longo do processamento do que se quer investigar, mas quando o processamento já foi concluído. O julgamento de aceitabilidade é o julgamento de indivíduos em relação à aceitabilidade de sentenças. Assim, a avaliação dos sujeitos é um dado pós-processual, mas que não deixa de refletir um processo mental. A avaliação do sujeito aos estímulos pode ser realizada por meio da Escala Likert, que é uma escala psicométrica constituída de vários pontos, sendo as escalas de cinco e sete pontos as mais utilizadas. Cada ponto se refere ao grau de aceitabilidade do indivíduo em relação ao estímulo apresentado. Outra opção é a Estimativa de Magnitude, em que os participantes devem atribuir valores que expressem a relação, proporção ou diferença entre os estímulos. O método Julgamento de Aceitabilidade apresenta confiabilidade e facilidade de aplicação, sendo um método de grande utilidade para estudos psicolinguísticos.

Palavras-chave: psicolinguística; escala Likert; estimativa de magnitude.

¹ Mestrandos da Pós-graduação em Estudos linguísticos da FALE/UFMG, orientados, respectivamente, pelos profs. Drs. Ricardo Augusto de Souza e Maria Luiza Cunha Lima. Trabalho desenvolvido no Laboratório Virtual de Psicolinguística da FALE/UFMG. E-mails: coliveira.ufmg@gmail.com; thaismaira@gmail.com

INTRODUÇÃO

Quem veio primeiro: a língua ou o pensamento? Como língua e pensamento são associados? Como as pessoas compreendem o que escutam ou leem? Quais são os processos mentais utilizados pelas pessoas quando elas falam ou escrevem algo? Como as crianças aprendem a entender o que ouvem e a falar o que sentem? A psicolinguística é uma ciência que se preocupa com questões como essas. Ela investiga como adquirimos, compreendemos e produzimos língua. Dessa forma, os estudos psicolinguísticos procuram entender quais são os processos mentais envolvidos na faculdade da linguagem. (cf. Glucksberg, 1975; Deese, 1976; Gernsbacher, 1994).

Vários métodos são utilizados em investigações sobre a língua. O mais tradicional deles é a introspecção, realizado pelo pesquisador que reflete sobre a língua e cria hipóteses a partir de sua própria reflexão (cf. Starr e Rayner, 2005). Não há, assim, uma metodologia que envolva sujeitos² ou dados do processamento comum obtidos de um experimento propriamente dito. Apesar de muito utilizada em estudos linguísticos, devemos questionar a confiabilidade da introspecção. Muitas vezes nos cercamos por tanto tempo de construções linguísticas que ficamos acostumados a elas e perdemos nossa capacidade de avaliação, o que acaba enviesando nossa reflexão. Assim, os psicolinguistas procuram associar a introspecção a métodos experimentais, pois permitem que o pesquisador obtenha dados que verifiquem sua hipótese sobre o funcionamento da língua. Os métodos de pesquisa sobre o processamento da língua podem ser divididos em: *off-line*, *on-line* e fisiológicos. (cf. Field, 2004; Gernsbacher, 1994; Leitão, 2008).

² Sujeitos serão aqui definidos como os voluntários livres e esclarecidos que contribuem com a pesquisa linguística, realizando a tarefa criada pelo experimentador. É a partir das respostas dos sujeitos que o experimentador obtém os dados de sua pesquisa.

Os métodos *off-line* são aqueles que investigam o processamento a partir de dados obtidos pós-processamento. Neles, os sujeitos são expostos às questões da pesquisa e após a exposição devem executar alguma tarefa. Ou seja, não se observam os processos mentais do sujeito ao longo do processamento do que se quer investigar, mas quando o processamento já foi concluído. Alguns exemplos de experimentos *off-line* são: o Julgamento de Aceitabilidade (descrito neste artigo), Testes de Associação de Palavras, Testes de Completação, entre outros.

Os métodos *on-line* e fisiológicos se contrapõem aos *off-line*, já que os dados são obtidos durante o processamento. Nos métodos *on-line* e fisiológicos os sujeitos são expostos a estímulos que apresentam questões da pesquisa e durante a exposição são avaliados. Nos *on-line*, uma tarefa é realizada ao longo da exposição, sendo que a tarefa depende do comportamento do sujeito. Tarefas como Nomeação, Priming e Leitura Auto-Cadenciada são exemplos de métodos *on-line*.

