

**UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGÜÍSTICOS**

ISABELLE DEOLINDA PEREIRA DE SOUSA

**A VERIFICAÇÃO DE PROFICIÊNCIA LINGÜÍSTICA DE BILÍNGUES DO PAR
PORTUGUÊS BRASILEIRO-FRANCÊS AFERIDA POR TESTE DE
VOCABULÁRIO, AUTODECLARAÇÃO DE PROFICIÊNCIA, TEMPO DE ESTUDO
E FREQUÊNCIA DE USO**

BELO HORIZONTE

2019

ISABELLE DEOLINDA PEREIRA DE SOUSA

A VERIFICAÇÃO DE PROFICIÊNCIA LINGUÍSTICA DE BILÍNGUES DO PAR
PORTUGUÊS BRASILEIRO-FRANCÊS AFERIDA POR TESTE DE VOCABULÁRIO,
AUTODECLARAÇÃO DE PROFICIÊNCIA, TEMPO DE ESTUDO E FREQUÊNCIA DE
USO

Dissertação apresentada ao Programa de Pós-Graduação em Estudos Linguísticos da Faculdade de Letras da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Linguística teórica e descritiva.

Área de concentração: Linguística Teórica e Descritiva

Linha de pesquisa: Processamento da Linguagem

Orientador: Prof. Dr. Ricardo Augusto de Souza

BELO HORIZONTE

2019

S725v

Sousa, Isabelle Deolinda Pereira de.

A verificação de proficiência linguística de bilíngues do par português brasileiro-francês aferida por teste de vocabulário, autodeclaração de proficiência, tempo de estudo e frequência de uso [manuscrito] / Isabelle Deolinda Pereira de Sousa. – 2019.

88 f., enc. : il., tabs., grafs., p&b.

Orientador: Ricardo Augusto de Souza.

Área de concentração: Linguística Teórica e Descritiva.

Linha de pesquisa: Processamento da Linguagem.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Faculdade de Letras.

Bibliografia: f. 66-70.

Apêndices: 71-83.

Anexos: f. 84-88.

1. Bilinguismo – Teses. 2. Aquisição da segunda linguagem – Teses. 3. Língua francesa – Estudo e ensino – Falantes de português – Teses. 4. Língua francesa – Testes de aptidão – Teses. 5. Língua francesa – Vocabulário – Teses. I. Souza, Ricardo Augusto de. II. Universidade Federal de Minas Gerais. Faculdade de Letras. III. Título.

CDD : 407



UNIVERSIDADE FEDERAL DE MINAS GERAIS

PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGÜÍSTICOS

PosLin

FOLHA DE APROVAÇÃO

A VERIFICAÇÃO DE PROFICIÊNCIA LINGÜÍSTICA DE BILÍNGUES DO PAR PORTUGUÊS BRASILEIRO-FRANCÊS AFERIDA POR TESTE DE VOCABULÁRIO, AUTODECLARAÇÃO DE PROFICIÊNCIA, TEMPO DE ESTUDO E FREQUÊNCIA DE USO

ISABELLE DEOLINDA PEREIRA DE SOUSA

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ESTUDOS LINGÜÍSTICOS, como requisito para obtenção do grau de Mestre em ESTUDOS LINGÜÍSTICOS, área de concentração LINGÜÍSTICA TEÓRICA E DESCRITIVA, linha de pesquisa Processamento da Linguagem.

Aprovada em 14 de fevereiro de 2019, pela banca constituída pelos membros:

Prof(a). Ricardo Augusto de Souza - Orientador
UFMG

Prof(a). Marcus Guilherme Pinto de Faria Valadares
Colégio Santo Agostinho

Prof(a). Daniel Márcio Rodrigues Silva
UFMG

Belo Horizonte, 14 de fevereiro de 2019.

Aos meus avós Maria das Graças e Zilton.

A minha mãe Zilmária.

Às professoras e aos professores de FLE no
Brasil.

AGRADECIMENTOS

À magnitude espiritual que rege o Cosmos e nós, seres humanos ainda em evolução, por ter-me trazido à terrinha do pão de queijo, proporcionando-me novas experiências ao lado de pessoas incríveis.

Ao estimado orientador Ricardo pela sua humanidade para comigo.

Aos meus avós, Maria das Graças e Zilton que, desde o início, apoiaram-me em todos os sentidos na realização do mestrado objetivo. Sem eles, esse objetivo seria impossível.

Ao meu querido Rodrigo Telles, que permaneceu ao meu lado com seu coração disponível nas horas tristes e felizes da minha trajetória acadêmica. Em várias ocasiões, relembrou-me de minhas conquistas pessoais e do quanto sou capaz de lutar pelo o que desejo, além de me conceder seu apoio em todas as esferas possíveis desde a minha graduação até o infinito presente.

Aos meus pais, Zilmária e Marcos, e aos meus sogros, Franceuri e Francisco José Telles, que se mantiveram no apoio e nas vibrações positivas constantes.

Aos “míguilis” companheiros de mestrado Jessica Queiroz, Flávia Alvarenga, Clarice Fernandes (em especial a elas pelo apoio incondicional nas horas de desânimo), além dos queridos João Henrique Souza, Marcus Vinícius, Matheus Freitas e Amália Perna.

A toda equipe do laboratório de Psicolinguística da UFMG, especialmente Mara Guimarães, Cláudia Brandão e Thaís de Sá, pela amizade, sororidade e paciência no auxílio teórico-científico.

Às professoras Cristina Bustamante, Lúcia Jacob e Daniela Hirakawa, da Faculdade de Letras da UFMG, que me acolheram carinhosamente e me deram a oportunidade de me aprimorar profissionalmente.

Aos amigos e professores de francês da Casa de Cultura Francesa e do Departamento de Letras Estrangeiras, ambos da Universidade Federal do Ceará.

Aos meus colegas de estágio do Núcleo de Línguas Estrangeiras da Universidade Estadual do Ceará (e, em especial, às coordenadoras Adriana Barros e Andreia Carneiro), da Casa de Cultura Francesa, do Instituto Poliglota e do Centro de Extensão da Faculdade de Letras da UFMG.

Às amigas e aos amigos de Fortaleza que se mantiveram nas vibrações positivas ao longo dessa jornada.

Às mulheres cientistas que percorrem o caminho acadêmico, especialmente as da área da Psicolinguística, pela coragem e pela determinação com as quais enfrentam todas as injustiças de uma academia ainda machista, misógina e patriarcal.

A todas as alunas e alunos que muito me ensinaram, de forma voluntária e involuntária, valores para dentro e para fora da sala de aula e com quem pude compartilhar meus conhecimentos em Língua Francesa.

Aos participantes brasileiros e francófonos que bondosamente dedicaram seu tempo em participar da pesquisa para essa dissertação.

À equipe administrativa do POSLIN pela atenção, transparência e empatia nos trâmites acadêmicos.

Ao apoio financeiro do CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brasil, no último ano de mestrado.

“All I could do was to offer you an opinion upon one minor point – a woman must have money and a room of her own if she is to write fiction; and that, as you will see, leaves the great problem of the true nature of woman and the true nature of fiction unsolved.”

(Virginia Woolf)

RESUMO

Testes de vocabulário receptivo se configuram como instrumento de mensuração de proficiência desde a década de 1960 (LADO, 1961; HARRIS, 1969), sobretudo como avaliação de aquisição e de uso de palavras em salas de aula de ensino de línguas. Revelaram-se igualmente importantes diante da possibilidade de se corresponderem com os estudos psicolinguísticos referentes ao bilinguismo, uma vez que podem trazer informações a respeito da proficiência linguística em segunda língua (L2). Durante o processo de comunicação, mecanismos cognitivos subjacentes podem ser estudados por meio da automaticidade do acesso às representações lexicais, constituindo-se como um dos aspectos de proficiência em L2. Para algumas pesquisas nesse campo de estudo, faz-se imprescindível o uso de testes que avaliem as habilidades linguísticas do aprendiz em L2, porém alguns têm acesso restrito por serem pagos e por demandarem metodologias mais complexas de aplicação, tais como TOEFL para o inglês e DELF/DALF para o francês. Diante desse fator, Souza, Duarte e Berg (2015) acrescentam que os laboratórios de pesquisa experimental necessitam de testes psicométricos com a mesma eficiência desses supracitados, mas de aplicação prática, confiável e acessível. Batista (2014), então, propõe o *Teste de la Taille du Vocabulaire* (TTV) para preencher uma lacuna nos estudos psicométricos de acesso vocabular referentes à língua francesa. Com o intuito de verificar a proficiência em L2 de bilíngues tardios do par português brasileiro e de investigar como as medidas externas de autodeclaração de proficiência, tempo de estudo de L2 e frequência de uso de L2 se relacionam com o conhecimento lexical e com o construto proficiência, aplicam-se o teste de amplitude vocabular TTV e questionários voltados para a avaliação dessas três variáveis. Os resultados apontam medidas significativas de correlação moderada e fraca entre a pontuação no TTV com as medidas de autodeclaração de proficiência e de tempo de estudo, respectivamente, enquanto não se verifica correlação entre o desempenho no teste e a frequência de uso de L2.

Palavras-chave: bilinguismo, proficiência linguística, amplitude vocabular, teste psicométrico.

ABSTRACT

Receptive vocabulary tests have been an instrument of proficiency measurement since the 1960s (LADO, 1961; HARRIS, 1969), particularly to evaluate the acquisition and the use of words in language teaching classrooms. They have been equally important in face of the possibility of relating the tests to psycholinguistic researches on bilingualism, once they may bring forth information concerning the linguistic proficiency in a second language (henceforth L2). During the process of communication, underlying cognitive mechanisms can be studied through the automaticity of the access to lexical representations, which is an aspect of L2 proficiency. To some researchers the use of tests that evaluate learner's L2 ability is indispensable, but some tests are restricted for being paid and for demanding complex application methodologies, such as the TOEFL for English and the DELF/DALF for French. For that reason, Souza, Duarte and Berg (2015) argue that experimental research laboratories need to have psychometric tests as efficient as the ones just mentioned but they must be practical, reliable, and accessible. Batista (2014) proposes the Teste de la Taille du Vocabulaire (TTV) to fill the gap in psychometric studies that access the vocabulary knowledge of L2 French learners. Aiming at verifying linguistic proficiency in late bilinguals Brazilian Portuguese-French and to investigate how three external measures (self-rating proficiency, time studying the L2, and usage frequency of the L2) relate to the lexical knowledge and to the proficiency as a construct, questionnaires that evaluated these measures and the TTV were given for learners to answer. The results show significant measures of moderate and weak correlation between the TTV scores and measures of self-rating and time studying the L2, respectively, while there is no correlation between the test scores and the usage frequency of the L2.

Keywords: bilingualism, language proficiency, vocabulary breadth, psychometric test.

LISTA DE ILUSTRAÇÕES

Figura 1 – The stretched Z.....	24
Figura 2 – Exemplo de bloco do Vocabulary Levels Test (NATION, 1990).....	29
Figura 3 – Exemplo de bloco do TTV.....	30
Figura 4 – Fase de instrução inicial para o TTV.....	39
Figura 5 – Qui-quadrado de <i>Goodness of Fit</i> para verificar a distribuição de classificação dos sujeitos entre os níveis.....	49
Figura 6 – Qui-quadrado de comparação entre desempenho no TTV e autodeclaração de proficiência.....	60
Figura 7 – Qui-quadrado de união dos níveis iniciais e separação dos níveis mais avançados...	61
Figura 8 – Qui-quadrado de união dos níveis mais avançados e separação dos níveis iniciais...	61
Gráfico 1 – Histograma de níveis de autodeclaração de proficiência.....	42
Gráfico 2 – Histograma de tempo de estudo.....	43
Gráfico 3 – Histograma de frequência de uso de língua francesa.....	44
Gráfico 4 – Classificação dos participantes brasileiros no TTV por nível (N = 133).....	48
Gráfico 5 – Classificação dos participantes nativos no TTV por nível (N = 34).....	49
Gráfico 6 – Diagrama de dispersão entre tempo de estudo e autodeclaração de proficiência...	51
Gráfico 7 – Diagrama de dispersão entre frequência de uso e autodeclaração de proficiência...	51
Gráfico 8 – Diagrama de dispersão entre autodeclaração de proficiência e escore no TTV.....	54
Gráfico 9 – Diagrama de dispersão entre tempo de estudo e escore no TTV.....	56
Gráfico 10 – Diagrama de dispersão entre frequência de uso do francês e escore no TTV.....	57

LISTA DE TABELAS

Tabela 1 – Estatística descritiva de autodeclaração de proficiência dos participantes nas quatro habilidades linguísticas	42
Tabela 2 – Estatística descritiva de tempo de estudo de francês.....	43
Tabela 3 – Estatística descritiva de frequência de uso de francês por atividade.....	45
Tabela 4 – Estatística descritiva de frequência de uso relacionadas às habilidades linguísticas.	47
Tabela 5 – Interpretação do valor do coeficiente de correlação de Pearson (r).....	50
Tabela 6 – Análise de regressão linear múltipla das variáveis independentes tempo de estudo e frequência de uso com autodeclaração de proficiência.....	52
Tabela 7 – Análise de regressão linear das variáveis autodeclaração de proficiência e TTV.....	55
Tabela 8 – Análise de regressão linear das variáveis tempo de estudo e TTV.....	56
Tabela 9 – Análise de regressão linear das variáveis frequência de uso e TTV.....	58
Tabela 10 – Análise de regressão linear múltipla das variáveis autodeclaração de proficiência e tempo de estudo na pontuação do TTV.....	59

SUMÁRIO

1. INTRODUÇÃO	15
2 REFERENCIAL TEÓRICO	17
2.1 Bilinguismo, proficiência e automaticidade	17
2.2 Autodeclaração de proficiência	21
2.3 Idade de exposição à L2 e tempo de estudo	23
2.4 Frequência de uso de L2	25
2.5 Testes de vocabulário e proficiência linguística.....	27
2.6 Teste de la Taille du Vocabulaire (TTV)	29
3 METODOLOGIA	35
3.1 Questionário aplicado para brasileiros	35
3.1.1 Sujeitos.....	35
3.1.2 Materiais.....	36
3.2 Questionário aplicado para nativos	37
3.2.1 Sujeitos.....	37
3.2.2 Materiais.....	37
3.3 Aplicação do <i>Teste de la Taille du Vocabulaire</i>	38
3.3.1 Sujeitos.....	38
3.3.2 Materiais.....	39
4 ANÁLISES E DISCUSSÕES	41
4.1 Questionários.....	41
4.1.1 Dados do questionário aplicado para brasileiros.....	41
4.1.2 Dados do questionário aplicado para nativos.....	48
4.2 Desempenho no TTV	48

4.2.1 Brasileiros	48
4.2.2 Nativos	49
4.3 A relação das variáveis tempo de estudo e frequência de uso com autodeclaração de proficiência	50
4.4 O papel das variáveis autodeclaração de proficiência, tempo de estudo e frequência de uso na variabilidade da pontuação do TTV	53
4.4.1 Análise individual da relação entre as variáveis autodeclaração de proficiência, tempo de estudo e frequência de uso com o desempenho no TTV	54
4.5 Autodeclaração de proficiência e tempo de estudo relacionadas à variabilidade da pontuação geral do TTV	59
4.6 Comparação entre classificação no teste e autodeclaração de proficiência	60
5 DISCUSSÃO GERAL, LIMITAÇÕES E PERSPECTIVAS FUTURAS	63
REFERÊNCIAS	66
APÊNDICES	71
ANEXOS	84

1. INTRODUÇÃO

Os estudos relacionados ao processamento da linguagem por bilíngues constituem uma área ampla de investigação, uma vez que ainda há pontos de discussão referentes a problemas de operacionalização de pesquisas do campo relacionados aos participantes, à ativação das línguas no cérebro, aos estímulos e às tarefas experimentais, bem como aos modelos de aquisição, de representação e de processamento bilíngue (GROSJEAN, 1998). Um dos tópicos de discussão levantado pelo autor se refere à insuficiência e à controvérsia dos critérios de recrutamento de participantes bilíngues para as pesquisas experimentais, pois os pesquisadores costumavam estabelecer diferentes critérios de escolha de sujeitos para seus estudos. Dentre os fatores de seleção, estavam o nível de fluência, medido por meio de escalas ou testes; o objetivo de uso das línguas (quais línguas eram usadas e para quê); a estabilidade linguística, para saber se o participante ainda estava em processo de aprendizagem do idioma, e o contexto de aprendizado das línguas.

Apesar da problemática e de suas possíveis soluções terem sido discutidas por Grosjean no final dos anos 90, Hulstijn (2012) reflete que as formas de mensuração de proficiência linguística continuam sendo um assunto pertinente, através da análise dos trabalhos publicados na revista *Bilingualism: Language and Cognition* até o ano de 2012. O autor acrescenta que a proficiência opera um papel crucial nos estudos cognitivos do bilinguismo, e que alguns autores a abordavam com o objetivo de recrutar participantes para as pesquisas experimentais ou de estabelecer grupos de sujeitos com diferentes níveis de proficiência linguística. Esta era avaliada pelos próprios sujeitos ou acessada por meio de testes de produção vocabular oral (*Peabody Picture Vocabulary Test*, *Mean Length of Utterance*, *Boston Naming Test*), de decisão lexical escrita (*LLEX Vocabulary Test*), de vocabulário receptivo (*Vocabulary Levels Test*), de tarefas de tradução e de julgamento.

Testes psicométricos de vocabulário receptivo se configuram como instrumentos de mensuração de habilidade linguística desde Lado (1961) e Harris (1969), sobretudo como avaliação de aquisição e de uso de palavras em salas de aula de ensino de línguas. Além de serem requeridos para fins educacionais no âmbito do ensino de segunda língua (doravante L2) devido ao interesse em se saber como os aprendizes desenvolvem o conhecimento e a habilidade de uso lexical em L2 (READ, 2000, p. 8), revelaram-se igualmente importantes diante da possibilidade de se corresponderem com os estudos psicolinguísticos referentes ao bilinguismo.

Assim, o presente estudo busca verificar a proficiência em L2 de bilíngues tardios do par português brasileiro e investigar como as medidas externas de autodeclaração de proficiência, tempo de estudo de L2 e frequência de uso de L2 se relacionam com o conhecimento lexical, aferido pelo teste de amplitude vocabular *Teste de la Taille du Vocabulaire* (TTV), de Batista (2014).

A hipótese que guia este estudo é a de que o desempenho no TTV pode demonstrar uma correlação direta com as medidas externas citadas anteriormente (autoavaliação de proficiência, tempo de estudo de L2 e frequência de uso de L2). Isso significa que bilíngues que declaram altos níveis de proficiência, que contabilizam maior tempo de estudo e maior uso da língua francesa podem ter maior pontuação no TTV, classificando-se em níveis mais altos no teste; já os que declaram baixos níveis de proficiência, apresentando menor tempo de estudo e menor frequência de uso, podem obter menor pontuação, classificando-se em níveis mais baixos.

Essa discussão ainda se faz relevante por haver o costume de se presumir que, para se ter um melhor desempenho em L2, o aprendiz precisa estar geograficamente situado em um país que tenha um determinado idioma como língua oficial para que ele possa ter o máximo de *input* e, com isso, conseguir exercer as habilidades ativas (como expressão oral e escrita) com maior êxito. Porém, com o aumento do acesso à tecnologia, a praticidade em seu uso e a motivação de aprendizagem do indivíduo, a imersão virtual (quando o aluno não está imerso geograficamente, mas encontra meios de obtenção de *input*) pode ser um fator de contribuição para o aprimoramento das habilidades em língua adicional do falante.

A organização deste trabalho está desenvolvida da seguinte maneira: o capítulo 2 traz a revisão dos referenciais teóricos para se compreender os conceitos de bilinguismo e de proficiência, bem como a relação dessa última com as medidas de autodeclaração, tempo de estudo de L2 e frequência de uso de L2, além do papel dos testes psicométricos de vocabulário que tratam o construto mental ligado ao processamento linguístico bilíngue. O capítulo 3 contém a descrição da metodologia desenvolvida e as estatísticas descritivas. No capítulo 4, há as análises e suas discussões. Por fim, o capítulo 5 versará sobre a verificação da hipótese principal atestada pelos resultados encontrados a partir do intervalo de dados fornecidos e alcançados pelos participantes da pesquisa, além de tratar as limitações do presente estudo e algumas perspectivas para futuras pesquisas.

2 REFERENCIAL TEÓRICO

2.1 Bilinguismo, proficiência e automaticidade

Diante de diferentes tipos de perfil de bilíngues, Grosjean (1998) já apontava alguns problemas nas pesquisas que tratavam e ainda tratam (HULSTIJN, 2012) sobre bilinguismo, principalmente no que diz respeito aos participantes, à ativação das línguas no cérebro¹, aos estímulos e tarefas experimentais, além dos modelos de aquisição, de representação e de processamento bilíngue. É evidente que todos esses setores são de extrema relevância, entretanto cabe ressaltar que duas questões apontadas sobre os participantes têm relação direta com esta pesquisa: (i) alguns pesquisadores apresentam dificuldades em saber quem de fato são os bilíngues e (ii) utilizam-se de diferentes fatores, insuficientes e até mesmo controversos na seleção de participantes para as pesquisas experimentais (GROSJEAN, 1998, p. 133).

A respeito do primeiro ponto de discussão sobre os participantes, um breve esclarecimento ainda se faz necessário sobre a definição de bilinguismo e sobre o perfil desses indivíduos. Os sujeitos que o presente estudo se propõe investigar são considerados bilíngues por usarem duas ou mais línguas ou dialetos (bi/multilinguismo) em seus cotidianos (GROSJEAN; LI, 2013, p. 5), mesmo em contexto de não-imersão geográfica. São, então, bilíngues brasileiros tardios do par português-francês, pois a aquisição do francês começou após o início da aquisição da primeira língua, doravante L1 (COSTA; SEBASTIÁN-GALLÉS, 2014, p. 337), o português brasileiro.

Para uma melhor compreensão do perfil dessa população, bilíngues adquirem e usam suas línguas para diferentes propósitos e em diversos domínios de suas vidas (acadêmico, familiar, profissional etc.), o que constitui o Princípio da Complementaridade (GROSJEAN, 1998; GROSJEAN; LI, 2013), e raramente são fluentes de maneira igualitária tanto nas línguas faladas como nas habilidades linguísticas relacionadas a um só idioma. Ainda em consonância com esse princípio, o repertório linguístico do bilíngue pode sofrer alteração à medida que o ambiente e as necessidades para uma certa habilidade linguística mudam (GROSJEAN, 1998, p. 132).