Já os fisiológicos não dependem do comportamento do sujeito, pois os dados são obtidos a partir de reações fisiológicas que o sujeito tem no momento em que é exposto ao que se quer investigar. Um exemplo de teste fisiológico é o ERP (Potenciais Relacionados a Eventos), que através de uma touca de eletrodos que é colocada na cabeça do sujeito, obtêm-se potenciais elétricos no momento em que o sujeito visualiza ou escuta frases ou palavras com o problema teórico que é investigado³.

Este artigo objetiva descrever uma das possibilidades da metodologia *off-line*: o Julgamento de Aceitabilidade, além de refletir sobre seus benefícios e seus problemas metodológicos.

³ Para saber mais sobre métodos *on-line* e métodos fisiológicos, ler Starr e Rayner (2005), Mitchell (2004).

1 A METODOLOGIA OFF-LINE

Imagine que você tem como hipótese que a construção “ir em” é aceita por estudantes de Letras falantes nativos do português brasileiro. Pela sua intuição você poderia pensar que a frequência da construção agramatical (segundo a gramática tradicional) “ir em” é tão grande que não haveria problemas em ser aceita. Contudo, a realidade poderia ser oposta à sua hipótese inicial e os estudantes de Letras, por terem maior conhecimento da gramática tradicional, não aceitariam a construção agramatical.

Como a intuição não é suficiente, você gostaria de comparar tal construção, que é agramatical e muito frequente, com a construção “ir a”, que é gramatical, mas pouco utilizada. Imagine que você tenha apresentado sentenças com ambas as regências do verbo “ir” para 30 alunos e tenha pedido que eles avaliassem de 0 a 10 as sentenças. Assim, os sujeitos leriam os estímulos e após a leitura eles atribuiriam as notas.

Durante a leitura dos estímulos ocorreria o processamento da sentença e só após o processamento é que você teria uma resposta do que você demonstrou. Dessa forma, você obteria dados que refletem algo sobre o processamento dos sujeitos, mas não no momento em que o processamento acontece. Ser pós-processual é o que define os experimentos *off-line*. Os dados pós-processuais podem ser dados pela *compreensão* dos sujeitos, como no exemplo acima, e pela *produção* dos sujeitos. Aqui focaremos em procedimentos que envolvem a compreensão ⁴. Antes de entrarmos nos experimentos propriamente ditos, discutiremos um pouco sobre tarefas de compreensão.

⁴ Por falta de espaço, optamos por descrever mais detalhadamente o Julgamento de Aceitabilidade, que é um experimento que envolve a compreensão do falante. Escolhemos tal método porque ele vem mostrando maior aplicabilidade. Para saber mais sobre produção, ver Gernsbacher (1994).

2 TAREFAS DE COMPREENSÃO

As tarefas de compreensão são aquelas que investigam o processamento em relação a o que o falante compreende do estímulo. Continuando no exemplo da aceitabilidade de “ir em”, ao perguntarmos aos sujeitos o que eles acharam das sentenças, se elas são aceitáveis ou não, queremos perceber como o sujeito *compreendeu* tal sentença.

Outro caminho que poderíamos recorrer seria pedir aos sujeitos que produzissem sentenças com o verbo ir. Se realizássemos tal tarefa, ao invés de pedir um julgamento sobre uma sentença dada, realizaríamos uma tarefa de produção, pois investigaríamos como o sujeito *produziu* a sentença e não como ele a entendeu.

Assim, tarefas de compreensão e produção divergem no que se espera avaliar sobre o processamento do sujeito, se é o *input*, a recepção e decodificação da informação, ou o *output*, a produção da informação.

3 JULGAMENTO DE ACEITABILIDADE

Como mencionado na introdução, muitas teorias foram e ainda são baseadas apenas no julgamento provindo da introspecção do próprio linguista. Encontram-se comumente na literatura propostas que baseiam sua argumentação em frases que são julgadas pelo próprio autor. Tais julgamentos são comumente representados a partir de símbolos tais como *, **, ?, e ?? ⁵.

⁵ O símbolo * comumente se refere a frases agramaticais e o símbolo ** se refere a frases consideradas ainda mais agramaticais do que as que. Já o símbolo ? se refere a frases cuja gramaticalidade ou aceitabilidade são duvidosas e o símbolo ?? é utilizado para aumentar o grau de intensidade de tal característica.

No desenvolvimento da dissertação de Oliveira⁶, por exemplo, ao se investigar a aquisição das construções resultativas, os pesquisadores tiveram que basear todo o seu estudo em obras cuja gramaticalidade das construções era relatada a partir da intuição de seus autores (WESCHLER, 2001; BROCCIAS, 2003; JACKENDOFF; GOLDBERG, 2004). Conseqüentemente, Oliveira encontrou o mesmo problema apontado por Bard (1996) em que itens pertencentes a um determinado grupo (*,**, ? ou ??) não parecem ter o mesmo nível de aceitabilidade e, além disso, o intervalo entre esses grupos não apresenta regularidade. Tal problema se deve a questões metodológicas.

A aceitabilidade de uma construção pode ter grande variabilidade, logo uma divisão em cerca de quatro grupos, por exemplo, pode não ser suficiente para uma classificação adequada. O corolário de tal metodologia, então, é uma grande variedade de itens com diferentes níveis de aceitabilidade dentro de um mesmo grupo.

Com o objetivo de se testar hipóteses na área da linguística aplicada e da linguística teórica e de tornar robustas as evidências relacionadas às hipóteses a partir da aceitabilidade de sentenças, o julgamento de vários indivíduos em relação à aceitabilidade de sentenças há anos é utilizado (BARD, 1996). Acredita-se que o julgamento de falantes em relação à aceitabilidade de algumas sentenças pode ser utilizado para se entender como certas construções (morfemas, palavras e, entre outros, sentenças) são representadas (ZARA, 2009; SOUZA, 2011) e o quão presente elas estão na gramática de um ou mais grupos de indivíduos.

É importante ressaltar que o julgamento de aceitabilidade não se refere à gramaticalidade de um estímulo. Apesar de haver trabalhos que utilizam os termos “julgamento de gramaticalidade” e “julgamento de aceitabilidade” com o mesmo significado (MYERS, 2008), essas duas expressões se referem a processos distintos.

⁶ Trabalho de dissertação em desenvolvimento no POSLIN/FALE/UFMG, linha 1D, sob orientação do prof. Dr. Ricardo Augusto de Souza, com previsão de término para março de 2013.

Enquanto a gramaticalidade de uma construção é uma característica inerente à mesma (SORACE, 2010), com embasamento teórico (MYERS, 2008) e independente do indivíduo, a aceitabilidade é a percepção que o indivíduo tem em relação a um item independentemente de qualquer construto teórico. Assim, os dados de um julgamento de aceitabilidade são relatos referentes às sensações dos participantes frente às construções em questão. Além disso, uma construção é julgada como aceitável ou não aceitável não apenas de acordo com sua gramaticalidade, mas também de acordo com seu custo de processamento, o seu significado e, entre outros, o contexto onde está inserida. Portanto, o termo julgamento de aceitabilidade parece ser o mais apropriado para se referir a estudos que abordam a gramática de um indivíduo, enquanto o termo julgamento de gramaticalidade poderia ser utilizado para se estudar o conhecimento das pessoas em relação à gramática (nesse caso, à gramática normativa) de uma língua.

Julgamentos de aceitabilidade são essenciais para as pesquisas linguísticas, já que a falta de paralelismo entre uso e conhecimento linguístico faz com que muitos fenômenos não possam ser percebidos de maneira fácil e natural (SORACE, 2010). Assim, experimentos que de alguma forma colocam os indivíduos em contato com os estímulos linguísticos em questão parecem ser a forma mais prática de verificar se a construção estudada compõe de alguma forma a gramática de um grupo específico.

Além disso, julgamentos de aceitabilidade permitem que linguistas estudem a reação de falantes frente a construções que não existem no uso comum da língua – informação negativa (SHUCTZE, 1996). Tal informação pode ser utilizada, por exemplo, para testar a influência de L2 na L1 de um falante bilíngue (OLIVEIRA; SOUZA, 2012). Para isso, o linguista elabora um experimento no qual o falante lidará com sentenças na sua língua nativa que são apenas licenciadas em sua L2. Oliveira e Souza (2012) e Oliveira, no desenvolvimento da sua dissertação, por exemplo, investigaram como falantes do Português Brasileiro (PB) representam construções específicas do inglês em PB: alternância de ação induzida e resultativa, respectivamente. Pouquíssimas

seriam as chances de serem encontrados usos naturais suficientes de sentenças como “ele marchou os soldados ao acampamento” e “ele martelou o metal torto reto”. Dessa forma, o julgamento de aceitabilidade fez com que o fenômeno fosse percebido mais facilmente.

Dois métodos comuns em psicolinguísticas para a obtenção de dados relativos a julgamentos de aceitabilidade são a Escala Likert e a Estimativa de Magnitude. A partir da experiência na utilização de tais métodos por Oliveira e na pesquisa bibliográfica desenvolvida para tal utilização, descrevemo-los nas próximas seções.