¹ De acordo com Grosjean e Li (2013, p. 15), há dois modos linguísticos envolvidos em uma interação com interlocutor ou em outras situações (escrita, leitura etc.). O modo monolíngue se caracteriza pelo uso de uma língua enquanto a(s) outra(s) permanecem inativas. Já no modo bilíngue, o indivíduo permanece com as línguas ativadas, mas com predominância da língua de base, podendo se utilizar de elementos da outra língua ou fazer mudanças de código linguístico.

O segundo tópico de discussão levantado por Grosjean (1998) se refere aos fatores de recrutamento de participantes bilíngues para as pesquisas experimentais. O autor comenta que, na época da publicação de seu artigo, havia pesquisadores que estabeleciam diferentes critérios de escolha de sujeitos para as pesquisas. Dentre os fatores de seleção, estavam o nível de fluência, medido por meio de escalas ou testes; o objetivo de uso das línguas (quais línguas eram usadas e para quê); a estabilidade linguística, para saber se o participante ainda estava em processo de aprendizagem do idioma, e o contexto de aprendizado das línguas. Conseqüentemente, devido à diversidade dos critérios usados ao se selecionar sujeitos e às diferentes tarefas experimentais, Grosjean (1998, p. 134) alerta sobre a possibilidade de variabilidade nos grupos de bilíngues das pesquisas mencionadas ao longo de seu artigo, tanto dentro de um determinado estudo quanto em comparação de grupos entre as pesquisas.

Para tentar solucionar os problemas supracitados, o autor propõe que as medidas de acesso bilíngue (fluência, uso das línguas, estabilidade e contexto de aprendizagem) tornem-se variáveis covariantes nas análises dos resultados, a fim de se verificar o efeito umas sobre as outras (GROSJEAN, 1998, p. 135) e que, principalmente, os trabalhos em Psicolinguística Experimental tragam informações dos indivíduos estudados, tais como dados biográficos (idade, sexo, nível instrucional), histórico linguístico (idade de aquisição das línguas, forma de aprendizado do idioma), estabilidade linguística (para saber se os participantes ainda estão em processo de aprendizagem da língua), objetivos de uso dos idiomas, proficiência nas habilidades linguísticas e modo de ativação linguística (monolíngue ou bilíngue).

Apesar da problemática e de suas possíveis soluções terem sido apontadas por Grosjean no final dos anos 90, o trabalho de Hulstijn (2012) demonstra que as formas de mensuração de proficiência linguística continuam sendo um assunto pertinente. A publicação analisa os trabalhos contidos na revista *Bilingualism: Language and Cognition* (do volume 1 ao 14) e foca nas questões sobre definição e mensuração de proficiência linguística, bem como suas implicações sobre elucidação e operacionalização da dominância linguística² (associado

² Vale salientar que os conceitos de proficiência e de dominância são diferentes. Segundo Montrul (2016, p. 16 e 17) o primeiro termo é definido como habilidade linguística e fluência (caracterizada pela velocidade de recuperação lexical em habilidades de produção e de recepção, como escrita, leitura, fala, escuta, tradução etc.) em uma língua, usado tanto em aquisição de primeira como de segunda língua. Em contrapartida, o conceito de dominância refere-se a uma certa influência e relação das línguas faladas pelo bilíngue, no que concerne ao uso e ao grau de proficiência. A língua dominante é geralmente aquela na qual o falante recebe mais *input* e a mais usada em contextos específicos, tal como previsto pelo Princípio da Complementaridade (GROSJEAN, 1998; GROSJEAN; LI, 2013).

ao conceito de bilíngue balanceado³). Fica evidente, então, o interesse de estudo do bilinguismo através de uma perspectiva cognitiva. Vários trabalhos também abordam o papel da proficiência em uma ou em todas as línguas dos participantes, seja avaliada pelos próprios sujeitos ou acessada por meio de outras ferramentas, como testes de produção vocabular oral (*Peabody Picture Vocabulary Test, Mean Length of Utterance, Boston Naming Test*), de decisão lexical escrita (*LLEX Vocabulary Test*), de vocabulário receptivo (*Vocabulary Levels Test*), tarefas de tradução e de julgamento, além de certificados de proficiência linguística em língua inglesa, como *Cambridge Certificate of Proficiency, International English Language Testing System - IELTS, Oxford Quick Placement Test* e *Test of English as a Foreign Language - TOEFL*⁴ (HULSTIJN, 2012, p. 423).

Tal como apontado por Silva (2016, p. 36), a proficiência em L2 é geralmente tomada como um construto multifacetado e correlacionado a diferentes medidas com o intuito de ser um conceito relativamente coesivo. É por isso que, como mostrado no parágrafo acima, testes que cobrem vários componentes de habilidades linguísticas são aplicados e correlacionados em uma tentativa de se estabelecer uma medida ampla para esse construto. O autor adiciona que é preciso entender a natureza dos conhecimentos envolvidos para se compreender a definição de proficiência (SILVA, 2016, p. 36) e sua relação com a automaticidade.

Segundo Ellis (2015, p. 3), o aprendizado implícito acontece naturalmente, de forma simples e inconsciente, no momento da aquisição do conhecimento de uma estrutura subjacente a partir de um ambiente de estímulo complexo. Por outro lado, o aprendizado explícito é consciente, pois o aprendiz pode formular e testar hipóteses para uma certa estrutura da língua. A aquisição do conhecimento de uma língua pode ocorrer, então, de maneira implícita (uma abstração automática e inconsciente da estrutura de um material proveniente de experiências linguísticas), explicitamente pela aprendizagem seletiva (quando o aprendiz busca por informação, construindo e testando hipóteses) ou explicitamente por meio de regras

³ O bilíngue balanceado seria o indivíduo igualmente fluente nos idiomas falados e nas habilidades pertinentes a cada língua, algo raro de ser encontrado de acordo com Grosjean (1998, p. 132).

⁴ Com a variedade de certificados de proficiência para diferentes idiomas - como o *Diplôme d'Études en Langue Française* (DELFL) para o francês, o *Diploma de Español como Lengua Extranjera* (DELE) para o espanhol, o *Certificazione di Italiano come Lingua Straniera* (CILS) para o italiano, entre outros -, pode-se questionar o porquê de não utilizá-los como única ferramenta para o recrutamento dos sujeitos de pesquisa e/ou para a mensuração de proficiência. Tal como apontado por Souza, Duarte e Berg (2015, p. 274), laboratórios de pesquisa experimental necessitam de testes psicométricos com a mesma eficiência desses mencionados, mas de aplicação prática, confiável e acessível, pois, além de serem pagos, sua aplicação pode levar horas, não se adequando às necessidades reais de laboratórios de pesquisa experimental brasileiros.

específicas (na observação de instruções explícitas para a assimilação de uma determinada regra).

O uso e a interação dinâmica entre esses dois tipos de aprendizado formam um complexo sistema de linguagem relacionado envolvendo similarmente a cognição, de acordo com Ellis (2015, p. 16), pois organiza-se, processa-se e transmite-se informação da mente de um locutor para a mente de um interlocutor através da linguagem.

Relacionando-se com a cognição, o impacto do bilinguismo vem sendo percebido por meio das funções executivas, ou seja, por processos que envolvem atenção, seleção, inibição, monitoramento e flexibilidade, desenvolvendo-se paralelamente com o córtex pré-frontal (BIALYSTOK; BARAC, 2013, p. 202). Ainda segundo os pesquisadores, três habilidades principais estão no centro do campo de estudo do bilinguismo: memória de trabalho ou atualização (manutenção de informação na mente e capacidade de manipulá-la mentalmente), controle inibitório (resistência para informações irrelevantes) e flexibilidade cognitiva (troca de objetivos e adaptação para mudanças de acordo com as exigências ou com as prioridades).

Bialystok e Barac (2013, p. 193) lembram que, durante 1970 e 1980, procuravam-se os efeitos do bilinguismo no domínio da competência linguística, precisamente na consciência metalinguística, em crianças monolíngues e bilíngues. A consciência metalinguística se relaciona com o aprendizado explícito de uma estrutura linguística e com a habilidade de acessá-la de forma intencional, permitindo que crianças separem o significado de uma palavra de sua forma e que façam julgamentos independentes sobre aspectos fonológicos, semânticos, sintáticos e morfológicos de uma língua⁵. Entretanto, parte do sucesso de crianças bilíngues nessa tarefa é resultado de suas vantagens no controle inibitório, isto é, na atenção seletiva da informação relevante – a forma da palavra – e na inibição da informação competidora para a tarefa – o significado da palavra (BIALYSTOK; PEETS; MORENO, 2014, p. 179).

Em contrapartida, a consciência metalinguística não estaria no centro dos estudos de desenvolvimento linguístico da criança, mas a proficiência. Bialystok, Peets e Moreno (2014, p. 179) comentam que as medidas de proficiência relacionadas à estrutura linguística, ao

⁵ Essa conclusão se baseia no estudo de avaliação de consciência sintática e semântica de Bialystok (1986). Nesse trabalho, crianças bilíngues entre 5 e 9 anos tiveram melhor performance que as monolíngues de mesma idade no reconhecimento de frases com violações de forma e de significado, assim como no reconhecimento das que apresentavam estrutura formal correta e sentido equivocado, como em *Apples grow on noses* (Maças crescem em narizes).

tamanho de vocabulário e ao acesso a palavras e conceitos tendem a favorecer os monolíngues. Especificamente sobre o vocabulário receptivo, crianças e adultos monolíngues têm melhor desempenho em testes de vocabulário do inglês que os bilíngues em idades equivalentes (BIALYSTOK *et al.*, 2010; BIALYSTOK; LUK, 2012). Os autores ainda ponderam que essas diferenças de desempenho entre a consciência metalinguística e o vocabulário receptivo podem ser explicadas pelo fato de que o desempenho metalinguístico pode ser intensificado pelo controle inibitório, porém restrito pelo conhecimento linguístico formal.

Ainda no que concerne à cognição, outro fator relacionado com a proficiência linguística é a automaticidade. Jiang (2007, p. 2) discute que nem todos os conhecimentos linguísticos de um aprendiz de L2 são recuperados e aplicados rapidamente em uma situação de comunicação espontânea. Apesar do aprendiz poder recuperar certos conhecimentos pertinentes à L2 sem esforço e sem consciência metalinguística, outros podem demandar mais recursos atencionais. Dessa forma, o autor define automaticidade como a habilidade de performance em L2 sem realização consciente ou de pouca utilização de recursos atencionais (JIANG, 2007, p. 2).

Para esse estudo, especificamente, automaticidade é um dos componentes de proficiência e significa menor custo de processamento linguístico em L2, envolvendo facilidade e velocidade de acesso na ativação de representações linguísticas, como abordado por Souza, Duarte e Berg (2015, p. 279), relacionando-se com um melhor desempenho das funções executivas (BIALYSTOK; CRAIK; LUK, 2012).

A partir do que foi exposto anteriormente, percebe-se que a proficiência não é um construto isolado, sendo analisada por meio de diversas ferramentas empregadas para as pesquisas experimentais (GROSJEAN, 1998; HULSTIJN, 2012). Algumas medidas externas que dizem respeito a esse construto servirão de parâmetro de correlação entre outras medidas no presente trabalho. Nas subseções a seguir, algumas dessas medidas, como a autodeclaração de proficiência, o tempo de estudo de L2 e a frequência de uso de L2 serão examinadas de acordo com o que vem sendo divulgado por certas pesquisas na área do bilinguismo.

2.2 Autodeclaração de proficiência

A autodeclaração de proficiência, também nomeada de autoavaliação, é um instrumento constantemente usado nas pesquisas científicas para se obter informações sobre a proficiência linguística – através da seleção de participantes ou da formação de grupos em

diferentes níveis de proficiência para a participação em estudos experimentais, por exemplo (GROSJEAN, 1998; HULSTIJN, 2012) – ou sobre a dominância linguística (TREFFERS-DALLER, 2016), como avaliação de performance dos estudantes e nas salas de aula de línguas (ROSS, 2006).

Treffers-Daller (2016, p. 254) pondera que as pesquisas de bilinguismo costumam pedir aos informantes adultos que estimem suas próprias competências na leitura, na escrita, na fala e na escuta. Contudo, é questionável se essas autoavaliações sejam detalhadas suficientemente para que os informantes estejam conscientes da complexidade da proficiência e do uso de suas línguas. Haveria, então, duas perspectivas diferentes ao lidar com a autodeclaração: a construção da avaliação e a competência de apreciação dos próprios informantes.

Hulstijn (2012, p. 425) comenta sobre o questionário de autoavaliação de Bachman e Palmer (1989)⁶ para a língua inglesa, no qual há três competências linguísticas (gramática, pragmática e sociolinguística) e três tipos de perguntas (sobre habilidade – o quanto da gramática da língua o sujeito sabe –, sobre dificuldade – quantos erros gramaticais de diferentes tipos o sujeito comete em inglês – e sobre reconhecimento – o quão frequentemente o sujeito percebe os erros gramaticais de outrem). O autor também aponta alguns possíveis problemas nas escalas globais de análise nesse e em outros questionários usados em algumas pesquisas de bilinguismo, como a falta de discriminação entre as quatro habilidades (escuta, fala, leitura e escrita); entre os controles de fonética, de prosódia, de fonologia, de morfologia e de sintaxe; a pouca informação sobre a fluência na língua, tomando-se, por fim, o falante nativo como modelo ideal, não havendo diferença de aspectos educacionais e/ou profissionais nesse tipo de falante.

Anos após a publicação do questionário supracitado, Bachman e Palmer (2010 *apud* TREFFERS-DALLER, 2016, p. 254)⁷ observam que a avaliação das habilidades linguísticas relacionadas apenas às esferas de escuta, fala, leitura e escrita ignoraria o fato de haver, por exemplo, distinções entre tarefas de escrita, uma vez que escrever um *e-mail* demanda diferentes competências se comparado à formulação de um artigo científico. Uma alternativa proposta por Treffers-Daller (2016, p. 254) seria o uso de questionários que contemplassem os vários componentes de habilidade linguística, porém Hulstijn (2012, p. 425)

⁶ O portal de periódicos da CAPES possui acesso aos artigos da revista no qual o trabalho foi publicado somente a partir do ano de 1999, não havendo possibilidade de acesso direto à obra.

⁷ Impossibilidade de acesso direto à obra devido a não disponibilidade do livro nem em meio físico nem *online*.

pondera que é incerto saber se os informantes seriam realmente qualificados em fornecer medidas válidas de acesso a respeito dos diversos nichos de proficiência, como pronúncia, fluência, tamanho de vocabulário e gramática.

Parte-se, assim, para o segundo ponto de tensão relacionado à aplicação desse tipo de avaliação: os sujeitos. Além do problema na construção das perguntas de mensuração de autoavaliação, indaga-se se os próprios indivíduos são capazes de fazer avaliações de si mesmos (ROSS, 2006; HULSTIJN, 2012). Zell e Krizan (2014) perceberam que a correlação entre a habilidade de autoavaliação e as medidas de performance eram mais fortes quando aquela dizia respeito a um contexto específico, ao invés de amplo, e quando a performance era de baixa complexidade ou objetiva. Essas conclusões, segundo os autores, podem indicar que os sujeitos compreendem suas habilidades de maneira modesta, mas colocam em evidência os fatores que possibilitam essa autopercepção.

Portanto, apesar do uso da autodeclaração nas pesquisas do campo do bilinguismo ter suas ressalvas, ela continua sendo usada frequentemente. O interesse em se utilizar essa variável para a presente pesquisa é motivado pela intenção de saber como ela se comporta em relação às outras medidas externas de proficiência.

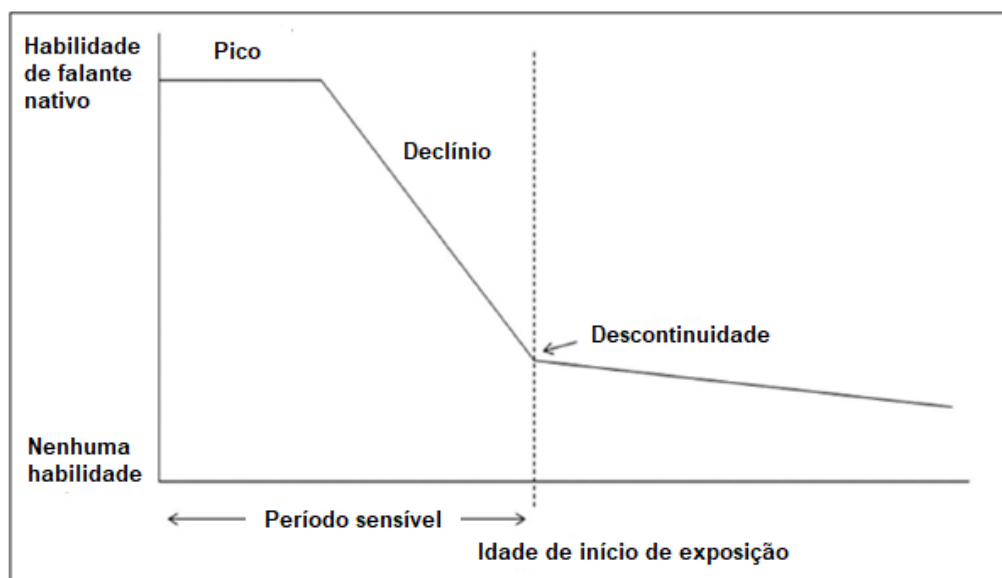
2.3 Idade de exposição à L2 e tempo de estudo

Uma outra medida usada para se verificar proficiência é a idade de início de exposição significativa em L2, uma vez que a predição seja a de que, quanto mais tardia for essa exposição, mais difícil seria para o indivíduo alcançar bons resultados no aprendizado de uma L2. Granera e Long (2012, p. 312) acrescentam que algumas pesquisas experimentais demonstram que esse *input* não seria um fator determinante para alta proficiência em L2. Um exemplo disso seriam os falantes não nativos com bastante tempo de residência em ambiente de L2, normalmente casados com nativos, usando a L2 em vários domínios de suas vidas, porém, mesmo nessas condições, podem não demonstrar um alto nível de proficiência.

Sabe-se que o bilinguismo influencia as habilidades cognitivas, modificando as estruturas e as funções cerebrais, pois várias pesquisas demonstram que o declínio de algumas dessas habilidades ocorre mais tardiamente em bilíngues que em monolíngues (BIALYSTOK; CRAIK; LUK, 2012). Entretanto, os efeitos trazidos pela idade do indivíduo são robustos independentemente do perfil psicossocial do aprendiz, da quantidade e da qualidade de *input* disponível (GRANERA; LONG, 2012, p. 312).

Granera e Long (2012, p. 313) mostram o que normalmente é o padrão, a partir de algumas pesquisas, a respeito do aprendizado de uma L2 com exposição precoce do indivíduo através da distribuição de Z esticado⁸ (*Stretched Z distribution*), mostrada na figura 1. O eixo horizontal representa temporalmente a idade do falante, enquanto o eixo vertical refere-se ao grau de habilidade linguística do sujeito, variando de nenhuma habilidade à habilidade tal qual de nativo da língua. O período de sensibilidade máxima ao *input* com obtenção plena das competências gramaticais em L2 ocorre na primeira infância (BIRDSONG, 2005, p. 112), pois quanto mais precoce for a exposição em L2, maiores serão as chances de o indivíduo desenvolver habilidades linguísticas mais próximas de um nativo. Em seguida, há o declínio do período sensível⁹, findando-se no momento em que a maturação cognitiva é alcançada (BIRDSONG, 2005, p. 113). Por fim, os efeitos da idade surgem após o término do período sensível à medida que as habilidades linguísticas do sujeito apresentam decaimento. Evidentemente a exposição à L2 não se configura como uma variável isolada durante a aquisição de uma L2, pois o aprendizado em L2 pode se manter constante ou decair a depender de outros fatores, como motivação, conhecimento metalinguístico, uso e exposição de língua (GRANERA; LONG, 2012, p. 314).

Figura 1 – *The stretched Z*



Fonte: GRANERA; LONG, 2012, p. 313

⁸ Tradução nossa.

⁹ Birdsong (2005) e Granera e Long (2012, p. 337) preferem a nomenclatura período sensível ao invés de período crítico (LENNEBERG, 1967) devido ao fato de estar sujeito à variação individual, além de o declínio da capacidade humana de aprendizado não ser tão acentuado tal qual a perda de algumas habilidades de desenvolvimento observada em animais.

Granera e Long (2012, p. 314) citam resultados de algumas pesquisas que discorrem sobre os períodos de aquisição do léxico e das colocações, relacionando-se a restrições de maturação cognitiva do sujeito, nas quais o pico do período sensível seria de 0 a 6 anos, seguido de seu declínio entre 6 e 12 anos de idade. Enquanto novos itens lexicais e colocações podem ser aprendidos constantemente por nativos e não nativos ao longo da vida, a habilidade de aprendizado lexical pode decair em função do aumento da idade.

A pesquisa dos autores supracitados examinou o desempenho de indivíduos de alta proficiência em L2 nas tarefas de produção e de recepção em fonologia, morfologia, sintaxe, léxico e colocações. O objetivo foi o de investigar se haveria relação entre a idade de exposição e esse desempenho, sendo consistente com a existência de períodos sensíveis, e se a aptidão linguística e/ou o tempo de residência em um país de L2 se relacionariam com a alta proficiência. No que concerne ao léxico (interesse principal para esta pesquisa) e às colocações, os resultados apontaram uma correlação inversa entre as pontuações das tarefas e a idade de exposição à L2 de crianças acima de 6 anos de idade, havendo também uma certa relação com o tempo de residência.

Diante do importante papel da idade de exposição de um indivíduo à L2, utiliza-se, na presente pesquisa, a medida de tempo de estudo em relação à idade de exposição¹⁰. Apesar da exposição à língua ser menor para os participantes desta pesquisa devido ao contato com a L2 ter sido tardio e ser diferente da que é tratada nos estudos acima, nos quais há em localidades de L2, o obstáculo geográfico não é considerado um grande problema, uma vez que o acesso *online* para filmes, séries de TV, músicas, *podcasts*, leitura, contato com nativos, etc. podem ser fatores de exposição à língua, além do ambiente de sala de aula.

2.4 Frequência de uso de L2

De acordo com Ellis (2012, p. 7), a frequência de uso é um outro fator que pode influenciar o aprendizado, a memória e a percepção cognitiva, uma vez que quanto mais se tem experiência em algo, menos recurso de memória é demandado e mais rápido é o acesso a tal conhecimento.