3.1 Escala Likert

Um dos métodos mais comuns na coleta de dados de julgamentos de gramaticalidade é a escala Likert (Fig. 1). É uma escala psicométrica que pode ser utilizada para se medir o nível de aceitabilidade de um indivíduo em relação a uma construção. Ela pode ser constituída de vários pontos sendo as escalas de cinco e sete pontos as mais utilizadas. Cada ponto se refere ao grau de aceitabilidade do indivíduo em relação ao estímulo apresentado.

Assim, o valor mínimo da escala corresponderia à rejeição total de um item e o valor máximo à aceitação total. O valor intermediário demonstraria uma posição neutra em relação ao item. Já os outros valores situados entre as extremidades e ponto médio seriam julgamentos de rejeição ou aceitação parcial do item.



Figura 1. Escala Likert de 7 pontos.

Métodos de julgamento de aceitabilidade como a escala Likert apresentam algumas vantagens em relação a outros métodos experimentais, tais como os métodos de processamento linguístico *on-line*, como Leitura Auto-Cadenciada, e métodos fisiológicos como ERP. O julgamento de aceitabilidade pode ser aplicado presencialmente ou via internet, por exemplo, em poucas sessões, o que facilita a obtenção de um número alto de sujeitos. Além disso, o teste é aplicado de forma rápida, o que faz com que julgamentos de aceitabilidade possam ser facilmente reaplicados e os dados reanalisados. Em testes de processamento *on-line*, diferentemente, a obtenção de um alto número de sujeitos não ocorre com a mesma facilidade, já que costumam demandar uma grande estrutura no momento da aplicação.

Na pesquisa de Oliveira, utiliza-se o paradigma da escala Likert para iluminar questões relacionadas à aquisição de estrutura argumental de uma construção rara na L1 e frequente na L2. Mais especificamente, investigou-se como indivíduos falantes nativos do português brasileiro com alto nível de proficiência em inglês representam construções resultativas gramaticais e agramaticais do inglês. Esta é uma construção pouco licenciada em português (LOBATO, 2004) e com rica variabilidade e aceitabilidade em inglês (JACKENDOFF; GOLDBERG, 2004). No trabalho em questão utilizaram-se apenas construções resultativas com o padrão sintático NP-V-NP-AP. A telicidade⁷ e, conseqüentemente, a gramaticalidade das sentenças que instanciam a construção resultativa com a estrutura acima são dependentes do adjetivo que a compõem (WECHSLER, 2001). Basicamente, o predicado resultativo precisa ser um adjetivo de escala fechada com ponto final máximo quando o verbo da sentença for durativo; ou pode ser um adjetivo não gradual caso o verbo da sentença seja pontual, como pode ser visto nos exemplos (1) e (2) abaixo:

⁷ Telicidade se refere à propriedade de um sintagma verbal que apresenta um evento ou uma ação descrito como completos.

(1) He wiped the table dry/ clean/ *wet/ *dirty

(2) He shot him dead/ *wounded⁸

No trabalho de Oliveira, buscou-se investigar se falantes nativos do português brasileiro com alta proficiência em inglês são sensíveis às idiossincrasias das construções resultativas que fazem com que a sentença acima varie em termos de gramaticalidade de acordo com o tipo de adjetivo utilizado ou se eles generalizam a regra de formação da construção em questão e aceitam a presença de quaisquer adjetivos. Os participantes do estudo (N=25), que consistiu em um experimento de Julgamento de Aceitabilidade com a escala Likert, tiveram a alta proficiência comprovada a partir do VLT (Vocabulary Levels Test) (NATION, 1990). Os participantes foram expostos a um conjunto de 48 frases, das quais 16 ilustravam a construção resultativa em questão, como nos exemplos acima, sendo oito sentenças agramaticais devido ao uso de adjetivos que não se enquadravam na construção, e oito gramaticais. Cada uma dessas 48 frases deveria ser classificada, a partir de uma escala Likert, de 1 a 7 em relação a sua gramaticalidade.

Os resultados do experimento indicaram que os falantes são capazes de diferenciar as construções resultativas gramaticais das agramaticais. O tratamento estatístico dos resultados foi realizado a partir de um teste t bicaudal ($t(24) = 3.94$, $df = 149$, $p \leq 0,01$) com amostras emparelhadas. Esse estudo, portanto, traz evidências de que, apesar do menor *input* em comparação com os falantes nativos de inglês, os bilíngues falantes nativos do português brasileiro e com alta proficiência em inglês são capazes de adquirir as especificidades que limitam os tipos de adjetivos que podem ser utilizados em construções resultativas.

⁸ Em português: (1) Ele esfregou a mesa seca/ limpa/*molha/ suja
(2) Ele o acertou (com um tiro) morto/ *ferido

A utilização de métodos estatísticos para análise dos dados obtidos em um julgamento de aceitabilidade é dependente de como o teste é entendido. Há duas visões diferentes em relação aos dados provindos de uma pesquisa que utiliza a Escala Likert: podemos considerá-los como dados ordinais ou intervalais. Dados ordinais não permitem a utilização de média e desvio padrão em sua análise, devido a sua natureza hierárquica. Com essa visão, é possível apenas utilizar dados como moda, mediana e porcentagem. Se os dados forem considerados intervalais, a média e o desvio padrão podem ser utilizados para o tratamento estatísticos dos dados, como foi feito no estudo de Oliveira. Uma opção para pesquisas que desejam evidenciar a intenção de se trabalhar com dados intervalais seria utilizar, no lugar de números, uma linha, na qual qualquer ponto do início ao fim pode ser marcado para julgar os estímulos.

Outra possível solução metodológica para se trabalhar com dados intervalais seria implementar um experimento cujo julgamento dos informantes fosse relativo a um outro item e não a um julgamento absoluto. O julgamento absoluto é aquele no qual os informantes precisam afirmar se um item contém ou não certa propriedade. Nesse caso, os informantes poderiam, por exemplo, classificar sentenças como gramaticais ou agramaticais, aceitáveis ou inaceitáveis, formais ou informais. Para isso, os informantes utilizariam um ponto de referência implícito a partir do qual os outros itens seriam julgados. Este tipo de julgamento apresenta algumas desvantagens, tais como a limitação das escalas e a inconsistência dos valores. Para obter resultados mais confiáveis é necessário que os pesquisadores utilizem julgamentos de aceitabilidade relativos, ou seja, em vez de caracterizar um item em termos absolutos é mais interessante julgá-lo em relação a outros. Dessa forma, é possível medir as diferenças entre aceitáveis e inaceitáveis; e, principalmente, medir o tamanho da preferência de um item em relação a outro. Sorace (2010) cita alguns experimentos que demonstram que as pessoas tendem a ter maior facilidade em fazer julgamentos relativos do que em

fazer julgamentos absolutos. Um método que pode ser utilizado para realizar experimentos com julgamentos relativos seria a estimativa de magnitude.

3.2 Estimativa de magnitude

A estimativa de magnitude é utilizada principalmente na psicofísica para determinar a sensação de uma pessoa a um estímulo (luz ou som, por exemplo) em relação a outro (SORACE, 2010; BARD et al., 1996). Basicamente, nesse tipo de experimento, um estímulo padrão é apresentado aos participantes que devem atribuir a esse estímulo padrão um valor que pode ser expresso por um número ou, até mesmo, por uma linha. Em seguida, outros estímulos são apresentados e os participantes devem atribuir valores que expressem a relação, proporção ou diferença entre os estímulos. Para ilustrar tal experimento, imagine que o participante é apresentado a um som e a este som ele atribui o valor 25. Em seguida, o som é apresentado duas vezes mais alto. Se o participante tem a sensação que o som está duas vezes mais alto então ele deve atribuir o valor 50, mas se a sensação dele foi que o som aumentou mais que duas ele pode atribuir qualquer valor acima de 50. Da mesma forma, se a sensação do informante foi que o som aumentou menos do que duas vezes ele atribuirá algum valor entre 25 e 50. Dessa forma, o paradigma da estimativa de magnitude é capaz de capturar a diferença entre estímulos tal como ela é percebida pelos participantes.

A maior diferença entre os estudos metafísicos e os estudos linguísticos está no fato de que na linguística só é possível trabalhar com os valores atribuídos pelos informantes já que não há “valores reais” em termos de gramaticalidade e aceitabilidade. De acordo com Sorace (2010), tal fato não gera problema algum às pesquisas linguísticas. A autora afirma que há diversos estudos que vêm corroborando

a ideia de que pesquisas que apenas trabalham com as opiniões dos participantes podem ser analisadas a partir da estimativa de magnitude. O uso desse paradigma em julgamentos de aceitabilidade forneceria dados mais confiáveis e precisos do que o uso das escalas convencionais, tais como a escala Likert, pois a estimativa de magnitude é mais sensível às diferenças sutis entre os estímulos (BARD *et al.*, 1996). Além disso, os informantes não precisam classificar as sentenças como boas ou ruins, apenas compará-las entre si.