Assim também ocorreria com o processamento de uma língua, a partir dos estudos psicolinguísticos (ELLIS, 2012, p. 9), por ser sensível à frequência de uso nos domínios de

¹⁰ Subtraindo-se o tempo de estudo da idade declarada pelos indivíduos adultos, é possível estimar a idade do aprendiz quando a exposição ao francês se iniciou.

representação da linguagem, como fonologia, fonotaxe, leitura, ortografia, léxico, morfossintaxe, escuta, gramaticalidade, produção oral e sintaxe. Quanto mais regular e frequente for uma estrutura, ela será aprendida e processada mais facilmente. Contudo, a percepção humana pode ser determinada por limitações. Um exemplo fornecido pelo autor (ELLIS, 2012, p. 20) seriam os artigos *the* e *a* do inglês, que, apesar de possuírem alta frequência na língua, ainda se caracterizam como construções de processamento cognitivo custoso para os aprendizes de inglês como L2.

Ellis (2012, p. 152) ainda acrescenta que os efeitos relacionados à prática são maiores nos estágios iniciais da aprendizagem, pois seria por meio da prática constante que haveria a incrementação no tempo de processamento de informação lexical, permitindo uma sintonia entre os sistemas motor e perceptivo moldada pela experiência linguística. O autor traça um breve apanhado pontual sobre algumas descobertas que abordam a interação entre léxico e frequência e mais alguns trabalhos podem reforçar esse tópico.

Para a escrita, palavras de alta frequência são nomeadas com mais rapidez que as de baixa frequência, são mais rápidas em serem reconhecidas como palavras em tarefas de julgamento e são mais bem escritas de acordo com as regras de ortografia da respectiva língua. Crianças e adultos demonstrariam rapidez e acurácia nos processos de reconhecimento e de produção lexical em L1 e em L2, como percepção de fala/sinal, leitura, nomeação de objetos, fala, escrita e sinalização de línguas de sinais, relacionados à frequência de palavras (KIRSNER, 1994 *apud* ELLIS, 2012, p. 152). Valadares (2017) verificou que bilíngues brasileiros tendo o inglês como L2 tendem a demonstrar melhor desempenho no *Vocabulary Size Test* (NATION; BEGLAR, 2007) quanto maior o envolvimento em atividades de letramento digital em inglês¹¹.

A relação entre frequência e léxico pode ser verificada inclusive na aquisição de L1, uma vez que o conhecimento subjacente ao uso fluente de uma língua não seria puramente gramatical, mas formado igualmente de lembranças de enunciados conhecidos (ELLIS, 2002, p. 166). Clark (2015), reforçando essa possível associação, assinala que crianças em tenra idade, vivenciam padrões de rotina, tais como situações de interação no banho, ligadas à água ou ao fato de sentir sede, usados por meio de inserção e de memorização devido às repetições dessas situações e da ação conjunta de *feedback* entre a criança e o adulto.

¹¹ Alguns exemplos desse tipo de atividade estão descritos na tabela 3 desta pesquisa, tendo por base o questionário aplicado por Valadares (2017).

Dessa forma, nota-se que a frequência de uso pode se configurar como indicadora de proficiência linguística, havendo indícios de relação com as habilidades cognitivas diante da maneira como as informações de uma língua são processadas pelos falantes.

Além dessa possível medida, testes de medição vocabular validados são amplamente aplicados no intuito de investigar se as habilidades lexicais de um falante em L2 podem ser preditoras de proficiência. A subseção seguinte trata esse ponto de vista.

2.5 Testes de vocabulário e proficiência linguística

Testes de vocabulário receptivo se configuram como instrumentos de mensuração de habilidade linguística desde a década de 1960 (LADO, 1961; HARRIS, 1969), sobretudo como avaliação de aquisição e de uso de palavras em salas de aula de ensino de línguas. Além de serem requeridos para fins educacionais no âmbito do ensino de L2 devido ao interesse em se saber como os aprendizes desenvolvem o conhecimento e a habilidade de uso lexical em L2 (READ, 2000, p. 8), revelaram-se igualmente importantes diante da possibilidade de se corresponderem com os estudos psicolinguísticos referentes ao bilinguismo.

A conexão entre conhecimento lexical e proficiência linguística não seria vista apenas pela óptica do conhecimento do significado das palavras, seus sinônimos e suas correspondências entre L1 e L2 (READ, 2000, p. 16). Ademais é igualmente relevante no tocante aos processos cognitivos envolvidos no processamento lexical, tais como recuperação de informação e planejamento de enunciados (SOUZA; DUARTE; BERG, 2015, p. 274). Sunderman e Kroll (2006) realizaram um estudo com falantes nativos de língua inglesa com baixa e alta proficiência em espanhol como L2 em uma tarefa de reconhecimento de tradução, na qual o tempo e a acurácia envolvidos na rejeição de traduções de pares de palavra podem refletir processos lexicais subjacentes à arquitetura cognitiva na aquisição do léxico de uma língua. Assim, é possível que a dificuldade no acesso lexical se vincule ao construto proficiência em L2. Trefers-Daller (2010) mostrou que medidas de amplitude lexical podem ser índices válidos de habilidade linguística nos idiomas falados por bilíngues, possibilitando seu uso para operacionalizar e mensurar dominância linguística nessa mesma população.

O conhecimento lexical, então, não é percebido como único preditor de proficiência, uma vez que um sistema linguístico não é constituído apenas de léxico, mas como uma das medidas para se aferir proficiência linguística. Nesse âmbito, Read (2000, p. 8) pondera que, para se interpretar o desempenho de sujeitos a respeito do conhecimento lexical medido

por diferentes tipos de teste, é necessário compreender as dimensões envolvidas no acesso vocabular e refletidas nos testes aplicados.

Os instrumentos de acesso ao conhecimento lexical, segundo as dimensões de Read (2000, p. 9), podem ser discretos (quando o vocabulário é tomado como um construto isolado, separado de outros componentes da competência linguística) ou *embedded*¹² (tratado como parte do acesso a um construto mais amplo); seletivo (quando itens específicos de vocabulário são o foco do acesso) ou compreensivo (quando o foco é todo o vocabulário contido em uma dada tarefa de leitura/escuta ou de escrita/fala); independente ou dependente de contexto (quando se espera que o indivíduo considere ou não a informação contextual para produzir a resposta esperada). Como bem frisado por Read (2000, p. 10), para se determinar as dimensões de um teste de vocabulário, seu intuito e o modo de interpretação dos resultados precisam ser considerados, pois não é possível falar das dimensões do teste apenas através da maneira com a qual ele se mostra para quem o responde.

Além das dimensões, outro fator de importância é que a pontuação de desempenho em um teste pode ser indicador das habilidades linguísticas do sujeito e, para que isso aconteça, o teste deve ser confiável e válido (BACHMAN, 1990 p. 24). De acordo com a autora, a confiabilidade refere-se à qualidade da pontuação de um teste, livre de erros de mensuração, pois fatores, como cansaço e ansiedade, podem afetar o desempenho do indivíduo. Já para que um teste seja válido, a pontuação deve refletir significativamente a respectiva habilidade de interesse de estudo e não conter interferência de outras habilidades (BACHMAN, 1990 p. 25). A autora usa o exemplo de quando se pede que o aluno faça um texto baseado em uma palestra assistida por ele. Sua produção escrita não seria baseada apenas em sua habilidade escrita, mas na de escuta também, concluindo-se que essa produção textual pode não fornecer uma medida válida de habilidade escrita.

Um teste de vocabulário de suma importância é o *Vocabulary Levels Test* (VLT) de Nation (1990). Trata-se de um teste de vocabulário receptivo que busca medir a amplitude (tamanho) de vocabulário¹³, relacionada a diferentes bandas de frequência de uso, de falantes de língua inglesa tanto como L2 como L1, servindo como controle para os resultados obtidos em pesquisas. O VLT se configura em cinco níveis de frequência de palavra, partindo da banda

¹² Prefere-se usar o termo original por não se encontrar um termo adequado para esse contexto e que não seja redundante com sua definição.

¹³ Schmitt, Schmitt e Clapham (2001, p. 62) comentam que testes de vocabulário receptivo não podem informar sobre o uso produtivo do léxico [pronúncia ou ortografias empregadas corretamente, reconhecimento ou emprego de palavras derivadas etc.] pelos participantes.

de maior para a de menor frequência: 2k, 3k, 5k, nível de palavras usadas em contexto acadêmico e 10k respectivamente. Cada nível de frequência traz seis blocos com seis palavras (três palavras-alvo e três distratoras). A cada correspondência correta entre a palavra e a definição, soma-se um ponto, com pontuação máxima de 90 escores por acerto. A figura 2 mostra um exemplo de bloco para uma melhor visualização da configuração do VLT.

Figura 2 – Exemplo de bloco do *Vocabulary Levels Test* (NATION, 1990)

This is a vocabulary test. You must choose the right word to go with each meaning. Write the number of that word next to its meaning. Here is an example.

- | | |
|-------------|----------------------------------|
| 1. business | |
| 2. clock | _____ part of a house |
| 3. horse | _____ animal with four legs |
| 4. pencil | _____ something used for writing |
| 5. shoe | |
| 6. wall | |

Fonte: NATION, 1990, p. 264

O VLT continua sendo objeto de interesse de pesquisas psicométricas relacionadas a novas validações de testes, tais como os estudos de Schmitt, Schmitt e Clapham (2001), Souza, Duarte e Berg (2015), Silva (2016) para o inglês e o espanhol, além de Batista (2014) para o francês. Esse último estudo é de suma relevância para esta pesquisa, pois foi o primeiro teste formulado para mensuração de vocabulário receptivo para a língua francesa de que se tem conhecimento. A subseção seguinte aborda a descrição e a importância desse teste que se configura como um dos objetos de estudo desta pesquisa.

2.6 Teste de la Taille du Vocabulaire (TTV)

O TTV (*Teste de la Taille du Vocabulaire*), é um teste de amplitude vocabular desenvolvido e validado por Batista (2014) para falantes de língua francesa como L2 em contexto de imersão geográfica. Seus objetivos basearam-se em fornecer estimativas úteis do número de palavras aprendidas por estudantes universitários e do CÉGEP (*Collège*

d'Enseignement Général et Professionnel)¹⁴ e em avaliar efetividade do teste por meio de sua aplicação a grupos de aprendizes de francês como L2 em diferentes níveis de proficiência. Até então, do que é de nosso conhecimento, não havia sido proposta uma ferramenta de medição de vocabulário receptivo aos moldes do VLT (*Vocabulary Levels Test*), de Nation (1990), para a língua francesa, por isso a importância desse teste para o campo do ensino de línguas.

O TTV dispõe de quatro níveis, 2k, 3k, 5k e 10k, relacionados às bandas de maior para a menor frequência de palavras respectivamente. Diferente do VLT, que apresenta cinco níveis de frequência, sendo o nível acadêmico referente ao quarto nível, apresentam-se quatro níveis no TTV porque, na época em que foi proposto, segundo a autora, ainda não havia sido identificada nenhuma lista acadêmica de palavras (BATISTA, 2014, p. 38). A tarefa dos participantes foi a de associar uma palavra para cada definição, como apresentado na figura 3. No total, há 10 blocos por nível, 120 definições, 60 palavras-alvo e 60 palavras consideradas distratoras (ANEXO B – *TESTE DE LA TAILLE DU VOCABULAIRE (TTV) APLICADA AOS PARTICIPANTES*), sendo possível 120 pontos por nível.

Figura 3 – Exemplo de bloco do TTV

Dans ce test de vocabulaire, vous devez trouver le mot qui correspond à la définition. Mettez le bon numéro à côté de chaque définition choisie. Voici un exemple:

- | | |
|---------------|----------------------------|
| 1. accomplir | |
| 2. bloquer | |
| 3. dormir | _____ arrêter |
| 4. encourager | _____ aider quelqu'un |
| 5. mêler | _____ réaliser ou terminer |
| 6. ressentir | |

Fonte: BATISTA, 2014, p. 94

Para as frequências do TTV, cada uma com dez blocos, a autora segue as seguintes proporções: cinco substantivos, três verbos e dois adjetivos nas bandas 2k, 3k e 10k; cinco substantivos, dois verbos e três adjetivos em 5k. Assim, somando-se todas as frequências, o teste é constituído de vinte substantivos, onze verbos e nove adjetivos. As respostas para cada

¹⁴ O CÉGEP é uma instituição pública que oferece cursos pré-universitários e técnicos. Fazendo uma comparação com o ensino brasileiro, ele está situado entre o ensino médio e o superior. O caráter público dos CÉGEPs, ao invés de versar sobre o acesso gratuito da população, refere-se à submissão a uma lei aplicada aos estabelecimentos de ensino médio e profissionalizante. Isso significa que o ensino do CÉGEP não é gratuito, mas os quebequenses pagam uma tarifa menor que os estrangeiros.

nível se encontram nos anexos dessa dissertação (ANEXO C – RESPOSTAS DO *TESTE DE LA TAILLE DU VOCABULAIRE* (TTV)).

Participaram da pesquisa 175 indivíduos inscritos no programa de “francização”¹⁵ de um CÉGEP em Montréal, cujas L1 eram variadas: espanhol (38), persa (26), romeno (26), mandarim (23), russo (20), árabe (15), tagalo (9), português (3), ucraniano (2), vietnamita (2), língua amárica (1), bengali (1), berbere (1), búlgaro (1), inglês (1), húngaro (1), coreano (1), quirguiz (1), nepalês (1), tamil (1) e teochew (1). Eles concluíram o teste em menos de 35 minutos, sendo que o máximo de tempo permitido era de 50 minutos.

Os sujeitos foram divididos em dois grupos de proficiência: nível intermediário e nível avançado. Os estudantes de nível intermediário frequentavam 30 horas semanais de curso de francês, na época em que o estudo foi aplicado, tendo frequentado 700 horas de instrução, em média, até então. Os avançados, por sua vez, frequentavam 12 horas de curso por semana e já tinham obtido 1.200 horas de instrução formal em média.

De acordo com Batista (2014, p. 46), apesar de 62% dos indivíduos residirem no Canadá por um ano ou menos até a data de aplicação do teste, o tempo médio de permanência foi de um ano e nove meses, com variação de três meses a 13 anos. Ou seja, essa informação registrada pela pesquisadora sugere que a maioria dos aprendizes que participaram de sua pesquisa não tinham longo tempo de residência no Québec. Além desses dados, ela ainda informa que 46% relataram ter estudado francês em seus países de origem antes de partirem para o Canadá, e a mesma porcentagem de participantes indicou que não falavam inglês.

Apesar de ter sido aplicado na província do Québec, no Canadá, a pesquisadora se utilizou de dois *corpora* cujas listas de palavras não se restringem a serem frequentes somente em território canadense francófono. Primeiramente, para as frequências de 2k a 5k, foi usado um dicionário com as cinco mil palavras mais frequentes (LONSDALE; LE BRAS, 2009), baseado num *corpus* de 23 milhões de palavras escritas e faladas na França e em seus territórios além-mar. Para a frequência de 10k, foi usado o dicionário de Baudot (1992), configurando-se numa lista de 22 mil palavras advindas das imprensas francesa (62%) e canadense (37%), além de outros territórios (1%) em meio escrito. Batista (2014, p. 33) salienta que, apesar desse segundo *corpus* parecer desatualizado por ter sido baseado em textos coletados entre 1906 e 1967, demonstrou-se útil para a construção de seu TTV pela sua consistência em tamanho e por retratar as palavras de conteúdo geral e acadêmico utilizadas no Québec.

¹⁵ Isso significa que os estudantes, geralmente imigrantes, estão em período de aprendizado ou de aprimoramento do francês em uma instituição de ensino no Québec, Canadá.

Vale observar que, um pouco antes da publicação do TTV, já era perceptível o avanço dos estudos de *corpora*, especificamente os relacionados ao ambiente acadêmico em língua francesa, uma vez que os estudos sobre esse assunto em língua inglesa estavam mais desenvolvidos¹⁶. Existiam, na época da construção do teste, pelo menos dois *corpora* acadêmicos¹⁷ que poderiam servir de base para o nível 5 do TTV, entretanto não se sabe a razão pela qual a autora não se utilizou desses materiais para construir o nível acadêmico em seu teste de vocabulário.

Batista (2014) relatou uma certa vantagem dos sujeitos que tinham como L1 línguas neolatinas, algo também discutido em Milton (2009), em Schmitt (2010) e em Souza, Duarte e Berg (2015)¹⁸. Uma quantidade de 67 falantes de português, romeno e espanhol como L1 respectivamente, equivalente a 38% de sua população, obtiveram pontuação média mais alta que os falantes de persa, russo, tagalo e outras 11 línguas não latinas, constituindo 46% da população, e que os indivíduos cujas primeiras línguas advêm do Leste da Ásia (coreano, vietnamita, mandarim e teochew), correspondendo a 15%. Batista (2014, p. 72) comenta que esse melhor desempenho do grupo cujas primeiras línguas são neolatinas sugere que as palavras cognatas em suas respectivas línguas podem ter sido uma vantagem para esses participantes.

As palavras foram randomizadas para cada frequência em quatro listas, e as definições foram elaboradas pela pesquisadora a partir de um dicionário escrito e de dois *online*. Como a maioria das palavras eram polissêmicas, Batista (2014, p. 41) escolheu os sentidos de acordo com o significado mais frequente, baseando-se em Lonsdale e Le Bras (2009) para as frequências de 2k, 3k e 5k. Para a de 10k, ela escolheu o primeiro sentido registrado no dicionário *online Portail Lexical* (CNRTL, 2012) por não haver definições ou contextos nas palavras-alvo selecionadas a partir de Baudot (1992). As palavras aparecem em ordem alfabética, enquanto que as definições estão em ordem crescente de extensão, tal como feito por Schmitt, Schmitt e Clapham (2001).

A versão final do teste conta com um questionário sociobiográfico em francês o qual os participantes responderam antes de fazerem o TTV. Um dia após a aplicação do teste,

¹⁶ No início da década de 60, os estudos de *corpora* relacionados ao Inglês despontaram na *Brown University*, nos EUA, por meio de Francis e Kučera (1964) trabalhando no *corpus Brown*, assim como na *University College London*, Inglaterra, através do *Survey of English Usage* (QUIRK; SVARTVIK, 1966).

¹⁷ O *Chambers-Le Baron Corpus of Research Articles in French* (CHAMBERS; LE BARON, 2007) e o *Learning and Teaching Corpus* (LETEC) FAVI (*Français académique virtuel international*), organizado por Yun e Chanier (2011).

¹⁸ Nessa pesquisa, aplicou-se o VLT com brasileiros falantes de Inglês.

os sujeitos participaram de uma entrevista gravada na qual versaram sobre o modo como perceberam o teste e como procederam para respondê-lo.

No projeto piloto do TTV, houve relatos de tentativa de adivinhação quando os indivíduos desconheciam a resposta. Então o procedimento adotado pela pesquisadora foi o de sinalizar que deixassem a resposta em branco caso não soubessem a resposta, seguindo a linha de raciocínio de Schmitt, Schmitt e Clapham (2001). Os pesquisadores esclarecem que, se o participante sabe o significado de uma palavra-alvo, ele tende a responder de forma direta, não considerando as outras opções de itens. Em contrapartida, se ele não sabe o significado ou se tem dúvidas a respeito, ele tende a avaliar as palavras disponíveis (SCHMITT; SCHMITT; CLAPHAM, 2001, p. 61).

Na entrevista pós-teste, a partir do que foi relatado por doze participantes, a hipótese de que os estudantes realmente conheciam o significado das palavras respondidas corretamente no TTV foi confirmada (BATISTA, 2014, p. 63). Houve um melhor desempenho desses sujeitos nessa fase da pesquisa se comparada ao teste, pois, do contrário, poderia haver indícios de um número maior de tentativas de adivinhações no teste.

O *Teste de la Taille du Vocabulaire* parece demonstrar alguns pontos a serem revisitados e discutidos, como a ausência de ponto de corte para a classificação dos participantes em níveis de acordo com as respostas, assim como a falta de controle temporal mais restrito para as respostas dos participantes.

Como a autora deixa claro, o TTV foi elaborado para estimar o tamanho do vocabulário dos aprendizes de francês como L2 em até 10 mil palavras (BATISTA, 2014, p. 65), talvez por isso não seja mencionado o ponto de corte para a distribuição dos participantes através dos níveis. Já o ponto de corte sugerido por Nation (1990) para o VLT é de 13 pontos (pontuação máxima de 18 por nível). Por exemplo, para estar no nível 1, o participante deve acertar menos de 13 palavras desse nível, uma vez que é atribuído um ponto por acerto de palavra¹⁹; para estar no nível 2, deve acertar, no mínimo, 13 palavras do nível 1 e, assim, sucessivamente até o nível 5. Entretanto, Souza, Duarte e Berg (2015) verificaram que 12, 13 ou 14 pontos de corte por nível são igualmente discriminatórios para o VLT.

Outro ponto inovador da pesquisa de Souza, Duarte e Berg (2015) foi a inclusão de uma janela temporal para os participantes. Como usaram a versão do VLT de Nation (1990), o

¹⁹ É importante lembrar que o VLT de Nation (1990) possui 5 níveis, cada um composto por 6 blocos. Cada bloco possui 6 palavras (metade delas sendo distratoras) e 3 definições, totalizando uma pontuação máxima de 18 pontos por nível.

limite de tempo estabelecido foi o de 20 segundo para cada bloco contendo 6 palavras, totalizando 10 minutos para a realização do teste. De acordo com os autores, a facilidade de ativação e o acesso rápido às representações lexicais são importantes fatores da competência lexical em bilíngues, além de serem igualmente significantes em segunda língua por advirem da automaticidade (SOUZA; DUARTE; BERG, 2015, p. 279).

No TTV de Batista (2014), como explicado anteriormente, os participantes não dispuseram de um limite preciso de tempo para realizarem a tarefa. O teste original foi aplicado presencialmente e em papel, eles tinham, no máximo, 50 minutos para concluí-lo, porém o fizeram em menos de 35 minutos. Em uma situação real na qual o aprendiz necessita decodificar uma mensagem e respondê-la, interagindo com seu interlocutor, ele não dispõe, muitas vezes, de tempo suficiente para pensar nas regras morfológicas de acordo de gênero, por exemplo, na estrutura das sentenças, no emprego dos vocábulos, no contexto pragmático etc. e em uma eventual avaliação ou correção após esses eventos. Sabe-se que a proficiência é constituída de diversos componentes cognitivos e psicológicos, e um dos fatores que a caracterizam é a automaticidade das habilidades linguísticas em L2. Por essa razão, é um tópico de grande significância que merece a devida atenção na aplicação de testes de vocabulário ou outras ferramentas de verificação de habilidade linguística em L2.

Os resultados encontrados por Batista (2014) tendem a confirmar a validade de seu teste. Quanto maior o nível de proficiência dos sujeitos, maior foi a pontuação no TTV, indicando, conseqüentemente, maior tamanho de vocabulário. A pontuação também decresceu conforme os aprendizes de Francês passaram de nível, ou seja, a maior pontuação foi registrada no nível 2k e diminuiu gradualmente até o nível 10k. Além desses fatores, houve uma consistência entre o desempenho no teste e na entrevista sobre a comprovação do conhecimento das palavras respondidas corretamente pelas pessoas que participaram dessas duas fases da pesquisa.

É preciso salientar que a pesquisadora reconhece que mais informações a respeito da validação de seu TTV são necessárias. Todavia o surgimento dessa ferramenta é de suma relevância tanto para uso de professores de francês avaliarem a aquisição de vocabulário de seus alunos quanto para os estudos do bilinguismo no que concerne à proficiência, assim como para os estudos psicométricos que servem de ferramentas de apoio às pesquisas experimentais em Psicolinguística.

3 METODOLOGIA

A fim de verificar o desempenho de brasileiros falantes de francês como L2, a presente pesquisa se constituiu em duas etapas para dois grupos distintos de população: brasileiros e francófonos.

Para os falantes brasileiros, aplicou-se um questionário cujo objetivo era averiguar três parâmetros: autoavaliação de proficiência, tempo de estudo e frequência de uso da língua francesa no dia a dia. Em seguida, os participantes realizaram o TTV de forma *online*.

Em contrapartida, foi elaborado um questionário diferente para os nativos, pois o intuito foi o de atestar se o francês era L1 e se era usado em contexto de imersão geográfica. Por essa razão, foram considerados participantes que, apesar de não residirem em seu local de origem, estivessem em países francófonos. Após responderem o questionário, eles procederam com a resolução do TTV.

Uma descrição sobre a metodologia utilizada no questionário para os brasileiros, no questionário para os nativos francófonos e no TTV é reportada nas subseções seguintes. As análises e os resultados dos questionários e do TTV, bem como as discussões, encontram-se na seção subsequente.

3.1 Questionário aplicado para brasileiros

3.1.1 *Sujeitos*

Para viabilizar a participação nessa pesquisa, os voluntários deveriam se enquadrar no seguinte perfil: terem, no mínimo, dezoito anos; serem brasileiros que estudam/falam²⁰/usam a língua francesa, independente de terem habilidades linguísticas em outros idiomas além da língua-alvo de investigação desse estudo. Nessa primeira etapa da pesquisa, houve 179 respostas por meio do formulário eletrônico do Google. Para que a resposta ao questionário fosse considerada, os participantes deveriam prosseguir para a tarefa do TTV após o formulário, através do endereço eletrônico e do código de acesso que lhe foram fornecidos. Os dados dos participantes que responderam somente a esse formulário e não fizeram o TTV não foram analisados.

²⁰ No sentido de considerarem o francês como um idioma no qual possuem certa habilidade linguística, mesmo que mínima.

Seguiram para essa fase 145 voluntários, porém as informações de quatro deles foram descartadas por mencionarem que residiam em um país francófono no momento da pesquisa²¹ e de outros dois por terem declarado que não tinham conhecimento de francês, que não tiveram instrução formal nem contato com o idioma (o que não correspondia ao perfil de população a ser investigado).

Assim, foram considerados 139 participantes a partir do questionário. Todavia, como os dados de mais alguns sujeitos ainda não puderam ser analisados, a descrição precisa da amostra dessa pesquisa encontra-se na subseção 3.3.1 dedicada à descrição dos sujeitos brasileiros na aplicação do TTV.

3.1.2 Materiais

Primeiramente, foi apresentado aos participantes o termo de consentimento livre e esclarecido (ANEXO A - Termo de consentimento livre e esclarecido) aprovado pelo Comitê de Ética na Pesquisa (COEP) da Universidade Federal de Minas Gerais (UFMG), sendo respondido de forma *online* por meio do formulário do Google. Em seguida, aplicou-se o questionário *online* (APÊNDICE A – QUESTIONÁRIO APLICADO PARA BRASILEIROS), tendo por base o questionário de Valadares (2017). O intuito foi o de extrair informações sobre o tempo de estudo, a frequência de uso do idioma e a percepção dos aprendizes a respeito de seus próprios níveis de proficiência em cada uma das quatro habilidades em língua francesa (compreensão e expressão oral; compreensão e expressão escrita), ou seja, suas autoavaliações em francês como L2. Os dados de perfil dos participantes, sobre os quais Grosjean (1998) sinalou importância, são apresentados no APÊNDICE B – Resumo dos dados sobre o perfil dos bilíngues, com exceção dos dados de autodeclaração de proficiência, de tempo de estudo e de frequência de uso por serem tratados individualmente nas tabelas de 1 a 4.

²¹ Uma vez que o objetivo desse estudo é o de validar o *Teste de la Taille du Vocabulaire* com brasileiros aprendizes de francês não imersos geograficamente em países francófonos, não puderam ser consideradas as respostas de quem declarou vínculo com alguma instituição de ensino francófona (informação posteriormente confirmada após contato direto por e-mail entre pesquisadora e participantes).

3.2 Questionário aplicado para nativos

3.2.1 *Sujeitos*

Através de *e-mail* enviado a algumas universidades francófonas²² com o intuito de convidar o corpo discente à realização da pesquisa, 61 pessoas cuja L1 é o francês responderam o questionário (APÊNDICE C – QUESTIONÁRIO APLICADO PARA NATIVOS) via formulário do Google. Todavia, uma pessoa declarou não ter origem francófona, o que inviabilizou sua participação. A caracterização dessa população se encontra na seção 4.2.2, referente à descrição dos sujeitos que responderam o TTV.

3.2.2 *Materiais*

O objetivo desse questionário foi o de verificar se os voluntários tinham o francês como L1 e se eles residiam em um país francófono, mesmo que fosse diferente do território de origem, caracterizando a imersão geográfica. Isso se deve ao fato apontado por Souza, Duarte e Berg (2015, p. 270) a respeito da superação da crença de que as influências translinguísticas seriam direcionalizadas somente da L1 para os outros idiomas dos falantes. Os autores relatam o grande interesse de estudo de algumas teorias de representações linguísticas que investigam as “influências do conhecimento de línguas não maternas sobre as representações da língua materna e os mecanismos que regulam a interação de ambas classes de representação no processamento linguístico.” (SOUZA; DUARTE; BERG, 2015, p. 271). Então, devido a essa conjectura, nativos imersos em outros territórios que não fossem francófonos poderiam constituir um perfil de controle diferente.

Os voluntários tinham acesso ao *link* e ao código de acesso ao TTV após finalizarem suas respostas no questionário, da mesma forma que os participantes brasileiros.

²² Devido ao fato de possuírem acordo bilateral de mobilidade acadêmica com a UFMG, o que poderia facilitar a divulgação da pesquisa, as seguintes instituições de ensino foram contactadas: *Université de Laval* (Canadá), *Université de Liège* (Bélgica), *Université de Lille*, *Université de Poitiers*, *Université de Rouen* e *Université Sorbonne Nouvelle – Paris 3* (França).

3.3 Aplicação do *Teste de la Taille du Vocabulaire*

3.3.1 *Sujeitos*

Dos 139 voluntários brasileiros que responderam o questionário de autoavaliação e frequência (APÊNDICE A – QUESTIONÁRIO APLICADO PARA BRASILEIROS), os dados de seis pessoas foram descartados por metade delas ter respondido às definições com a palavra escrita ao invés de seu número correspondente no TTV²³, enquanto a outra metade não respondeu o teste após abri-lo.

Assim, 133 participantes (86 do sexo feminino) responderam o TTV, tendo idades entre 18 e 66 anos²⁴ (média: 26,87; desvio padrão: 7,86; mediana: 24; moda: 23), sendo a maioria estudantes universitários (ver APÊNDICE B – RESUMO DOS DADOS SOBRE O PERFIL DOS BILÍNGUES) e oriundos de vários estados brasileiros (Minas Gerais, Ceará, Amazonas, Mato Grosso, Santa Catarina, Paraíba e São Paulo). O leque dos possíveis graus de escolaridade considerados para o estudo foi amplo a fim de se conseguir diferentes níveis de competência linguística em francês, do básico ao avançado. Os alunos da Universidade Federal de Minas Gerais (UFMG) que registraram seus números de matrícula receberam certificado de participação na pesquisa para compor os créditos de atividades de cunho científico-cultural.

O TTV foi igualmente aplicado para nativos francófonos, e a relevância desse procedimento consiste em examinar se o teste está possível de ser resolvido, ou seja, se os itens contidos nele estão aceitáveis. No trabalho original, Batista (2014, p. 46) aplicou a versão piloto com apenas cinco francófonos. Para essa pesquisa, pretendemos aumentar o quanto fosse possível essa amostra. Todavia, dos 60 nativos que responderam o respectivo questionário, 22 não realizaram o TTV, dois responderam às definições com a palavra escrita ao invés de seu número correspondente no TTV e dois não o finalizaram. Portanto foram considerados 34 participantes nativos (31 do sexo feminino), com idade de 19 a 72 anos (média: 31,08; desvio padrão: 13,10; mediana: 27; moda: 22), todos universitários de licenciatura, bacharelado ou com pós-graduação.

²³ O modelo de resposta pode ser visto no ANEXO C – RESPOSTAS DO *TESTE DE LA TAILLE DU VOCABULAIRE* (TTV)

²⁴ Por ser uma tarefa de processamento *offline*, foram considerados falantes com, no mínimo, 18 anos e sem limite máximo de idade pré-estabelecido. Em tarefas cujo processamento é *online*, contudo, deve haver esse limite, uma vez que, as funções cognitivas do ser humano vão decaindo com o passar do tempo.

3.3.2 *Materiais*

Para essa pesquisa, foi usado o TTV desenvolvido por Batista (2014), entretanto algumas modificações foram realizadas.

Assim como aplicado em Souza, Duarte e Berg (2015) para o VLT, optou-se por incluir uma janela temporal para o TTV por assumir, tal qual explicado anteriormente, que a proficiência em uma língua está também relacionada com a automaticidade da ativação de representações lexicais e semânticas (SOUZA, DUARTE e BERG, 2015, p. 279). Foram seguidas as indicações desse mesmo estudo que optou por restringir vinte segundos para cada bloco de seis palavras. Uma vez que há, no TTV, dez blocos de seis palavras para cada banda de frequência (quatro ao total: 2k, 3k, 5k e 10k), o tempo total estipulado para a realização do teste foi de 14 minutos, incluindo a fase de treinamento exemplificada na figura 4.

Figura 4 - Fase de instrução inicial para o TTV

Dans ce test de vocabulaire, vous devez trouver le mot qui correspond à la définition. Mettez le bon numéro à côté de chaque définition choisie. Voici un exemple:

- | | |
|---------------|----------------------------|
| 1. accomplir | |
| 2. bloquer | |
| 3. dormir | _____ arrêter |
| 4. encourager | _____ aider quelqu'un |
| 5. mêler | _____ réaliser ou terminer |
| 6. ressentir | |

Vous devez inscrire votre réponse de la façon suivante:

- | | |
|---------------|-----------------------------------|
| 1. accomplir | |
| 2. bloquer | |
| 3. dormir | <u> 2 </u> arrêter |
| 4. encourager | <u> 4 </u> aider quelqu'un |
| 5. mêler | <u> 1 </u> réaliser ou terminer |
| 6. ressentir | |

Complétez toutes les sections du test. Si vous ne connaissez pas un mot, laissez la réponse en blanc.

Fonte: BATISTA, 2014, p. 94

Para incluir o componente temporal no TTV, foi necessário usar a plataforma *online* de criação e gerenciamento de testes *EasyTestMaker*, cuja licença é de responsabilidade do Laboratório de Psicolinguística da UFMG. Dessa forma, a viabilidade de aplicação do teste

ficou melhor, pois, ao término dos questionários respondidos pelo formulário do Google, os participantes recebiam o endereço e o código de acesso para essa plataforma a fim de realizarem o teste, estando cientes do limite de tempo da tarefa.

Além da inclusão da janela temporal e da viabilidade *online* do teste, ao invés de ser repassado em papel e na presença da pesquisadora, tal qual realizado no estudo original, outra diferença foi o estabelecimento do ponto de corte para os níveis do TTV.

Apesar de não haver menção de ponto de corte por níveis no estudo original de Batista (2014), os participantes deveriam acertar o mínimo de 23 palavras (77%) do total de 30 para a classificação no nível²⁵. No VLT, o ponto de corte mínimo proposto por Nation (1990, p. 262) foi de 12 acertos sobre o total de 18, ou seja, pouco mais de 30% das palavras de um dado nível poderiam não ser reconhecidas pelo aprendiz. Por outro lado, o ponto de corte empregado por Silva para o VLT (2016, p. 63) foi o de 14 (78%), sendo o ponto de corte mais conservador. Como foi provado por Souza, Duarte e Berg (2015, p. 288 e 289), a distribuição de classificação dos indivíduos sob o ponto de corte 15 é semelhante ao ponto de corte 13 para o VLT aplicado para brasileiros que tenham o inglês como L2, mesclando-se os níveis quatro e cinco em um só. Sendo assim, a partir das investigações desses três estudos, optou-se por estabelecer a proporção de acerto de 77% de palavras por nível para o TTV.

Caso houvesse dúvida em estabelecer a relação de correspondência entre a palavra e a definição, os participantes foram encorajados a deixarem a resposta em branco, tal qual seguido por Batista (2014) a partir de Schmitt, Schmitt e Clapham (2001).

²⁵ A estrutura permaneceu a mesma do TTV elaborado por Batista (2014): dez blocos por nível de frequência (totalizando 40 blocos); cada bloco com seis palavras (metade delas sendo distratoras) e três definições; um ponto para cada correspondência correta entre palavra e definição.

4 ANÁLISES E DISCUSSÕES

4.1 Questionários

4.1.1 *Dados do questionário aplicado para brasileiros*

As informações a serem tratadas nessa seção serão autoavaliação de proficiência, tempo de estudo e frequência de uso, consideradas como fatores externos a serem relacionados com o desempenho no teste de vocabulário.

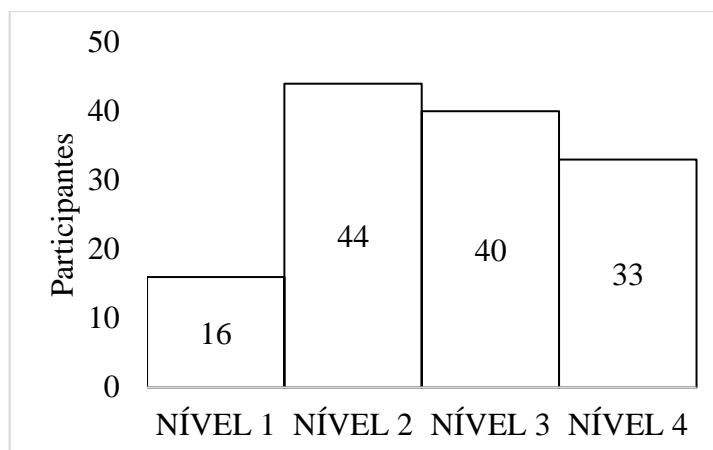
Para transformar os dados de autodeclaração de proficiência em categorias numéricas, foi atribuído um valor para cada categoria no programa Excel²⁶. Nas perguntas de 12 a 15 (APÊNDICE A – QUESTIONÁRIO APLICADO PARA BRASILEIROS), atribuiu-se valor 1 para a categoria “Principiante”, valor 2 para “Básica”, valor 3 para “Média”, valor 4 para “Avançada” e valor 5 para “Fluente”. Assim, reunindo as informações a respeito das quatro habilidades linguísticas (compreensão e expressão oral e escrita), o valor mínimo seria de 4, se o voluntário se considerasse principiante em todas as habilidades, enquanto o valor máximo seria 20, caso ele se considerasse fluente.

Em seguida, essas habilidades foram divididas em intervalos de valores²⁷, variando do nível de menor proficiência para a de maior proficiência: se a soma das habilidades estivesse entre um e cinco, seria nível 1, ou seja, menor proficiência; se estivesse entre seis e dez, nível 2; entre 11 e 15, nível 3; entre 16 e 20, nível 4, caracterizando-se como o nível de maior proficiência. O gráfico 1 mostra a quantidade de pessoas em distribuição nos níveis de 1 a 4, em ordem crescente de proficiência, de acordo com a autodeclaração de proficiência (média: 2,67; desvio padrão: 0,98): 16 participantes (12,03%) no nível 1, 44 (33,08%) no nível 2, 40 (30,08%) no nível 3 e 33 (24,81%) no nível 4 (N = 133).

²⁶ Todos os cálculos estatísticos e as atribuições de níveis reportados nas seções de análise relacionadas aos questionários e à performance no teste de vocabulário foram realizados através do Excel.

²⁷ Essa divisão foi útil para a constituição do histograma (GRÁFICO 1) e para a estatística inferencial realizada na seção 4.3 e na subseção 4.4.1.

Gráfico 1 - Histograma de níveis de autodeclaração de proficiência



Na análise descritiva da tabela 1, percebe-se que a maior média foi atribuída à habilidade de compreensão escrita, seguida da compreensão oral. Esse resultado pode refletir o fato do possível desenvolvimento das competências passivas de compreensão ocorrerem anteriormente às competências ativas de expressão através da quantidade de *input* ao qual o sujeito é submetido. Em contrapartida, a expressão escrita foi a habilidade na qual os participantes demonstraram menor domínio. Apesar do contexto de não imersão geográfica, a expressão oral foi uma das habilidades ativas com melhor percepção de fluência de acordo com a opinião dos participantes.

Tabela 1 – Estatística descritiva de autodeclaração de proficiência dos participantes nas quatro habilidades linguísticas

Habilidade	Valor mínimo	Valor máximo	Média	Desvio padrão	Mediana
Compreensão escrita	1	5	3,14	1,13	3
Compreensão oral	1	5	2,96	1,25	3
Produção escrita	1	5	2,58	1,17	3
Produção oral	1	5	2,73	1,21	3

Quanto ao fator tempo de estudo, as informações declaradas na questão dez (APÊNDICE A – QUESTIONÁRIO APLICADO PARA BRASILEIROS) estão descritas no gráfico 2 e na tabela 2. Os dados foram coletados em anos e transformados para meses, variando de 3 a 276 meses (23 anos) de estudo da língua francesa, seja por autodidatismo e/ou por

instrução formal em instituições de ensino público ou privado. A distribuição da quantidade de tempo de estudo em intervalos de 24,8 meses para cada classe é mostrada no gráfico 2, cuja informação aponta que 63 (47,36%) dos 133 voluntários estudam francês há pouco mais de um ano (moda: 12).

Gráfico 2 - Histograma de tempo de estudo

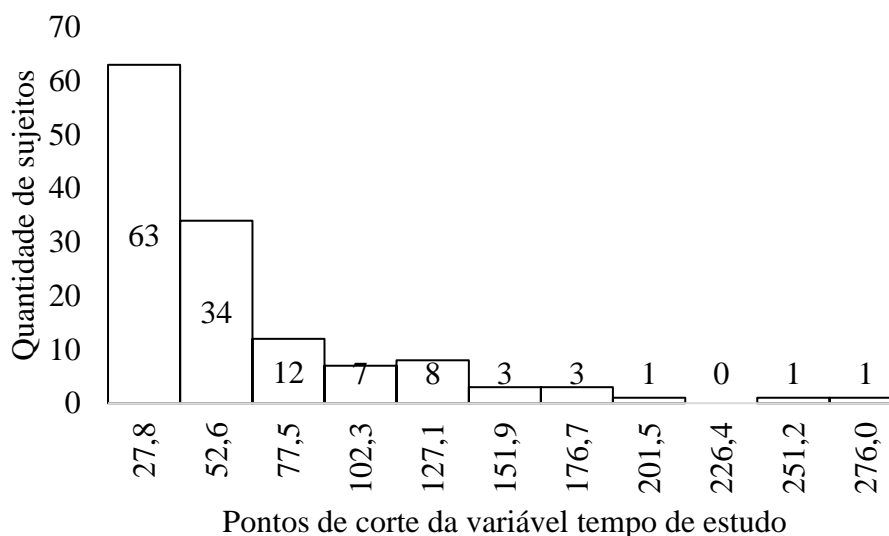


Tabela 2 – Estatística descritiva de tempo de estudo de francês

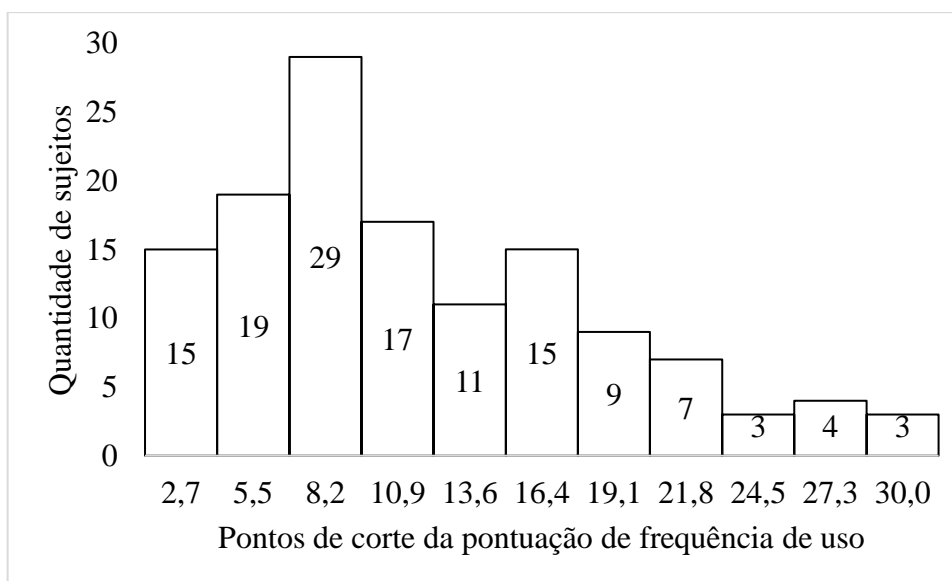
	Valor mínimo (em meses)	Valor máximo (em meses)	Média	Desvio padrão	Mediana
Tempo de estudo	3	276	46,80	48,18	30

Com relação à frequência de uso da língua francesa, foram atribuídos valores numéricos para as categorias adverbiais na questão 16 (APÊNDICE A – QUESTIONÁRIO APLICADO PARA BRASILEIROS): “Nunca” e “Raramente” receberam valor 0; “Ocasionalmente”, valor 1; “Frequentemente”, valor 2; “Sempre”, valor 3. O valor mínimo, então, seria de 0 se o sujeito assinalasse uma frequência de “Nunca” ou “Raramente”, e o valor máximo seria de 3 caso o sujeito marcasse “Sempre” para a frequência em uma determinada atividade.