O pesquisador tem diferentes opções de realização de experimentos de julgamento de aceitabilidade com o paradigma de estimativa de magnitude em relação ao estímulo padrão. Este pode ser escolhido previamente pelo pesquisador, que, usualmente, opta por estímulos com valores intermediários. Assim, os valores seguintes poderão ser maiores ou menores do que o estímulo padrão. Nas pesquisas metafísicas há uma preferência pela randomização ⁹ de todos os itens e, logo, o estímulo padrão não é definido previamente. Uma vantagem da primeira opção seria o fato de que ela apresenta melhor validade aparente já que ela parece exigir menos dos informantes. Independentemente da opção pela randomização, o pesquisador pode escolher se o estímulo padrão estará disponível durante todo o experimento ou se ele será apresentado apenas no início. Bem como a não randomização, a presença do estímulo padrão durante toda a condução do experimento facilita a execução do experimento por parte do participante e, logo, apresenta maior validade aparente.

Um cuidado que deve ser tomando ao se realizar um estudo com a estimativa de magnitude é a seleção do perfil dos participantes do estudo. Não se sabe exatamente se todas as pessoas têm a capacidade de fazer de forma confiável essa relação entre estímulos. É bem possível que crianças tenham maior dificuldade para atribuir valores e comparar diferentes estímulos. Assim, sugere-se que tal paradigma

⁹ Seleção de forma aleatória.

seja utilizado apenas para estudo com adultos. Além disso, é possível que a formação escolar tenha grande influência no julgamento de aceitabilidade linguística. Acredita-se, por isso, que pessoas com baixa alfabetização não devam compor este tipo de experimento, ao menos que elas sejam o alvo da pesquisa. Para obter dados mais confiáveis, o ideal é que todos os participantes tenham um perfil similar.

Outro fator que deve receber atenção na realização de experimentos com estimativa de magnitude é a normalização dos dados para que possa ser feito um teste estatístico. De acordo com Sorace (2010), tal procedimento é necessário porque as pessoas utilizam diferentes intervalos de estimativas e as distribuições tendem a ser distorcidas. Para que isso não interfira nos resultados o pesquisador tem duas opções. Ou ele transforma os valores crus em logaritmos a partir das médias geométricas de tendência central; ou ele divide cada valor numérico pelo valor do estímulo padrão.

Há diversas vantagens e desvantagens na utilização do paradigma da estimativa de magnitude. Este é capaz de captar com maior precisão o grau de variabilidade da aceitação dos itens e não impõe nenhum tipo de restrição ao julgamento dos informantes, que podem atribuir os valores que acharem mais condizentes a suas percepções. Além disso, não há necessidade de atribuir valores absolutos tais como “bom”, “agramatical”, “**”, ou “?”, é necessário apenas que os itens sejam comparados entre si. O paradigma permite a aplicação de estatística paramétrica e que os experimentos contenham um número de participantes considerável. Contudo, este tipo de experimento tende a requerer uma sessão de treinamento extensa, devido ao fato de que os participantes provavelmente não estão acostumados como esse tipo de tarefa e ela exige certo nível de habilidade por parte dos informantes para ser bem executada e, dessa forma, informantes devem ter um perfil bem similar.

4 OS CUIDADOS METODOLÓGICOS EM JULGAMENTO DE ACEITABILIDADE

Schutze (1996) apresenta uma série de cuidados que linguistas devem ter ao realizar experimentos que envolvam julgamento de aceitabilidade que se aplicam tanto a testes com escala Likert como aqueles que utilizam estimativa de magnitude. Segundo o autor, se todas as pesquisas seguissem as sugestões apontadas por ele, os dados provindos desse tipo de pesquisa científica seriam mais confiáveis.

Primeiramente, é preciso organizar todos os estímulos, incluindo os distratores, de forma que nem a ordem nem a gramaticalidade dos itens afetem os resultados. Para isso, é interessante que a ordem dos itens seja randomizada a cada sessão para que o nervosismo do início e a fadiga do final não interfiram nos julgamentos. Além disso, os itens podem ser controlados de forma que 50% deles sejam gramaticais e os outros 50% agramaticais. Dessa forma, os informantes terão menos chance de perceber algum tipo de tendência em relação à gramaticalidade dos itens. Tanto a proporção de sentenças gramaticais e agramaticais e a ordem dos estímulos, como aponta Schutze (1996), são comprovadamente fatores que influenciam no julgamento. Logo, esses cuidados são importantes para que os dados apontem o funcionamento da gramática do informante.

O julgamento de uma sentença pode ser influenciado pelo vocabulário que nela está presente. A presença de palavras com baixa frequência pode interferir no julgamento da frase como um todo. Uma forma de evitar que isso ocorra seria incluir no experimento apenas palavras cuja alta frequência seja comprovada por *corpus*. Alternativamente, o linguista poderia autorizar que os informantes indagassem sobre o significado de qualquer vocábulo. Independentemente de qual método escolhido, é fundamental que os participantes da pesquisa tenham consciência do significado de

todas as palavras da sentença julgada para que aspectos lexicais não venham a influenciar na aceitabilidade de uma frase.