Para essa pesquisa, foram verificadas 12 situações de frequência de uso de francês, ao invés das 16 situações estudadas por Valadares (2017) referentes à investigação de aquisição de vocabulário por meio do letramento digital em língua inglesa, considerando as que estão relacionadas às habilidades linguísticas presentes no aprendizado de um idioma, além da acessibilidade à língua francesa por meio de ferramentas (tecnológicas ou não) possivelmente presentes na realidade da maioria dos membros da população sobre a qual esse estudo pretende fazer inferências.

Como o valor mínimo para cada atividade é 0, e o valor máximo é 3, a pontuação mínima para todas as atividades é 0 se o sujeito marca “Nunca” ou “Raramente” em todos os itens, enquanto a máxima é 36 se ele responde “Sempre” para as 12 alternativas. O gráfico 3 mostra que 79,70% dos sujeitos obtiveram pontuação de frequência próxima a 16, enquanto 19,55% tiveram pontuação acima de 19 (média: 10,47; desvio padrão: 7,16; mediana: 9).

Gráfico 3 - Histograma de frequência de uso da língua francesa



A análise descritiva presente na tabela 3 demonstra que as maiores médias de hábitos de uso do francês são relacionadas ao contato com filmes, seriados e programas (atividade 6) e à busca por conteúdo escrito na internet (4) respectivamente, exceto quando se trata de leitura de jornais ou revistas (1). Quando se trata da visualização de filmes ou de séries, o uso da legenda em português (10) apresentou maior média, sendo seguido da legenda em francês (11) e da não utilização de legendas (12).

No tocante ao hábito da leitura, o motivo de lazer (3) obteve a maior média se comparado ao motivo acadêmico ou profissional (2). Entretanto a pesquisa *online* por conteúdo falado (5) aparenta ser menor que a de conteúdo escrito (4).

Em contrapartida, as atividades que envolvem o ato de escrever em diferentes tipos textuais - narração, descrição, dissertação etc. - em língua francesa (7), seja pela internet ou por outros meios (8), apresentam-se com médias mais baixas possivelmente por requererem habilidades linguísticas mais custosas aos indivíduos.

A média relativamente baixa para o treino da oralidade com alguém que fale o francês em uma escala de intermediário a fluente (9) pode advir do fato da prática oral ser restrita à sala de aula ou à pouca interação com nativos francófonos, apesar dessa barreira ser cada vez mais possível de ser ultrapassada graças à tecnologia e à internet.

Por fim, para o exercício de assistir a filmes e séries (de 10 a 12), o uso da legenda em português apresentou maior média.

Tabela 3 – Estatística descritiva de frequência de uso de francês por atividade

Atividades	Valor mínimo	Valor máximo	Média	Desvio padrão	Mediana
1. Leio artigos de jornais e revistas (impresso ou online)	0	3	0,83	0,89	1
2. Leio materiais para fins profissionais (trabalho) e/ou acadêmicos (estudos)	0	3	0,94	1,03	1
3. Leio por prazer e como atividade de lazer	0	3	1,06	0,87	1
4. Pesquiso por informações na internet cujo conteúdo é escrito	0	3	1,19	0,91	1
5. Pesquiso por informações na internet cujo conteúdo é falado	0	3	0,92	0,95	1

6. Procuo assistir a filmes, a seriados e a outros programas	0	3	1,35	0,94	1
7. Escrevo qualquer tipo de texto	0	3	0,54	0,75	0
8. Troco mensagens pela Internet ou em outros meios	0	3	0,60	0,90	0
9. Treino a oralidade com quem fala a língua em nível intermediário, avançado ou fluente	0	3	0,75	0,95	0
10. Assisto a programas / filmes / seriados falados em francês com legendas em português	0	3	1,00	0,90	1
11. Assisto a programas / filmes / seriados falados em francês com legenda em francês	0	3	0,74	0,89	0
12. Assisto a programas / filmes / seriados falados em francês sem legenda	0	3	0,56	0,87	0

Para facilitar a interpretação desses dados, as atividades supracitadas foram organizadas de acordo com as quatro competências requeridas no aprendizado de uma língua: compreensão oral e escrita; produção oral e escrita.

As atividades de 1 a 4 (“Leio artigos de jornais e revistas”, “Leio materiais para fins profissionais (trabalho) e/ou acadêmicos”, “Leio por prazer e como atividade de lazer” e “Pesquisa por informações na internet cujo conteúdo é escrito”) foram relacionadas com a habilidade de compreensão escrita. As atividades 5, 6, 10, 11 e 12 (“Pesquisa por informações na internet cujo conteúdo é falado”, “Procuo assistir a filmes, a seriados e a outros programas”, “Assisto a programas / filmes / seriados falados em francês com legendas em português”,

“Assisto a programas / filmes / seriados falados em francês com legenda em francês”²⁸, “Assisto a programas / filmes / seriados falados em francês sem legenda”) relacionam-se à compreensão oral. Já as atividades 7 e 8 (“Escrevo qualquer tipo de texto” e “Troco mensagens pela Internet ou em outros meios”) compreendem a produção escrita, enquanto a 9 (“Treino a oralidade com quem fala a língua em nível intermediário, avançado ou fluente”) envolve a produção oral.

A partir dessas associações, as categorias adverbiais foram somadas de acordo com a atribuição de valores explicitada anteriormente. A análise descritiva da tabela 4 mostra que as atividades relacionadas à compreensão oral e à compreensão escrita apresentam as maiores médias respectivamente, ou seja, os sujeitos parecem praticar a escuta com mais frequência que a leitura. No entanto, a atividade de menor frequência de uso é a de produção oral de acordo com o detalhamento das atividades e as respectivas discussões realizadas nos parágrafos anteriores.

Tabela 4 – Estatística descritiva de frequência de uso relacionadas às habilidades linguísticas

Atividades	Valor mínimo	Valor máximo	Média	Desvio padrão	Mediana
Atividades de compreensão escrita	0	9	4,01	2,90	4
Atividades de compreensão oral	0	15	4,56	3,19	4
Atividades de produção escrita	0	6	1,14	1,45	0
Atividades de produção oral	0	3	0,75	0,95	0

²⁸ Para relacionar essas atividades com a habilidade de compreensão oral, optou-se por assumir que o *input* principal é a escuta dos programas, filmes e seriados, ao invés da compreensão escrita das legendas em francês.

4.1.2 Dados do questionário aplicado para nativos

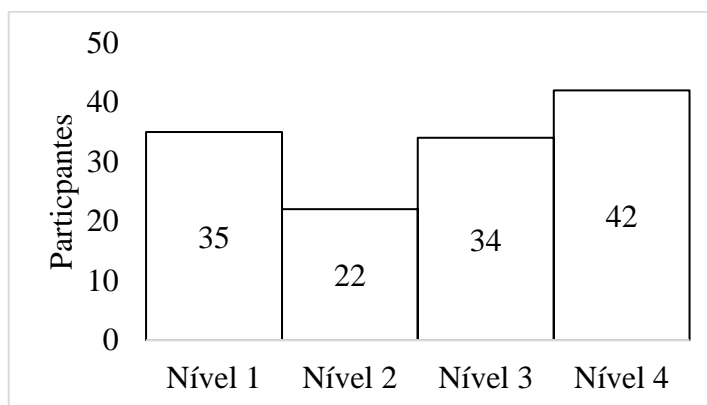
As informações relevantes sobre esse questionário são abordadas na subseção 4.2.2, pois seu objetivo foi o de verificar se o participante francófono estaria em imersão geográfica que possibilitasse o uso frequente da língua francesa.

4.2 Desempenho no TTV

4.2.1 Brasileiros

Com ponto de corte de 23 pontos, a classificação dos brasileiros por nível no TTV é mostrada no gráfico 4. Dos 133 participantes, 35 (26,32%) foram classificados no nível 1; 22 (16,54%) no nível 2; 34 (25,56%) no nível 3 e 42 (31,58%) no nível 4. Percebe-se então que a maioria dos participantes foram classificados no último nível, correspondente a maior proficiência, seguido pelo nível 1, de menor proficiência.

Gráfico 4 - Classificação dos participantes brasileiros no TTV por nível (N = 133)



A fim de saber se o teste captura de maneira semelhante o total esperado de pessoas classificadas por nível, verificando se há discrepância na distribuição da figura 5, foi aplicado o qui-quadrado de *Goodness of Fit*²⁹, baseando-se na expectativa de que as habilidades linguísticas dos aprendizes se distribuem de maneira razoavelmente uniforme dentro da distribuição da amostra. Assim, não foi encontrada significância estatística (qui-quadrado:

²⁹ Disponível em <https://www.socscistatistics.com/tests/goodnessoffit/Default2.aspx>

6,218; $p > 0,05$), ou seja, é possível que o teste demonstre uma distribuição na qual os sujeitos estão ajustados de forma equilibrada ao longo dos quatro níveis de classificação.

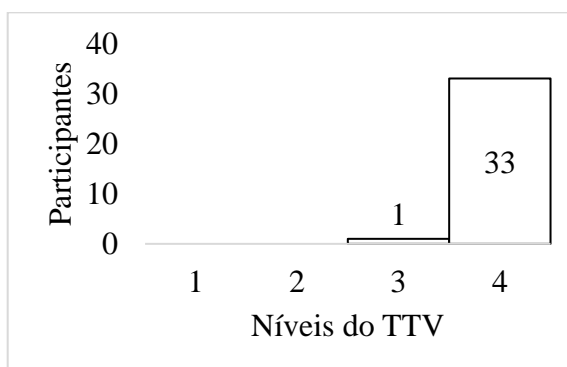
Figura 5 - Qui-quadrado de *Goodness of Fit* para verificar a distribuição de classificação dos sujeitos entre os níveis

	<i>Observed</i>	<i>Expected</i>	<i>Difference</i>	<i>Difference Sq.</i>	<i>Diff. Sq. / Exp Fr.</i>
Nível 1	35	0.25 (33.25)	1.75	3.06	0.09
Nível 2	22	0.25 (33.25)	-11.25	126.56	3.81
Nível 3	34	0.25 (33.25)	0.75	0.56	0.02
Nível 4	42	0.25 (33.25)	8.75	76.56	2.30
					6.218

4.2.2 Nativos

Seguindo o mesmo ponto de corte para os brasileiros, 23 acertos por nível no TTV, de todos os nativos que responderam o teste ($N = 34$), 33 foram classificados no nível 4, enquanto apenas um foi classificado no nível 3, como mostrado no gráfico 5 (média: 3,97; desvio padrão: 0,17). O total de escore do grupo variou de 75 à pontuação máxima de 120 (média: 113,14; desvio padrão: 11,23; mediana: 118). Uma vez que o objetivo de aplicar o teste para esse grupo era o de averiguar se os itens contidos nele estavam aceitáveis, o desempenho dessa categoria foi satisfatório para esse quesito, embora fosse esperado que houvesse mais resultados próximos ao escore máximo de 120.

Gráfico 5 - Classificação dos participantes nativos no TTV por nível ($N = 34$)



4.3 A relação das variáveis tempo de estudo e frequência de uso com autodeclaração de proficiência

Relembrando o que foi discutido na discussão teórica, é necessário ter cautela ao tomar testes de autoavaliação de proficiência como parâmetro para medir a competência linguística do falante em um certo idioma. Com o intuito de verificar se tempo de estudo e frequência de uso do francês estão correlacionados com a autodeclaração de proficiência e se aquelas variáveis conseguem explicar a variabilidade dessa, utilizou-se a regressão múltipla para tentar responder essa questão. Para tanto, a classificação total de autodeclaração (com valor mínimo de 4 e máximo de 20 na soma das competências de leitura, escuta, escrita e fala) foi requerida em detrimento da estratificação em níveis.

De antemão, as análises de correlação presentes nesse estudo seguiram o critério de interpretação do coeficiente de correlação de Pearson (r) presente em Vieira (2013), apresentado na tabela 5.

Tabela 5 – Interpretação do valor do coeficiente de correlação de Pearson (r)

Valores de r	Tipo de correlação
$0 < r < 0,25$ ou $-0,25 < r < 0$	Correlação pequena ou nula
$0,25 < r < 0,50$ ou $-0,50 < r < -0,25$	Correlação fraca
$0,50 < r < 0,75$ ou $-0,75 < r < -0,50$	Correlação moderada
$0,75 < r < 1,00$ ou $-1 < r < -0,75$	Correlação forte
$r = -1$ ou $r = 1$	Correlação perfeita

Fonte: Vieira, 2013, p. 83

Os gráficos 6 e 7 mostram os diagramas de dispersão de cada variável independente em relação à variável dependente autodeclaração de proficiência, juntamente com os dados da regressão na tabela 6.

Gráfico 6 - Diagrama de dispersão entre tempo de estudo e autodeclaração de proficiência

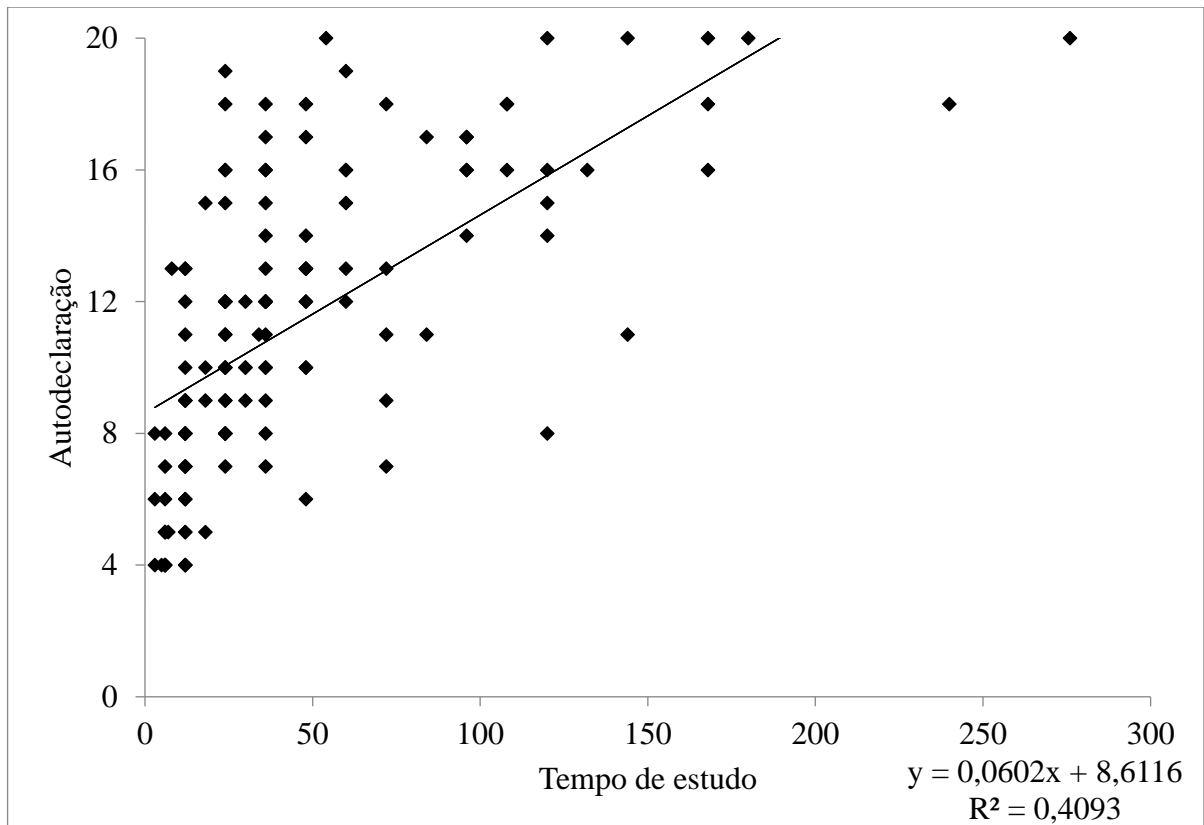


Gráfico 7 - Diagrama de dispersão entre frequência de uso e autodeclaração de proficiência

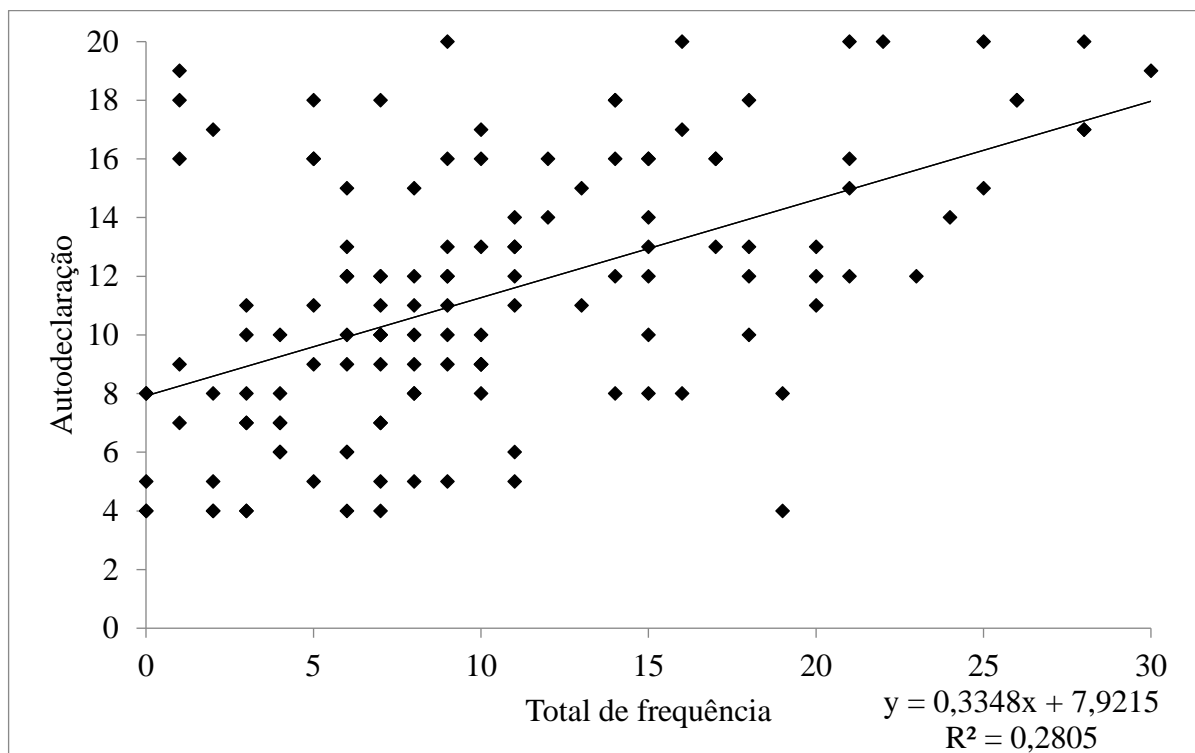


Tabela 6 – Análise de regressão linear múltipla das variáveis independentes tempo de estudo e frequência de uso com autodeclaração de proficiência

Estatística de regressão	
R múltiplo	0,710
R-Quadrado	0,505
R-quadrado ajustado	0,497
Erro padrão	3,212
Observações	133

	gl	SQ	MQ	F	F de significação
Regressão	2	1370,91	685,46	66,42	0,000
Resíduo	130	1341,66	10,32
Total	132	2712,57

	Coefficientes	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores
Interseção	6,95	0,51	13,63	0,000	5,943	7,962
Tempo de estudo	0,05	0,01	7,69	0,000	0,036	0,061
Frequência de uso	0,21	0,04	5,02	0,000	0,128	0,295

A partir dos gráficos 6 e 7, além da tabela 6, parecem haver evidências significativas ($p < 0,05$) de uma correlação moderada ($r = 0,71$) entre o tempo de estudo e a frequência de uso desse idioma com a autodeclaração de proficiência, além de ser provável que aquelas duas variáveis independentes possam explicar 50% da variabilidade da autodeclaração ($r^2 = 0,50$).

Dessa forma, percebe-se que outros fatores, além dessas variáveis investigadas, podem estar envolvidos na autoavaliação dos participantes. Através do intervalo de dados observados até então, tempo de estudo e frequência de uso podem explicar metade da autodeclaração. Apesar de sua confiabilidade como instrumento de medição de competência linguística ser colocada frequentemente em questão, o intuito é de fazer uma relação entre essa medida e o teste de vocabulário para que seja uma contribuição no campo de estudo dos testes psicométricos, principalmente para a área do francês.

Após a verificação de como a medida de autodeclaração de proficiência se correlaciona com as outras medidas externas mencionadas pelos sujeitos da pesquisa, o

interesse das próximas seções é investigar se há relação desses fatores com o desempenho dos participantes no teste de vocabulário.

4.4 O papel das variáveis autodeclaração de proficiência, tempo de estudo e frequência de uso na variabilidade da pontuação do TTV

Nessa etapa da análise, buscaram-se respostas para as seguintes perguntas: há algum tipo de relação entre a autodeclaração de proficiência, o tempo de estudo e a frequência de uso com o desempenho no *Teste de la Taille du Vocabulaire*? Esses três fatores (autodeclaração de proficiência, o tempo de estudo e a frequência de uso) podem ser preditores de amplitude vocabular (um dos elementos associados ao construto multifacetado proficiência)? Se positivo, quais seriam os tipos de relação e as intensidades dessas relações?

Para responder essas questões, foi necessário levar em consideração a pontuação total do TTV, eliminando as classificações por níveis³⁰. Isso significa que a classificação por nível não se torna interessante para esse propósito, por ser uma variável categórica ordinal. Para verificar a correlação, é necessário que as variáveis investigadas sejam quantitativas. Assim a estratificação de níveis foi desconsiderada, lançando-se mão da pontuação total resultante do desempenho dos sujeitos³¹ para transformar a variável dependente categórica ordinal em quantitativa discreta. Com relação às variáveis independentes, as categóricas ordinais (autodeclaração de proficiência e frequência de uso) foram igualmente transformadas em quantitativas discretas por meio da atribuição de valores numéricos, como explicado no início desse capítulo, para que os resultados fossem melhor interpretados. Já a variável tempo de estudo não precisou ser modificada por ser de natureza quantitativa discreta, pois o tempo em anos foi transformado em meses.

Nas subseções a seguir, será investigado se as variáveis independentes estão relacionadas com a variável dependente de desempenho no TTV a partir dos diagramas individuais de dispersão, utilizando-se os dados dos 133 participantes apresentados

³⁰ Para verificar em que nível (de um a quatro) o participante se encontrava, o ponto de corte foi de 23, ou seja, ele teria que acertar, no mínimo, 23 correspondências por nível entre palavras e suas definições. Para isso, o escore total não era importante, pois bastava-se saber a pontuação mínima.

³¹ É possível pensar que um participante poderia ter uma pontuação relativamente alta em relação a sua classificação por nível. Entretanto não se deve esquecer o componente temporal adicionado a essa aplicação do TTV. Isso implica dizer que é provável que um sujeito com baixa proficiência não consiga chegar aos níveis mais altos de frequência pelo tempo que ele deva levar ao tentar responder todos os itens, mesmo os que ele tente acertar por tentativa de adivinhação.

individualmente na tabela 11 (APÊNDICE D – DADOS DAS VARIÁVEIS DE ANÁLISE DOS BRASILEIROS). Após os diagramas abaixo, as regressões lineares serão investigadas.

4.4.1 *Análise individual da relação entre as variáveis autodeclaração de proficiência, tempo de estudo e frequência de uso com o desempenho no TTV*

Para verificar se autodeclaração de proficiência, tempo de estudo e frequência de uso do francês podem ser preditores de proficiência (baseada no escore do TTV), foram feitos três diagramas de dispersão e as análises de regressão linear de cada uma das variáveis independentes, tendo como objetivo investigar se elas podem responder pela variabilidade na pontuação do TTV.

Através do gráfico 8, que analisa a relação entre autodeclaração de proficiência e escore no TTV, apesar de haver uma certa dispersão entre as variáveis, parece existir uma correlação positiva entre elas, ou seja, quanto maior a proficiência autodeclarada pelo participante, maior seu desempenho no TTV. Entretanto é possível verificar igualmente sujeitos que relataram um baixo nível de proficiência e obtiveram um bom desempenho no TTV e vice-versa, podendo caracterizar uma correlação positiva fraca.

Gráfico 8 – Diagrama de dispersão entre autodeclaração de proficiência e escore no TTV

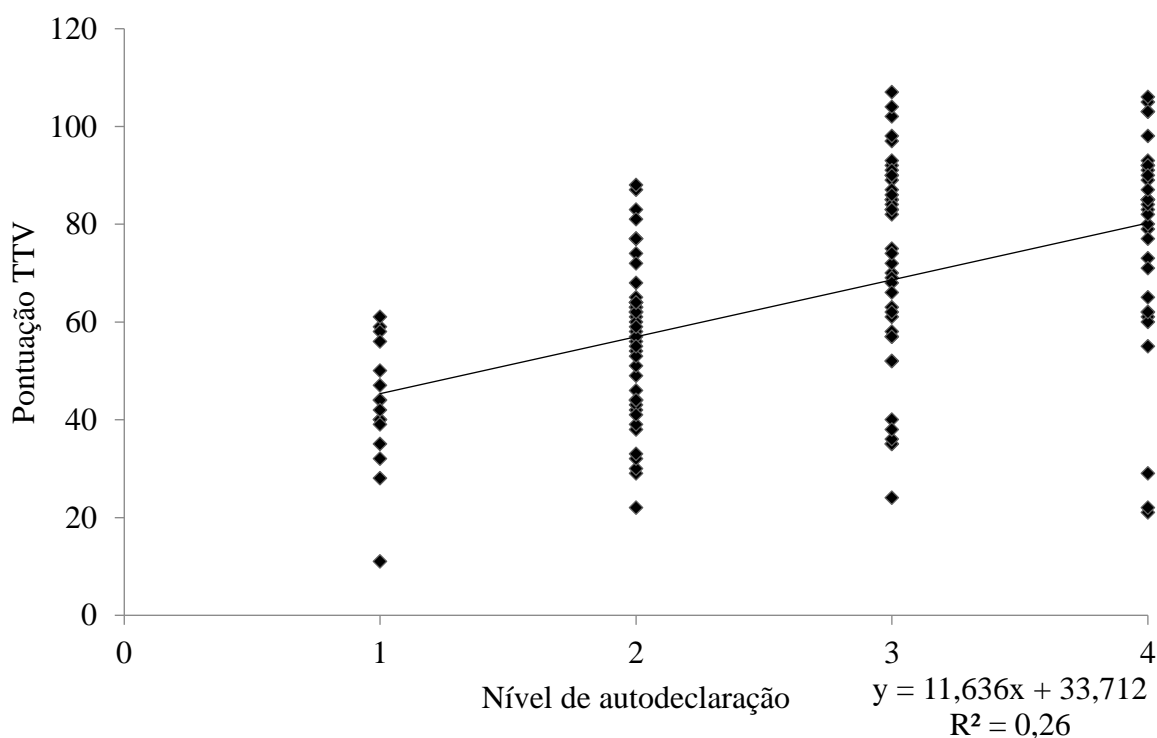


Tabela 7 – Análise de regressão linear das variáveis autodeclaração de proficiência e TTV

Estatística de regressão	
R múltiplo	0,512
R-Quadrado	0,262
R-quadrado ajustado	0,257
Erro padrão	19,187
Observações	133

	gl	SQ	MQ	F	F de significação
Regressão	1	17207,37	17207,37	46,74	0,000
Resíduo	131	48228,91	368,16
Total	132	65436,29

	Coefficientes	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores
Interseção	33,71	4,85	6,95	0,000	24,12	43,31
Autodecl	11,636	1,70	6,84	0,000	8,269	15,002

A partir do exposto na tabela 7, parece haver evidência significativa ($p < 0,05$) de uma correlação moderada ($r = 0,51$) entre a autodeclaração de proficiência e o desempenho no TTV, além de ser provável que 26% da variabilidade da pontuação no TTV ($r^2 = 0,26$) possa ser explicada por meio dessa variável independente.

Para a verificação do papel da segunda variável independente, o gráfico 9 mostra a correlação positiva entre as variáveis tempo de estudo e pontuação geral no TTV. Em comparação ao gráfico anterior, parece haver uma correlação de tempo de estudo e TTV mais fraca que a de autodeclaração de proficiência e TTV por conta da maior dispersão dos participantes.

Gráfico 9 – Diagrama de dispersão entre tempo de estudo e escore no TTV

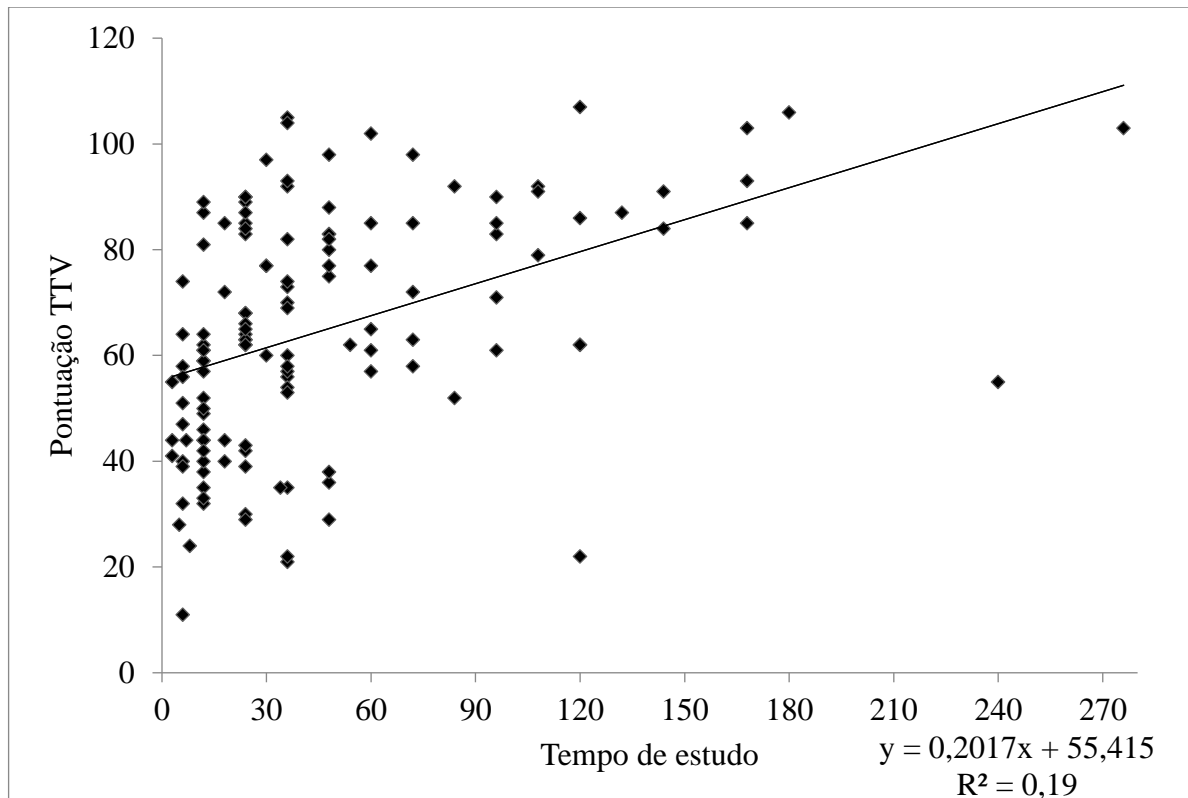


Tabela 8 – Análise de regressão linear das variáveis tempo de estudo e TTV

Estatística de regressão	
R múltiplo	0,436
R-Quadrado	0,190
R-quadrado ajustado	0,184
Erro padrão	20,107
Observações	133

	gl	SQ	MQ	F	F de significação
Regressão	1	12473,88	12473,88	30,85	0,000
Resíduo	131	52962,41	404,29
Total	132	65436,29

	Coefficientes	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores
Interseção	55,415	2,435	22,758	0,000	50,60	60,23
Tempo de estudo	0,2017	0,036	5,555	0,000	0,13	0,27

A partir da tabela 8, parece haver evidência significativa ($p < 0,05$) de uma correlação fraca ($r = 0,43$) entre o tempo de estudo de língua francesa e o desempenho no TTV, além de ser provável que 19% da variabilidade da pontuação no TTV ($r^2 = 0,19$) possa ser explicada por meio dessa variável independente.

Após as análises das variáveis autodeclaração de proficiência e tempo de estudo, a variável a ser investigada a seguir é a frequência de uso do francês. O gráfico 10 mostra que muito provavelmente haja pouca ou nenhuma influência dessa variável independente sobre a pontuação total do teste devido à grande dispersão dos dados.

Gráfico 10 – Diagrama de dispersão entre frequência de uso do francês e escore no TTV

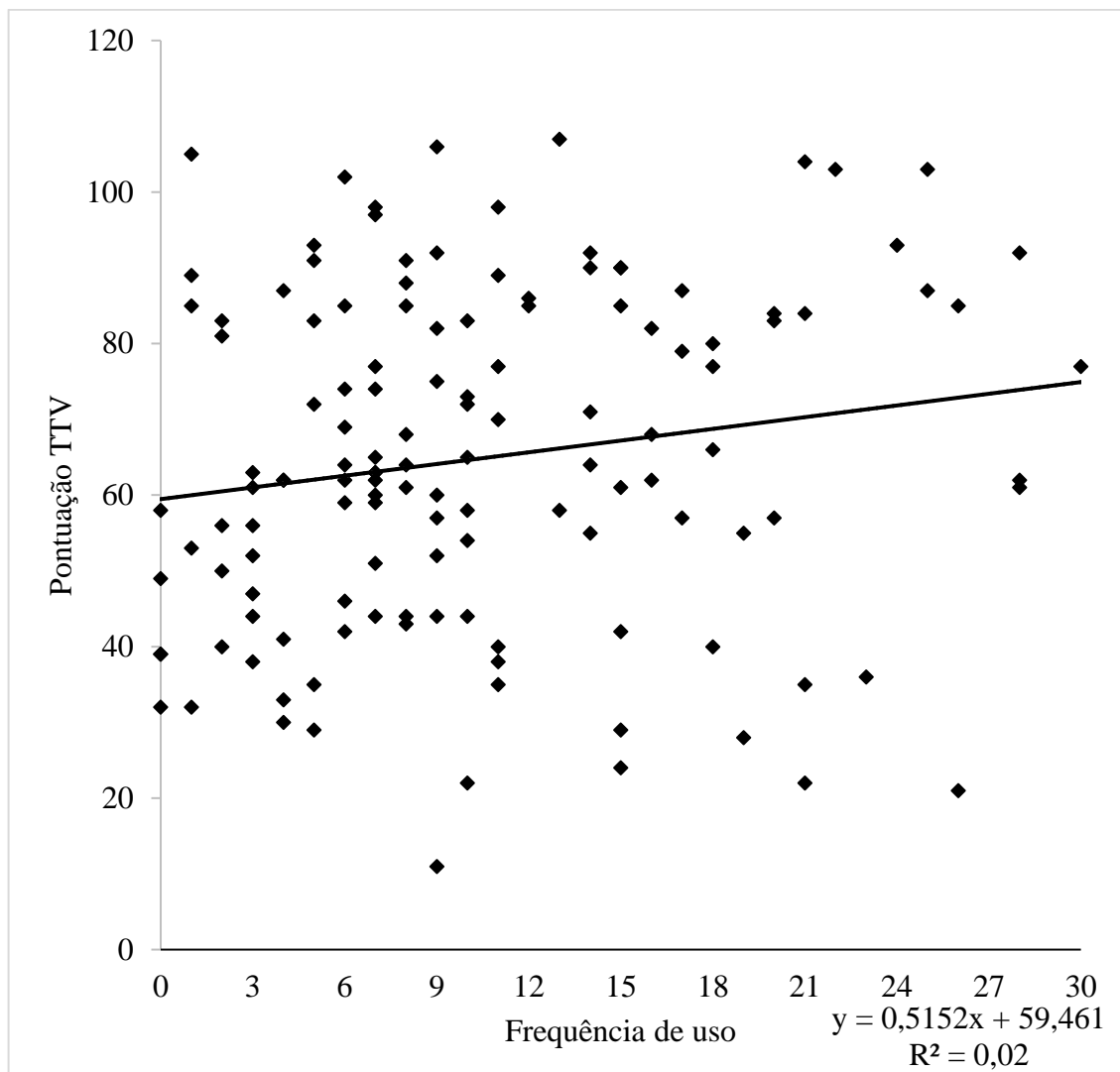


Tabela 9 – Análise de regressão linear das variáveis tempo de estudo e TTV

Estatística de regressão	
R múltiplo	0,165
R-Quadrado	0,027
R-quadrado ajustado	0,020
Erro padrão	22,040
Observações	133

	gl	SQ	MQ	F	F de significação
Regressão	1	1801,29	1801,29	3,71	0,056
Resíduo	131	63635,00	485,76
Total	132	65436,29

	Coefficientes	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores
Interseção	59,46	3,39	17,53	0,000	52,75	66,17
Frequência	0,52	0,27	1,93	0,056	-0,014	1,044

Os dados da tabela 9 sugerem evidências insuficientes ($p > 0,05$), apesar de muito próximas, de uma correlação entre frequência de uso de língua francesa e o desempenho no TTV. Mesmo que houvesse evidências significativas, essa variável responderia por apenas 2% da variabilidade da pontuação no TTV ($r^2 = 0,02$).

Por meio das análises acima, percebe-se que, das três variáveis independentes, autodeclaração de proficiência, tempo de estudo e frequência de uso do idioma, somente as duas primeiras demonstram possibilidade de relação com a pontuação no TTV. Porém autodeclaração de proficiência e tempo de estudo apresentam coeficientes de determinação (r^2) baixos (0,26 e 0,19 respectivamente), não conseguindo explicar grande parte da variabilidade nos escores do TTV. Assim, com o intuito de analisar melhor essa questão, utilizou-se a regressão linear múltipla.

4.5 Autodeclaração de proficiência e tempo de estudo relacionadas à variabilidade da pontuação geral do TTV

Com base nas análises anteriores, a variável preditora que melhor explicaria a variabilidade da pontuação do TTV seria a autodeclaração de proficiência, com 26% ($r^2 = 0,26$), seguida do tempo de estudo com 19% ($r^2 = 0,19$). Contudo as variáveis não conseguiriam explicar individualmente grande parte da variabilidade do desempenho total no TTV. Dessa forma, a regressão linear múltipla pode fornecer indícios do quanto essas duas variáveis independentes podem esclarecer juntas a pontuação no TTV.

Tabela 10 – Análise de regressão linear múltipla das variáveis autodeclaração de proficiência e tempo de estudo na pontuação do TTV

Estatística de regressão	
R múltiplo	0,537
R-Quadrado	0,289
R-quadrado ajustado	0,278
Erro padrão	18,923
Observações	133

	gl	SQ	MQ	F	F de significação
Regressão	2	18884,41	9442,21	26,37	0,000
Resíduo	130	46551,87	358,09
Total	132	65436,29

	Coefficientes	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores
Interseção	36,710	4,980	7,372	0,000	26,858	46,561
Nível de autodecl	8,896	2,102	4,231	0,000	4,736	13,055
Tempo de estudo	0,093	0,043	2,164	0,032	0,008	0,177

Com os dados da tabela 10, parece haver evidências estatísticas ($p < 0,05$) de que o modelo preveja que as variáveis autodeclaração de proficiência ($p > 0,05$) e tempo de estudo ($p < 0,05$) se relacionam com a pontuação global do TTV. Entretanto, juntas, essas variáveis podem explicar somente 29% ($r^2 = 0,289$) dessa variabilidade no escore do teste de vocabulário

aplicado, demonstrando uma correlação moderada ($r = 0,54$). Além disso, não há evidências suficientes de que a frequência de uso seja uma variável preditora da pontuação do teste ($p > 0,05$). Sendo assim, torna-se necessária a investigação futura de outras variáveis que possam estar relacionadas com o desempenho no teste de vocabulário, pois as previsões discutidas acima são válidas somente para o intervalo de dados observados.

4.6 Comparação entre classificação no teste e autodeclaração de proficiência

Essa subseção se dedica a responder se o TTV classifica os sujeitos de acordo com a autodeclaração de proficiência. Assim foi considerada a classificação dos participantes por nível³², ao invés da pontuação total obtida no teste, devido à possibilidade de se trabalhar com duas variáveis categóricas ordinais baseadas em níveis. Para tentar responder essa pergunta, foi realizado o teste de qui-quadrado (figura 6).

Figura 6 - Qui-quadrado de comparação entre desempenho no TTV e autodeclaração de proficiência

Results						
	Autodeclaração	TTV				Row Totals
Nível 1	16 (25.50) [3.54]	35 (25.50) [3.54]				51
Nível 2	44 (33.00) [3.67]	22 (33.00) [3.67]				66
Nível 3	40 (37.00) [0.24]	34 (37.00) [0.24]				74
Nível 4	33 (37.50) [0.54]	42 (37.50) [0.54]				75
<i>Column Totals</i>	133	133				266 (Grand Total)

A figura 6 mostra o número de participantes (133) nas duas variáveis e o número total de observações (266), bem como os valores esperados de distribuição dos sujeitos por cada nível entre parênteses e as estatísticas de qui-quadrado entre colchetes para cada célula. Pelos dados, parece haver uma diferença significativa de distribuição de classificação entre os níveis de autodeclaração de proficiência e o teste de vocabulário (qui-quadrado = 15,978; $p < 0,05$; $gl = 3$), pois a hipótese nula de que as distribuições seriam similares parece ter sido rejeitada.

Na tentativa de averiguar uma explicação para a distribuição não similar entre os níveis, foram feitas análises *post hoc* de qui-quadrado com a união dos níveis iniciais e a separação dos avançados separados e vice-versa. Os dados da figura 7 (qui-quadrado = 2,380,

³² Para se classificar em um certo nível, o participante teria de acertar 23 das 30 palavras. A distribuição de classificação por nível se encontra no gráfico 4.

$p > 0,05$) parecem confirmar a hipótese nula de não haveria diferença de distribuição nos níveis mais avançados, 3 e 4.

Figura 7 - Qui-quadrado de união dos níveis iniciais e separação dos níveis mais avançados

Results						
	Autodeclaração	TTV				Row Totals
Níveis 1 e 2	60 (59.37) [0.01]	57 (57.83) [0.01]				117
Nível 3	44 (39.58) [0.49]	34 (38.42) [0.51]				78
Nível 4	33 (38.06) [0.67]	42 (38.94) [0.69]				75
Column Totals	137	133				270 (Grand Total)

Por outro lado, separando-se os níveis iniciais e unindo os níveis avançados, parece haver diferença significativa (qui-quadrado = 14,472; $p < 0,05$) entre o quão proficiente o participante se declara e o nível de classificação atingido por ele no teste de vocabulário dentro dos níveis básicos.

Figura 8 - Qui-quadrado de união dos níveis mais avançados e separação dos níveis iniciais

Results						
	Autodeclaração	TTV				Row Totals
Nível 1	18 (25.50) [3.54]	35 (25.50) [3.54]				51
Nível 2	44 (33.00) [3.67]	22 (33.00) [3.67]				66
Níveis 3 e 4	73 (74.50) [0.03]	76 (74.50) [0.03]				149
Column Totals	133	133				266 (Grand Total)

Os dados das figuras 7 e 8 mostram a possibilidade dos níveis mais baixos de autodeclaração de proficiência não estarem de acordo com a classificação por níveis do TTV. Ou seja, há uma diferença significativa no número de participantes classificados nos níveis 1 e 2 em relação à autodeclaração de proficiência e o TTV. Já para os níveis 3 e 4, parecem existir indicações significativas de que os sujeitos têm tendências a serem classificados de maneira semelhante na autodeclaração de proficiência e no TTV.

É possível, então, a partir da interpretação desses dados, que não haja semelhança de classificação dos participantes, tanto na autodeclaração, como no TTV, nos níveis baixos de proficiência. Duas prováveis explicações para esse fenômeno podem ser estas: ou a

autodeclaração de proficiência não é eficaz ou o TTV não discrimina bem a proficiência (relacionada ao tamanho de vocabulário) nos níveis mais iniciantes. Cabe, dessa forma, a necessidade de estudos futuros que possam investigar essa questão.

5 DISCUSSÃO GERAL, LIMITAÇÕES E PERSPECTIVAS FUTURAS

Esta pesquisa teve o intuito de avaliar a proficiência linguística em L2 de bilíngues adultos do par português brasileiro-francês por meio do teste de amplitude vocabular *Teste de la Taille du Vocabulaire* (TTV) de Batista (2014). Outras medidas que costumam ser usadas para aferir proficiência foram correlacionadas com a pontuação de desempenho no teste, como a autodeclaração de proficiência, o tempo de estudo e a frequência de uso de L2.