Bem como o vocabulário, o contexto de uma construção pode fazer com que esta seja menos aceita. Uma mesma construção pode ser inserida em uma grande variedade de contextos. No entanto, nem todos podem ser compreendidos com a mesma facilidade. Em outras palavras, certos contextos demonstram ser mais apropriados para determinadas construções. Dessa forma, para obter dados confiáveis sobre uma construção é aconselhável que o linguista selecione os contextos que são pragmaticamente relacionados com a construção para evitar grande variação entre os sujeitos.

Para que pequenas diferenças entre as construções alvo sejam percebidas é essencial que haja paralelismo entre todas elas. No que se referem à sintaxe, as sentenças alvos devem apresentar os mesmos tipos de sintagmas, seguindo a mesma ordem, ao menos que a diferença de sintagmas ou de ordem seja o objeto de estudo. Além disso, sentenças sintaticamente complexas podem sobrecarregar a memória de trabalho de alguns indivíduos e conseqüentemente serem julgadas inadequadamente. Uma frase como “Maria demonstrou satisfação em mudar” não sobrecarrega a memória de trabalho da mesma forma que uma frase como “o fato de Maria ter dito para seus familiares que ela vai se mudar demonstra que é provável que ela foi sincera quando ela disse que não estava satisfeita com a situação na qual a cidade se encontrava”.

Considera-se importante também a manutenção da semântica para que construções não tenham suas estruturas julgadas inadequadamente devido ao seu sentido excêntrico, como em “Sonhos verdes comiam pensamentos ásperos”. Portanto, se a sintaxe e a semântica dos itens receberem os cuidados apropriados, maiores serão as chances de se chegar a resultados confiáveis.

Bem como os estímulos do experimento, os informantes precisam ser escolhidos criteriosamente para que eles não enviesem os dados. É bem possível que linguistas, por estarem envolvidos constantemente com questões de aceitabilidade e gramaticalidade, tenham julgamento diferenciado e, assim, não reflitam a população a que pertencem. Dessa forma, é preferível que não haja linguistas nos testes.

Acredita-se, ainda, que fatores como idade e educação também possam interferir no julgamento, logo, o ideal é que todos os informantes tenham nível de escolaridade e idade semelhantes. Além disso, o número de informantes deve ser o maior possível, para que os dados tratados estatisticamente tenham maiores chances de confiabilidade da amostra da população em questão. Assim, em uma pesquisa ideal, além das observações feitas anteriormente em relação aos estímulos, o linguista também precisa ser cuidadoso na seleção dos seus informantes. Ele deve ser capaz de reunir o número máximo possível de informantes com perfis semelhantes.

A aplicação do teste também deve ser controlada para que não comprometa os dados. Como foi dito anteriormente os termos “aceitabilidade” e “gramaticalidade” ainda são confundidos entre linguistas e podem também ser mal interpretados pelos informantes. Dessa forma, o ideal é que o experimentador explique o que ele quer dizer com aceitabilidade e julgue algumas sentenças para exemplificar o que foi proposto.

Para dar maior segurança, é necessário que haja alguns exemplos para que os indivíduos possam praticar a tarefa de julgamento de aceitabilidade, de forma que quando a tarefa se iniciar, todos os participantes estejam confortáveis. Portanto, é importante que o experimentador deixe todos os participantes certos do que eles devem fazer na tarefa.

Todos os experimentos que utilizam julgamento de aceitabilidade ainda têm a ajuda da tecnologia para facilitar e intensificar o seu uso. Há vários programas

disponíveis na *internet*, sendo alguns pagos e outros gratuitos, que facilitam a utilização de testes com escala Likert e estimativa de magnitude. O Webexp, por exemplo, é um software gratuito que permite que dados da escala Likert e da estimativa de magnitude sejam coletados via *internet*. Dessa forma, é possível se obter um número considerável de informantes e o programa facilita o acesso a populações (estrangeiros, por exemplo) com as quais a coleta presencial seria mais restrita.

CONCLUSÕES

O julgamento de aceitabilidade é um método *off-line* tão eficaz quanto outros métodos *on-line* e fisiológicos para iluminar questões relacionadas ao processamento e à estrutura da língua. Uma grande vantagem de se utilizar o julgamento de aceitabilidade é que este pode ser conduzido a partir de um computador sem o contato físico entre experimentador e informante. Pelo que sabemos, isso ainda não pode ser feito com métodos *on-line* ou com métodos fisiológicos.