A hipótese que norteou este estudo foi a de que o desempenho no TTV poderia demonstrar uma correlação direta com as medidas externas de autoavaliação de proficiência, tempo de estudo de L2 e frequência de uso de L2, significando que bilíngues que declarassem altos níveis de proficiência e que contabilizassem maior tempo de estudo e maior uso da língua francesa poderiam ter maior pontuação no TTV, classificando-se em níveis mais altos no teste; já os que declarassem baixos níveis de proficiência e tivessem menor tempo de estudo e menor frequência de uso poderiam demonstrar menor pontuação, classificando-se em níveis mais baixos do teste.

Apesar da obviedade da hipótese acima, as análises realizadas apontaram indícios de correlação moderada e fraca, respectivamente, entre os escores do TTV com autodeclaração e com o tempo de estudo.

Na relação entre o teste e a autopercepção sobre suas respectivas habilidades linguísticas, parece que os sujeitos da pesquisa têm uma autoapreciação compatível com suas pontuações no teste, indo de encontro com a discussão de Hulstijn (2012). Porém ainda caberia melhorar o questionário de autoavaliação, tal como mencionado por Bachman e Palmer (2010), Hulstijn (2012) e Treffers-Daller (2016), com o objetivo de ponderar até que ponto essa autoavaliação de proficiência seria fidedigna ao desempenho em testes de vocabulário.

Já no que concerne à relação entre desempenho no TTV e tempo de estudo, esperava-se uma correlação mais forte. Entretanto os resultados indicaram uma correlação fraca entre essas duas variáveis, permitindo uma certa abertura para questionamentos ligados ao papel da não-imersão geográfica no aprendizado de L2, ao modo de como a instrução formal e o acesso às atividades fora do contexto de sala de aula se realizam, estabelecendo-se um diálogo com a medida de frequência de uso.

Como também não houve evidências estatísticas de correlação entre a medida de frequência de uso e a pontuação no teste, diferentemente do que era previsto a partir de estudos anteriores (ELLIS, 2002, 2012; VALADARES, 2017), percebe-se que outros fatores, além das

três medidas investigadas, podem estar envolvidos no construto de proficiência, como a motivação em se aprender uma L2 ou até mesmo a performance de funções executivas. Através do intervalo de dados observados, tempo de estudo e frequência de uso explicariam somente 29% do desempenho no TTV, o que possibilita a manutenção da investigação de elementos envolvidos no bilinguismo.

Uma vez que não houve uma distribuição equilibrada de participantes entre os níveis do teste, uma análise posterior foi efetuada com o objetivo de buscar possíveis explicações para essa situação. Verificou-se que uma explanação plausível seria a de que o problema parece estar na diferenciação dos níveis mais baixos de proficiência. É como se o participante que declarasse proficiência nível 1 pudesse estar no nível 2 do TTV ou vice-versa. Essa possível não distinção dos níveis de baixa proficiência pode ser um problema do teste ou da autodeclaração, tal como discutido a respeito de sua confiabilidade em refletir proficiência. Entretanto, para confirmar tais suspeitas, são necessários mais estudos.

A respeito da metodologia aplicada para esta pesquisa, algumas ponderações sobre a quantidade de participantes e sobre a verificação do real conhecimento das palavras por parte deles cabem nesta seção. Reconhece-se que mais participantes seriam necessários para se obter um resultado mais robusto, uma vez que o estudo original contou com 175 participantes, e este, com 133.

O segundo comentário refere-se ao fato de não se saber realmente se os participantes tinham conhecimento sobre o significado das palavras apresentadas no teste, como averiguado por Batista (2014) através das entrevistas e dos questionários pós-teste. Apesar de serem orientados a deixarem a resposta em branco caso não soubessem responder, é possível haver acertos relacionados a tentativas de adivinhação, mesmo com a janela temporal estando presente, demonstrando a importância da entrevista pós-teste para verificar essa questão.

Já, na descrição dos participantes que responderam a aplicação do TTV original, a pesquisadora mencionou (BATISTA, 2014, p. 46) que 62% dos indivíduos residiam no Canadá por um ano ou menos até a data de aplicação do teste e que o tempo médio de permanência foi de um ano e nove meses, sugerindo que a maioria dos aprendizes que participaram de sua pesquisa não tinham longo tempo de residência no Québec. Isso pode gerar uma discussão produtiva a respeito do papel da imersão geográfica na proficiência sobre seu possível enfraquecimento. Contudo, tendo em mente que o objetivo do estudo original era o de medir a amplitude vocabular dos aprendizes de francês como L2 em níveis diferentes de proficiência, essa discussão pode ser revisitada futuramente.

Outro assunto importante de ser comentado refere-se à pergunta nove (ver APÊNDICE A – QUESTIONÁRIO APLICADO PARA BRASILEIROS) que procurou investigar as línguas faladas pelos sujeitos e a sua ordem de proficiência de acordo com seus julgamentos pessoais. Ao escreverem a sequência de línguas faladas da mais fluente para a menos fluente, admitindo-se que o português brasileiro fosse a língua mais fluente, 27,06% dos participantes registraram o francês em segundo lugar contra 64,66% que registraram o inglês (outras línguas correspondem ao restante minoritário dos dados). Apesar de a maioria dos voluntários não considerarem o francês como segundo idioma em termos de proficiência, o TTV não pode responder sobre o mecanismo de acesso a esses outros sistemas linguísticos por não apresentar componentes capazes de realizar esse tipo de mensuração. O que se sabe, como discutido na descrição do TTV original e no presente estudo, é que pode acontecer uma facilidade no desempenho de pontuação no teste para os falantes de línguas derivadas do latim pela semelhança encontrada em vocábulos.

De fato, o TTV preencheu uma lacuna relacionada ao desenvolvimento e à validação de uma medida psicométrica relacionada ao tamanho de vocabulário de francês como L2, contribuindo igualmente para os estudos de ensino de línguas. Já esta pesquisa pretendeu ser uma contribuição no campo do bilinguismo e do ensino de línguas estrangeiras através da discussão sobre a verificação da proficiência a partir das medidas de medição vocabular (por meio do TTV), de autodeclaração de proficiência, de tempo de estudo de L2 e de frequência de uso de L2.

Estudos futuros sobre os pontos de discussão aqui postulados continuam sendo necessários e indispensáveis para mais contribuições científicas sobre validação de testes de psicométricos de francês como L2 e sobre a configuração do conhecimento lexical como possível preditor do construto proficiência na área da psicolinguística.

REFERÊNCIAS

BACHMAN, L. F.; PALMER, A. S. The construct validation of self-ratings of communicative language ability. **Language Testing**, [s.l.], v. 6, n. 1, p. 14–29, June 1989.

BACHMAN, L. F. **Fundamental Considerations in Language Testing**. Oxford: Oxford University Press, 1990.

_____. **Language Assessment in Practice**. Oxford: Oxford University Press, 2010.

BATISTA, Roselene. **A receptive vocabulary knowledge test for french L2 learners with academic reading goals**. 2014. Dissertação (Mestrado em Linguística Aplicada) – Departement of Education, Concordia University, Montréal, 2014. Disponível em: <https://spectrum.library.concordia.ca/978313/> Acesso em: 26 maio 2018.

BAUDOT, J. **Les fréquences d'utilisation des mots en français écrit contemporain**. Montréal: Presses de l'Université de Montréal, 1992.

BIALYSTOK, Ellen. Factors in the growth of linguistic awareness. **Child Development**, Oxford, vol. 57, n. 2, p. 498–510, Apr. 1986.

BIALYSTOK, Ellen; BARAC, Raluca. Cognitive Effects. *In*: GROSJEAN, François; LI, Ping. **The Psycholinguistics of Bilingualism**. Malden: Wiley-Blackwell, 2013.

BIALYSTOK, Ellen; CRAIK, F. I. M.; LUK, G. Bilingualism: Consequences for Mind and Brain. **Trends in Cognitive Sciences**, Cambridge, v. 16, n. 4, p. 240-250, Apr. 2012.

BIALYSTOK, Ellen, *et al.* Receptive vocabulary differences in monolingual and bilingual children. **Bilingualism: Language and Cognition**, [s.l.], v. 13, n. 4, p. 525–531, Oct. 2010.

BIALYSTOK, Ellen; LUK, Gigi. Receptive vocabulary differences in monolingual and bilingual adults. **Bilingualism: Language and Cognition**, [s.l.], v. 15, n. 2, p. 397-401, Apr. 2012.

BIALYSTOK, Ellen; PEETS, Kathleen F.; MORENO, Sylvain. Producing bilinguals through immersion education: Development of metalinguistic awareness. **Applied Psycholinguistics**, Cambridge, v. 35, n. 1, p. 177-191, Jan. 2014.

BIRDSOING, D. Interpreting Age Effects in Second Language Acquisition. *In*: KROLL, J. F.; DE GROOT, A. M. B. (ed.) **Handbook of Bilingualism: Psycholinguistic Approaches**. Oxford: Oxford University Press: John Benjamins, 2005.

CENTRE NATIONAL DE RESSOURCES TEXTUELLES ET LEXICALES (CNRTL). Portail lexical [Web site], Nancy, 2012. Disponível em: <http://www.cnrtl.fr/definition/> Acesso em: 14 dez. 2018.

CHAMBERS, A.; LE BARON, F. **The Chambers–Le Baron Corpus of Research Articles in French / Le Corpus Chambers–Le Baron d'articles de recherche en français**. Oxford:

Oxford Text Archive, 2007. Disponível em: <http://ota.ahds.ac.uk/headers/2527.xml>. Acesso em: 4 out. 2018.

CLARK, E. Common ground. In: MACWHINNEY, B.; O'GRADY, W. (eds.). **The Handbook of Language Emergence**. Malden, MA: Wiley-Blackwell, 2015.

COSTA, A.; SEBASTIÁN-GALLÉS, N. How does the bilingual experience sculpt the brain? **Nature Reviews Neuroscience**, London, v. 15, p. 336-345, May 2014.

DIWERSY, S. **Varitext - Plateforme d'analyse lexico-statistique de variétés linguistiques**. Université de Cologne, Institut des Langues Romanes / Praxiling, Université Paul-Valéry Montpellier 3, 2013. Disponível em: <http://syrah.uni-koeln.de/varitext/>. Acesso em: 4 out. 2018.

ESHKOL-TARAVELLA, I.; BAUDE, O.; MAUREL, D.; HRIBA, L.; DUGUA, C.; TELLIER, I. 2011. Un grand corpus oral « disponible » : le corpus d'Orléans 1 1968–2012. **Traitement Automatique des Langues**, [s.l.], v. 53, n. 2, p. 17-46, 2011. Disponível em: <https://halshs.archives-ouvertes.fr/halshs-01163053/document> Acesso em: 14 dez. 2018.

ELLIS, Nick. Frequency Effects in Language Processing: A review with implications for theories of implicit and explicit language acquisition. **Studies in second language acquisition**, Cambridge, v. 24, n. 2, p. 143-188, June, 2002.

_____. What can we count in language, and what counts in language acquisition, cognition, and use? In: GRIES, S. T.; DIVJAK, D. **Frequency Effects in Language Learning and Processing**. Berli: De Gruyter Mouton, 2012.

_____. Implicit AND explicit language learning: Their dynamic interface and complexity. In: REBUSCHAT, Patrick (ed.). **Implicit and Explicit Learning of Languages**. Amsterdam: John Benjamins Publishing Company, 2015.

FRANCIS, W. N.; KUČERA, H. **Manual of Information to Accompany 'A Standard Sample of Present-Day Edited American English, for Use with Digital Computers'**. Ed. rev. 1979. Providence, RI: Department of Linguistics, Brown University, 1964.

GADET, F. Stylistic and syntactic variation: introduction. In: BEECHING K.; ARMSTRONG, N.; GADET, F. (ed.) **Sociolinguistic Variation in Contemporary French**. Amsterdam: John Benjamins, 2009, p. 115-120.

GRANERA, G.; LONG, M. H. Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. **Second Language Research**, [s.l.], v. 29, n. 3, p. 311-343, Dec. 2012.

GROSJEAN, F. Studying Bilinguals: Methodological and Conceptual Issues. **Bilingualism: Language and Cognition**, [s.l.], v. 1, n. 2, p. 131-149, Aug. 1998.

GROSJEAN, François; LI, Ping. **The Psycholinguistics of Bilingualism**. Malden: Wiley-Blackwell, 2013.

- HARRIS, D. P. **Testing English as a Second Language**. New York: McGraw-Hill, 1969.
- HULSTIJN, Jan H. The Construct of Language Proficiency in the Study of Bilingualism from a Cognitive Perspective. **Bilingualism: Language and Cognition**, [s.l.], v. 15, n. 2, p. 422-433, Apr. 2012.
- KIRSNER, K. Implicit processes in second language learning. *In*: ELLIS, N. (ed.), **Implicit and explicit learning of languages**. San Diego, CA: Academic Press, 1994.
- LADO, R. **Language Testing**. London: Longman, 1961.
- LENNEBERG, Eric. **Biological foundations of language**. New York, NY: Wiley, 1967.
- LONSDALE, D.; LE BRAS, Y. **A frequency dictionary of French**. New York, NY: Routledge, 2009.
- MILTON, J. **Measuring second language vocabulary acquisition**. Bristol: Multilingual Matters, 2009.
- MONTRUL, Silvina. Dominance and proficiency in early and late bilinguals. *In*: SILVA-CORVALÁN, Carmen; TREFFERS-DALLER, Jeanine (ed.). **Language dominance in bilinguals: issues of measurement and operationalization**. Cambridge: Cambridge University Press, 2016.
- NATION, P. **Teaching and Learning Vocabulary**. Boston, MA: Heinle & Heinle, 1990.
- NATION, P.; BEGLAR, D. A vocabulary size test. **The Language Teacher**, [s.l.], v. 31, n. 7, p. 9-13, July 2007.
- QUIRK, Randolph; SVARTVIK, Jan. **Investigating linguistic acceptability**. The Hague: Mouton, 1966.
- READ, John. **Assessing Vocabulary**. Cambridge: Cambridge University Press, 2000.
- ROSS, John A. The Reliability, Validity, and Utility of Self-Assessment. **Practical Assessment, Research & Evaluation**, [s.l.], v. 11, n. 10, p. 1-13, Nov. 2006.
- SCHMITT, N. **Researching vocabulary: a vocabulary research manual**. Hampshire: Palgrave and Macmillan, 2010.
- SCHMITT, N.; SCHMITT, D.; CLAPHAM, C. Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. **Language Testing**, [s.l.], v. 18, n. 1, p. 55-88, Jan. 2001.
- SEGALOWITZ, N. Automaticity and Second Languages. *In*: DOUGHTY, Catherine J.; LONG, Michael H. (ed.) **The Handbook of Second Language Acquisition**. Malden: Blackwell Publishing, 2005.

SIEPMANN, D.; BÜRCEL, C.; DIWERSY, S. The Corpus de Référence du Français Contemporain (CRFC) as the first genre-diverse mega-corpus of French. **International Journal of Lexicography**, Oxford, v. 30, n. 1, p. 63-84, Dec. 2015. Disponível em: <https://academic.oup.com/ijl/article-abstract/30/1/63/2584094> Acesso em: 4 out. 2018.

_____. Le Corpus de référence du français contemporain (CRFC), un corpus massif du français largement diversifié par genres. **SHS Web of Conferences**, [s.l.], v. 27, n. 11002, Jan. 2016, Disponível em: <https://doaj.org/article/2c37920598f84269bebaca5747a1352b> Acesso em: 4 out. 2018.

SILVA, Jesiel Soares. **Exploring a Vocabulary Test and a Judgment Task as Diagnoses of Early and Late Bilinguals' L2 Proficiency**. 2016. 143 f. Tese (Doutorado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2016. Disponível em <http://hdl.handle.net/1843/MGSS-A8RN2Z>. Acesso em: 4 out. 2018.

SOUZA, R.; DUARTE, D.; BERG, I. A avaliação de habilidades linguísticas em L2: uma questão metodológica em estudos de influências translinguísticas. In: BUCHWEITZ, A. & MOTA, M. B. (ed.) **Linguagem e Cognição: Processamento, Aquisição e Cérebro**. Porto Alegre: EdUPUCRS, 2015.

SUNDERMAN, G.; KROLL, J. First language activation during second language lexical processing: an investigation of lexical form, meaning, and grammatical class. **Studies in Second Language Acquisition**, [s.l.], v. 28, n. 3, p. 387-422, Sept. 2006.

TYNE, H. *et al.* **French through Corpora: Ecological and Data-Driven Perspectives in French Language Studies**. Newcastle upon Tyne: Cambridge Scholars Publishing, 2014.

TREFFERS-DALLER, Jeanine. Operationalizing and measuring language dominance. **International Journal of Bilingualism**, v. 15, n. 2, p. 147–163, June 2011.

TREFFERS-DALLER, J. Language dominance: The constructo, its measurement, and operationalization. In: SILVA-CORVALÁN, Carmen; TREFFERS-DALLER, Jeanine (ed.). **Language dominance in bilinguals: issues of measurement and operationalization**. Cambridge: Cambridge University Press, 2016.

VALADARES, Marcus Guilherme Pinto de Faria. **Padrões emergentes de dominância linguística em português e inglês: o impacto de práticas socioculturais de letramento (digital) na amplitude lexical de brasileiros falantes de língua inglesa como L2**. 2017. 146 f. Tese (Doutorado em Linguística Aplicada) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2017. Disponível em: <http://hdl.handle.net/1843/LETR-AP5HTS>. Acesso em: 4 out. 2018.

VIEIRA, Sônia. **Estatística básica**. São Paulo: Cengage Learning, 2013.

YUN, H.; CHANIER, T. **Corpus d'apprentissage FAVI (Français académique virtuel international)**. Banque de corpus CoMeRe. Nancy: Ortolang.fr, 2011. Disponível em: <http://olac ldc.upenn.edu/item/oai:mulce.org:mce-favi-letec-all>. Acesso em: 3 dez. 2018.

ZELL, E.; KRIZAN, Z. Do people have insight into their abilities? A metasynthesis. **Perspectives on Psychological Science**, [s.l.], v. 9, n. 2, p. 111-125, Mar. 2014.

APÊNDICE A – QUESTIONÁRIO APLICADO PARA BRASILEIROS

Apresentação

Convidamos-lhe a participar de um experimento que compõe uma pesquisa de mestrado sobre o processamento da linguagem por bilíngues, de Isabelle Deolinda Pereira de Sousa, do Programa de Pós-Graduação em Estudos Linguísticos da UFMG, sob a orientação do Prof. Dr. Ricardo Augusto de Souza (Faculdade de Letras - UFMG). É necessário que você seja brasileiro falante de Francês em qualquer nível (iniciante, intermediário ou proficiente).

Esse questionário é a primeira fase de coleta de dados e, ao final dele, você receberá um link e um código de acesso para o experimento em si e que só poderá ser acessado através de seu computador (o acesso por smartphone ou tablet pode comprometer seu desenvolvimento). A tarefa consiste na associação de palavras em Francês, com duração total de 14 minutos. Ela exige concentração, podendo portanto parecer cansativa. A qualquer instante você poderá interromper sua participação, sem nenhuma espécie de ônus, porém sua participação é muito importante para a condução desse estudo.

A qualquer momento, você pode esclarecer dúvidas através do e-mail isabelledpsousa2@gmail.com.

Esse experimento também pode ser revertido em créditos para atividades de cunho científico-cultural para os estudantes da UFMG.

Caso concorde com as explicações acima e queira dar continuidade ao questionário e, em seguida, ao experimento, informe seu e-mail abaixo.

* Required

1. **Email address ***

Dados do participante

Todos os dados coletados serão convertidos em informações numéricas e quantitativas que não serão associadas a nomes ou indivíduos em momento algum. As informações coletadas serão acessadas exclusivamente pela pesquisadora desse estudo. Portanto, todos os participantes terão anonimato garantido, e sob circunstância alguma, sua identidade como participante será revelada.

2. **Nome completo ***

3. **Sexo ***

Mark only one oval.

Feminino

Masculino

4. **Idade (em anos) ***

5. Escolaridade *

Mark only one oval.

- Ensino médio incompleto
- Ensino médio completo
- Ensino superior incompleto
- Ensino superior completo
- Especialização incompleta
- Especialização completa
- Mestrado incompleto
- Mestrado completo
- Doutorado incompleto
- Doutorado completo

6. Curso? *

7. Se você é estudante da UFMG e pretende aproveitar esse experimento como crédito para atividade de cunho científico-cultural, informe seu número de matrícula. Caso contrário, continue para a próxima seção.

8. Se você não é aluno da UFMG, mencione a instituição de ensino a qual você é/foi vinculado(a):

9. Escreva as línguas que você fala na sequência da mais fluente (assumimos que seja o Português) para a menos fluente de acordo com sua percepção. *

Por exemplo: 1- Português; 2 - Inglês; 3 - Francês

10. Há quanto tempo (em anos) você estuda francês? *

11. Se você tiver uma certificação de proficiência em Francês, mencione abaixo o diploma (DELF, DALF, TCF etc.), o nível de classificação (A1, A2, B1, B2, C1 ou C2) e o ano de obtenção. Se não, deixe a resposta em branco.

Questionário de uso de língua francesa

12. Classifico a minha habilidade de leitura em francês como *

Mark only one oval.

- Fluente
 Avançada
 Média
 Básica
 Principiante

13. Classifico a minha habilidade de compreensão do francês falado como *

Mark only one oval.

- Fluente
 Avançada
 Média
 Básica
 Principiante

14. Classifico a minha habilidade de escrever em francês como *

Mark only one oval.

- Fluente
 Avançada
 Média
 Básica
 Principiante

15. **Classifico a minha habilidade de falar em francês como ***

Mark only one oval.

- Fluente
 Avançada
 Média
 Básica
 Princiante

16. **Identifique a frequência aproximada com a qual você realiza as atividades abaixo EM FRANCÊS. ***

Mark only one oval per row.