ABSTRACT

Psycholinguistic studies shed light on what mental processes are connected with language. A vast array of methods is utilized in order to address this issue. The aim of this article is to analyze an offline method named Acceptability Judgment. Offline methods are those which are used to investigate language processing by using post-processing data. In this kind of method, subjects do tasks after being exposed to the research questions. In other words, the mental processes are not observed during the processing itself; instead, assumptions are made from what is perceived after the processing have been concluded. Acceptability Judgment refers to individuals' judgement concerning the acceptability of different sentences. In turn, the judgements are post processing data, which are used as a reflection of different mental process. The acceptability of the stimuli can be measured by using the Likert Scale paradigm, which is a psychometric scale made of different points, the most common ones are the

5-point-scale and the 7-point-scale. Each of these scale points refers to how acceptable each sentence is according to the judge. Another paradigm that can be utilized is the Magnitude Estimation, through which participants assign grades to each sentence to express the proportion and differences between each item. The Acceptability Judgement methodology is easily and reliably applied and it is considerably useful for psycholinguistics studies.

Keywords: Psycholinguistics, Likert scale, magnitude estimation

REFERÊNCIAS

- BARD, Ellen G.; ROBERTSON, Dan; SORACE, Antonella. Magnitude estimation of linguistic acceptability. *Language*, 72, p. 32-68. 1996.
- BROCCIAS, Cristiano. The English Change Network. Berlin/New York: Mouton de Gruyter. *Cognitive Linguistics Research Series*, 22, 2003.
- DEESE, James. *Psicolinguística*. Petrópolis: Vozes, 1976.
- FIELD, Jonh. *Psycholinguistics: The key concepts*. New York: Routledge, 2004.
- GERNSBACHER, Morton A. *Handbook of psycholinguistics*. San Diego : Academic Press, 1994.
- GLUCKSBERG, Sam. *Experimental psycholinguistics*. New Jersey: Lawrence Erlbaum Associates, 1975.
- JACKENDOFF, Ray; GOLDBERG; Adele. The English Resultative as a family of constructions. *Language*, Washington D.C., n. 80. setembro 2004, p. 523-567.
- LEITÃO, Márcio M. Psicolinguística Experimental: Focalizando o processamento da linguagem. In: MARTELOTTA, Mário Eduardo (org.) *Manual de Lingüística*. São Paulo: Contexto, 2008.
- LOBATO, Lúcia. Afinal, existe a construção resultativa em português? In: NEGRI, Lígia; FOLTRAN, Maria José; OLIVEIRA, Roberta Pires de (orgs.). *Sentido e Significação*. São Paulo: Contexto, 142-179, 2004.

MITCHELL, Don. On-line methods in language processing: Introduction and historical review. In: CARREIRAS, Manuel; CLIFTON, Charles (Eds.). *The On-line Study Comprehension: Eyetracking, ERP and Beyond*. Brighton, UK: Psychology Press, 2004.

NATION, Paul. *Teaching and Learning Vocabulary*. Boston, MA: Heinle & Heinle, 1990.

SCHULTZE, Carson. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago, IL: University of Chicago Press, 1996.

SOUZA, Ricardo; OLIVEIRA, Fernando de. Is knowledge of a non-dominant L2 activated by bilinguals using their dominant L1? Insights from an on-line psycholinguistic study. *Organon*, Porto Alegre, n. 51, 2011. p. 103-128.

SOUZA, Ricardo. Argument structure in L2 acquisition: Language transfer re-visited in a semantics and syntax perspective. Ilha do Desterro. *Journal of English Language, Literatures in English and Cultural Studies*, Florianópolis, n. 6. p. 153-188, 2011.

SORACE, Antonella. Using Magnitude Estimation in developmental linguistics research. IN: BLOM, Elma; UNSWORTH, Sharon (orgs.). *Experimental Methods in Language Acquisition Research*. Amsterdam/Philadelphia: John Benjamins, 2010. p. 57-72.

STARR, Matthew; RAYNER, Keith. Language Comprehension, Methodologies for Studying. In: *Encyclopedia of Cognitive Science*. p. 1358-1392. 2005.

WECHSLER, Stephen. An analysis of English resultatives under the event-argument homomorphism model of telicity. In: *Proceedings of the 3rd Workshop on Text Structure*. University of Texas at Austin, 2001.

ZARA, Julia. *Estudo da Expressão de Eventos de Transferência de Posse na Interlíngua de Brasileiros Aprendizes de Inglês*. Dissertação (Mestrado em Linguística Aplicada), 2009. 159 f., UFMG.