	Nunca	Raramente	Ocasionalmente	Frequentemente	Sempre
Leio artigos de jornais e revistas (impresso ou online)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Leio materiais para fins profissionais (trabalho) e/ou acadêmicos (estudos)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Leio por prazer e como atividade de lazer	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pesquisa por informações na internet cujo conteúdo é escrito	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pesquisa por informações na internet cujo conteúdo é falado	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Procuo assistir a filmes, a seriados e a outros programas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Escrevo qualquer tipo de texto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Troco mensagens pela Internet ou em outros meios	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Treino a oralidade com quem fala a língua em nível intermediário, avançado ou fluente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Assisto a programas / filmes / seriados falados em francês com legendas em português	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Assisto a programas / filmes / seriados falados em francês com legenda em francês	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Assisto a programas / filmes / seriados falados em francês sem legenda	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

APÊNDICE B – RESUMO DOS DADOS SOBRE O PERFIL DOS BILÍNGUES

ESCOLARIDADE	QUANTIDADE	
Ensino médio completo	2	1,5%
Ensino superior incompleto	71	53,4%
Ensino superior completo	18	13,5%
Especialização incompleta	3	2,3%
Especialização completa	6	4,5%
Mestrado incompleto	11	8,3%
Mestrado completo	7	5,3%
Doutorado incompleto	11	8,3%
Doutorado completo	4	3,0%
TOTAL	133	100,0%

LÍNGUAS DE MAIOR A MENOR PROFICIÊNCIA SEGUNDO A PERCEPÇÃO DOS PARTICIPANTES

Português	2	1,5%
Português - Francês	4	3,0%
Português - Francês - Inglês	16	12,0%
Português - Inglês - Francês	43	32,3%
Português - Espanhol - Francês - Inglês	1	0,8%
Português - Espanhol - Inglês - Francês	7	5,3%
Português - Francês - Espanhol - Inglês	1	0,8%
Português - Francês - Inglês - Espanhol	8	6,0%
Português - Francês - Italiano - Inglês	2	1,5%
Português - Inglês - Espanhol - Mandarim	1	0,8%
Português - Inglês - Francês - Alemão	2	1,5%
Português - Italiano - Inglês - Francês	1	0,8%
Português - Inglês - Alemão - Francês	1	0,8%
Português - Inglês - Espanhol - Francês	11	8,3%
Português - Inglês - Francês - Cantonês	1	0,8%
Português - Inglês - Francês - Coreano	1	0,8%
Português - Inglês - Francês - Dinamarques	1	0,8%
Português - Inglês - Francês - Espanhol	5	3,8%
Português - Inglês - Francês - Italiano	4	3,0%
Português - Inglês - Francês - Japonês	1	0,8%
Português - Inglês - Italiano - Francês	1	0,8%
Português - Francês - Inglês - Espanhol - Italiano	1	0,8%
Português - Francês - Inglês - Espanhol - Japonês	1	0,8%
Português - Francês - Inglês - Italiano - Espanhol	1	0,8%
Português - Francês - Inglês - Norueguês - Espanhol	1	0,8%
Português - Inglês - Espanhol - Francês - Italiano	1	0,8%
Português - Inglês - Espanhol - Francês - Mandarim	1	0,8%
Português - Inglês - Francês - Espanhol - Alemão	2	1,5%
Português - Inglês - Francês - Espanhol - Italiano	2	1,5%
Português - Inglês - Francês - Espanhol - Latim	1	0,8%
Português - Inglês - Francês - Espanhol - Mandarim	1	0,8%
Português - Inglês - Italiano - Francês - Alemão	1	0,8%
Português - Inglês - Espanhol - Francês - Alemão - Japonês	1	0,8%
Português - Inglês - Francês - Espanhol - Italiano - Alemão	1	0,8%
Português - Inglês - Francês - Espanhol - Mandarim - Alemão	1	0,8%
Inglês - Espanhol - Italiano - Alemão - Francês	1	0,8%
Sem resposta	1	0,8%
TOTAL	133	100,0%

	QUANTIDADE	
Pessoas com certificação	30	22,6%

APÊNDICE C – QUESTIONÁRIO APLICADO PARA NATIVOS

Présentation

Nous vous invitons à participer à cette recherche sur le traitement linguistique par des personnes bilingues, étude réalisée par l'étudiante de master Isabelle Deolinda PEREIRA DE SOUSA, attachée au Programme d'Études Supérieures en Linguistique à l'Université Fédérale du Minas Gerais (Brésil), avec M. Ricardo Augusto DE SOUZA (Faculté de Lettres - UFMG) comme directeur de recherche. Pour participer à cette recherche, vous devez avoir le français comme première langue, peu importe le pays francophone dont vous venez. Il est important que vous habitiez dans votre pays d'origine ou dans un autre pays francophone pour que l'utilisation de votre première langue soit toujours fréquente.

Ce questionnaire est la première phase de collecte de données. A la fin de celui-ci, vous verrez un lien et un code d'accès pour la tâche elle-même, qui n'est accessible que sur ordinateur (l'accès par smartphone ou tablette peut compromettre le développement). La tâche consiste à associer des mots en français, et a une durée totale de 14 minutes. Il faut être concentré et cela peut donc sembler fatigant. A tout moment vous pouvez arrêter votre participation, sans aucune charge, mais votre participation est très importante pour la conduite de cette étude.

A tout moment, vous pouvez poser des questions en envoyant un e-mai à isabelledpsousa2@gmail.com

Si vous êtes d'accord avec les explications ci-dessus et si vous voulez poursuivre le questionnaire et la tâche suivante, entrez votre adresse e-mail ci-dessous. Nous vous remercions pour votre participation !

* Required

1. Email address *

Données du participant

Toutes les données collectées seront converties en informations numériques et quantitatives qui ne seront associées à aucun nom ni à aucun individu. Les informations recueillies seront exclusivement accessibles à la chercheuse de cette étude. L'anonymat est donc garanti à tous les participants.

3. Nom *

4. Sexe *

Mark only one oval.

Homme

Femme

5. Âge *

6. Ville et pays d'origine *

7. Si vous habitez dans un pays différent de celui de votre origine, écrivez la ville et le pays où vous habitez maintenant.

8. Si vous habitez dans un pays différent de celui de votre origine, il y a combien de temps que vous êtes parti (e) ?

9. Scolarité *

Mark only one oval.

- Lycée incomplet
- Lycée complet
- Licence / Bachelier / Baccalauréat incomplet (e)
- Licence / Bachelier / Baccalauréat complet (e)
- Master / Maîtrise incomplet (e)
- Master / Maîtrise complet (e)
- Doctorat incomplet
- Doctorat complet

10. Nom du cours *

APÊNDICE D – DADOS DAS VARIÁVEIS DE ANÁLISE DOS BRASILEIROS

Tabela 11 – Dados de autodeclaração de proficiência, tempo de estudo, frequência de uso do francês e pontuação total no TTV

(Continua)

Sujeito	Autodecl. total	Nível autodecl.	TE	Freq	Acertos TTV	Nível TTV
1	4	1	12	7	59	2
2	7	2	12	3	61	1
3	5	1	18	2	40	1
4	5	1	6	9	11	1
5	17	4	96	2	83	4
6	4	1	6	0	32	1
7	12	3	36	11	70	3
8	9	2	24	8	64	3
9	10	2	24	7	62	1
10	8	2	12	0	49	1
11	9	2	72	10	58	3
12	16	4	108	17	79	4
13	7	2	72	7	63	3
14	12	3	60	15	61	3
15	18	4	168	26	85	4
16	13	3	36	9	92	4
17	18	4	108	14	92	4
18	14	3	36	24	93	4
19	10	2	36	3	56	2
20	17	4	36	10	73	3
21	13	3	72	6	85	4
22	8	2	24	15	42	2
23	10	2	48	15	29	1
24	10	2	30	18	77	1
25	11	3	72	5	72	2
26	8	2	12	3	38	2

Tabela 11 – Dados de autodeclaração de proficiência, tempo de estudo, frequência de uso do francês e pontuação total no TTV

(Continua)

Sujeito	Autodecl. total	Nível autodecl.	TE	Freq	Acertos TTV	Nível TTV
27	7	2	12	4	87	4
28	7	2	6	7	51	1
29	18	4	36	26	21	1
30	12	3	36	21	35	1
31	10	2	24	4	30	2
32	7	2	12	4	62	1
33	5	1	6	11	40	2
34	8	2	24	8	43	1
35	10	2	30	7	60	1
36	12	3	30	7	97	4
37	8	2	24	16	68	3
38	18	4	48	18	80	4
39	18	4	24	1	85	4
40	11	3	12	9	52	3
41	13	3	12	11	89	4
42	14	3	48	11	98	4
43	4	1	5	19	28	1
44	20	4	276	25	103	4
45	11	3	34	11	35	1
46	11	3	84	3	52	3
47	6	2	12	6	46	1
48	19	4	24	1	89	4
49	16	4	36	1	105	4
50	15	3	18	8	85	4
51	9	2	24	5	83	4
52	5	1	7	7	44	1

Tabela 11 – Dados de autodeclaração de proficiência, tempo de estudo, frequência de uso do francês e pontuação total no TTV

(Continua)

Sujeito	Autodecl. total	Nível autodecl.	TE	Freq	Acertos TTV	Nível TTV
53	10	2	36	9	57	3
54	10	2	48	8	88	4
55	11	3	24	20	84	3
56	18	4	72	7	98	4
57	16	4	96	15	85	4
58	18	4	108	5	91	4
59	12	3	12	20	57	3
60	13	3	48	10	83	3
61	16	4	60	10	65	3
62	7	2	24	3	63	3
63	8	2	120	4	62	1
64	12	3	36	9	82	4
65	12	3	48	9	75	3
66	20	4	54	16	62	3
67	17	4	96	28	61	3
68	15	3	120	13	107	4
69	8	2	12	2	81	4
70	13	3	48	20	83	4
71	4	1	12	2	50	1
72	8	2	12	8	44	2
73	12	3	36	7	74	3
74	20	4	180	9	106	4
75	9	2	36	10	54	1
76	9	2	12	10	44	1
77	8	2	24	0	39	2
78	13	3	12	18	40	2
79	16	4	36	9	60	2
80	9	2	30	7	77	3

Tabela 11 – Dados de autodeclaração de proficiência, tempo de estudo, frequência de uso do francês e pontuação total no TTV

(Continua)

Sujeito	Autodecl. total	Nível autodecl.	TE	Freq	Acertos TTV	Nível TTV
81	13	3	8	15	24	2
82	11	3	36	13	58	1
83	4	1	6	0	39	1
84	12	3	36	6	69	2
85	16	4	168	5	93	4
86	14	3	96	15	90	4
87	4	1	3	3	44	1
88	16	4	24	5	29	2
89	6	2	48	11	77	4
90	9	2	18	9	44	2
91	15	3	24	25	87	3
92	9	2	12	1	32	1
93	11	3	24	7	63	3
94	12	3	24	8	68	2
95	20	4	120	28	62	3
96	14	3	120	12	86	3
97	10	2	12	6	64	3
98	19	4	60	30	77	3
99	17	4	48	16	82	4
100	7	2	36	1	53	3
101	4	1	6	3	47	3
102	9	2	12	6	59	2
103	12	3	24	6	62	3
104	5	1	12	8	61	2
105	20	4	168	22	103	4
106	12	3	24	18	66	3
107	15	3	60	6	102	4
108	20	4	144	21	84	4

Tabela 11 – Dados de autodeclaração de proficiência, tempo de estudo, frequência de uso do francês e pontuação total no TTV

(Conclusão)

Sujeito	Autodecl. total	Nível autodecl.	TE	Freq	Acertos TTV	Nível TTV
109	16	4	96	14	71	3
110	5	1	6	0	58	1
111	6	2	6	6	74	4
112	10	2	18	10	72	2
113	15	3	36	21	104	4
114	12	3	48	23	36	2
115	16	4	60	12	85	4
116	4	1	6	2	56	1
117	11	3	144	8	91	4
118	6	2	3	4	41	1
119	10	2	24	7	65	3
120	16	4	132	17	87	4
121	5	1	12	5	35	1
122	6	2	12	4	33	1
123	8	2	3	19	55	1
124	18	4	240	14	55	3
125	8	2	6	14	64	1
126	16	4	120	21	22	1
127	8	2	36	10	22	1
128	4	1	12	6	42	2
129	17	4	84	28	92	4
130	12	3	24	14	90	4
131	13	3	48	11	38	2
132	16	4	24	15	90	4
133	13	3	60	17	57	3

ANEXO A - TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Convidamos-lhe a participar de um experimento que compõe uma pesquisa sobre o processamento da linguagem por bilíngues. O experimento do qual você participará envolverá a realização de uma tarefa no computador de associação de palavras. No início da tarefa, o experimentador fornecerá as instruções e esclarecerá dúvidas eventuais sobre sua execução. Todos os dados coletados serão convertidos em informações numéricas e quantitativas que não serão associadas a nomes ou indivíduos em momento algum. As informações coletadas serão acessadas exclusivamente pelos pesquisadores deste estudo. Portanto, todos os participantes terão anonimato garantido, e sob circunstância alguma, sua identidade como participante será revelada. A tarefa tem duração de aproximadamente 20 minutos e exige concentração, podendo portanto parecer-lhes cansativa. A qualquer instante você poderá interromper sua participação, sem nenhuma espécie de ônus. Por outro lado, sua participação neste estudo não envolve nenhuma forma de risco a sua saúde, não lhe trará qualquer desconforto físico ou mental, e nem tampouco poderá ser comprometedor de sua reputação como cidadão, como estudante ou como profissional. Sua participação deste estudo não acarretará benefícios materiais imediatos para você, individualmente. Os resultados deste estudo, obtidos após tratamento estatístico da totalidade das informações coletadas, poderão vir a público na forma de apresentações de comunicações orais ou pôster em congressos científicos; artigos científicos publicados em periódicos especializados; capítulos de livros; monografias, dissertações ou teses orientadas pelo pesquisador principal. Contudo, pelas razões acima enumeradas as identidades dos participantes permanecerão ocultas em toda e qualquer forma de publicação e divulgação do estudo.

Caso você tenha qualquer dúvida, sinta-se à vontade para fazer contato com o pesquisador responsável. Você também pode consultar o COEP (Comitê de Ética e Pesquisa da Universidade Federal de Minas Gerais) para tirar dúvidas sobre aspectos éticos da pesquisa.

Pesquisador responsável:

Prof. Dr. Ricardo Augusto de Souza
 Faculdade de Letras /Programa de Pós-Graduação em Estudos Linguísticos
 Universidade Federal de Minas Gerais – Brasil
 ricsouza@ufmg.br

Prof. Dr. Ricardo Augusto de Souza

Comitê de Ética na Pesquisa (COEP):

Av. Antônio Carlos, 6627
 Unidade Administrativa II - 2º andar - Sala 2005
 Campus Pampulha
 Belo Horizonte, MG – Brasil
 31270-901
 (31) 3409-4592

Se você aceitar participar, por favor preencha e assine o formulário de consentimento abaixo. Agradecemos-lhe por sua participação!

FORMULÁRIO DE CONSENTIMENTO:

Eu, _____, concordo em participar da tarefa que compõe o estudo experimental sobre a linguagem no bilinguismo supervisionado pelo Prof. Dr. Ricardo A. de Souza, da Universidade Federal de Minas Gerais. Estou ciente de que minha participação do estudo não implica danos ou riscos a minha saúde e bem-estar, que ela não implica riscos à minha reputação e imagem, assim como que ela não implica benefícios materiais imediatos para mim, individualmente. Estou também ciente de que posso interromper minha participação do experimento a qualquer momento, sem ônus algum, e de que minha identidade como participante não será exposta em nenhuma forma de divulgação dos resultados do estudo.

Data: ___/___/___

Assinatura: _____

ANEXO B – TESTE DE LA TAILLE DU VOCABULAIRE (TTV) APLICADO AOS PARTICIPANTES

TTV - Test de la taille du vocabulaire

Étude de maîtrise de Roselene Batista, Université Concordia, février 2014

La deuxième tranche de mille mots

1. concours
2. division
3. joie
4. phase
5. stade
6. véhicule

_____ grand plaisir
 _____ un moyen de transport
 _____ séparation en deux parties

1. autorisation
2. bonjour
3. confusion
4. faim
5. rupture
6. tribunal

_____ erreur
 _____ le besoin de manger
 _____ la maison de la justice

1. adapter
2. crier
3. distribuer
4. formuler
5. procéder
6. traverser

_____ partager
 _____ parler très fort
 _____ aller d'un côté à l'autre

1. bras
2. circuit
3. détermination
4. match
5. réception
6. théorie

_____ tour
 _____ principe et règle
 _____ une partie du corps

1. bâtiment
2. consultation
3. essence
4. habitude
5. leçon
6. prestation

_____ enquête
 _____ construction
 _____ la façon normale de faire

1. brûler
2. distinguer
3. examiner
4. mentionner
5. rêver
6. supprimer

_____ imaginer
 _____ remarquer
 _____ détruire par le feu

1. fondamental
2. global
3. moderne
4. prudent
5. récent
6. traditionnel

_____ complet
 _____ qui est la base
 _____ qui ne prend pas de risques

1. attaque
2. contribution
3. dommage
4. église
5. incident
6. mécanisme

_____ institution
 _____ action violente
 _____ ensemble de pièces

1. actif
2. inutile
3. fier
4. majeur
5. puissant
6. scolaire

_____ occupé
 _____ qui ne sert à rien
 _____ qui a un grand pouvoir

1. dégager
2. élire
3. mériter
4. persuader
5. résumer
6. soulever

_____ choisir
 _____ convaincre
 _____ raconter de façon brève

La troisième tranche de mille mots

1. ambassadeur
2. enfance
3. portrait
4. rayon
5. trouble
6. vœu

_____ conflit
 _____ première partie de la vie
 _____ représentant du gouvernement

1. aube
2. docteur
3. issue
4. législation
5. préparation
6. reportage

_____ lever du soleil
 _____ science des lois
 _____ article d'un journaliste

1. amateur
2. cellule
3. expansion
4. profil
5. sondage
6. terrorisme

_____ petite chambre
 _____ un visage vu de côté
 _____ des questions et des réponses

1. brutal
2. formidable
3. impressionnant
4. mobile
5. obligatoire
6. raisonnable

_____ dur
 _____ juste
 _____ nécessaire

1. automobile
2. barre
3. dominant
4. paquet
5. race
6. tourisme

_____ espèce
 _____ objet long et étroit
 _____ activité d'une personne qui voyage

1. chéri
2. décisif
3. magnifique
4. rigoureux
5. tragique
6. vital

_____ strict
 _____ terrible
 _____ celui qui est très aimé

1. douleur
2. minorité
3. permanence
4. rédaction
5. sagesse
6. transition

_____ passage
 _____ texte écrit
 _____ bon sens et connaissance

1. annuler
2. cultiver
3. défaire
4. mentir
5. plaider
6. siéger

_____ travailler la terre
 _____ ne pas dire la vérité
 _____ prendre la défense d'une cause

1. commenter
2. engendrer
3. promouvoir
4. remporter
5. songer
6. téléphoner

_____ faire naître
 _____ gagner un jeu
 _____ élever à un rang supérieur

1. avertir
2. blesser
3. caractériser
4. déclencher
5. serrer
6. trahir

_____ définir
 _____ signaler
 _____ provoquer un phénomène

La cinquième tranche de mille mots

1. cote
2. foumiture
3. goutte _____ le milieu entre deux extrêmes
4. hépatite _____ très petite quantité d'un liquide
5. juridiction _____ des objets utilisés en salle de classe
6. moyenne

1. anormal
2. convaincant
3. éminent _____ froid
4. insensible _____ bizarre
5. mensuel _____ qui arrive une fois par mois
6. séparatiste

1. affection
2. chancelier
3. fouet _____ premier ministre
4. golfe _____ un type de sentiment
5. lance _____ compétence et expérience
6. qualification

1. amusant
2. boursier
3. irrégulier _____ mauvais
4. malin _____ qui n'est pas uniforme
5. naïf _____ qui concerne le marché financier
6. solidaire

1. aviser
2. capter
3. desservir _____ informer
4. empresser _____ changer le sens
5. inverser _____ formuler des conditions précises
6. stipuler

1. analogue
2. barbare
3. échéant _____ pareil
4. intact _____ qui n'a pas été touché
5. philosophique _____ qui cause une grande surprise
6. stupéfiant

1. brouillard
2. coïncidence
3. farce _____ une histoire qui fait rire
4. instituteur _____ ce qui empêche de voir loin
5. pneu _____ un professionnel de l'éducation
6. soumission

1. coffre
2. convoi
3. duc _____ abri
4. inondation _____ caisse où l'on met de l'argent
5. nuance _____ différence entre deux choses semblables
6. tente

1. cage
2. panneau
3. repli _____ défaut
4. sou _____ prison
5. traite _____ monnaie
6. vice

1. confiner
2. énumérer
3. imputer _____ isoler
4. octroyer _____ être suspendu
5. pendre _____ dresser une liste
6. venger

La dixième tranche de mille mots

1. amas
2. dortoir
3. loque _____ mots, phrases, règles
4. huissier _____ chambre à plusieurs lits
5. nain _____ personne de très petite taille
6. syntaxe

1. armure
2. charpentier
3. épinette _____ espèce d'arbre
4. granit _____ vêtement en métal
5. hypothèque _____ celui qui construit des immeubles
6. torsade

1. cafard
2. étang
3. fenouil _____ lac
4. ouïe _____ plante
5. répit _____ pause
6. sangle

1. buée
2. dicton
3. fût _____ ce que l'on utilise pour frapper
4. gamiture _____ destiné à contenir des liquides
5. maillet _____ homme qui marche dans les rues
6. piéton

1. avortement
2. matrice
3. persévérance _____ jeune soldat
4. ravitaillement _____ morceau coupé d'un objet
5. recrue _____ qualité de celui qui a de la patience
6. tronçon

1. astiquer
2. foisonner
3. incarcérer _____ enfermer quelqu'un
4. marteler _____ avoir en grande quantité
5. railler _____ faire hésiter entre deux choses
6. tirailler

1. bourrer
2. crouler
3. émanciper _____ remplir
4. mijoter _____ détruire ou tomber
5. patauger _____ revêtir d'une couleur
6. teinter

1. culbuter
2. décupler
3. grelotter _____ souffrir du froid
4. poncer _____ couper un arbre
5. pulvériser _____ briser ou réduire en morceaux
6. scier

1. âpre
2. fulgurant
3. hébété _____ violent ou rapide
4. rusé _____ fort, gros et solide
5. trapu _____ qui est désagréable au toucher
6. utopique

1. ahuri
2. dru
3. nutritif _____ étonné
4. opaque _____ qui n'est pas transparent
5. pondéré _____ très animé ou mouvementé
6. trépidant

ANEXO C – RESPOSTAS DO *TESTE DE LA TAILLE DU VOCABULAIRE* (TTV)

La deuxième tranche de mille mots

1. concours
2. division
3. joie 3 grand plaisir
4. phase 6 un moyen de transport
5. stade 2 séparation en deux parties
6. véhicule

1. autorisation
2. bonjour
3. confusion 3 erreur
4. faim 4 le besoin de manger
5. rupture 6 la maison de la justice
6. tribunal

1. adapter
2. crier
3. distribuer 3 partager
4. formuler 2 parler très fort
5. procéder 6 aller d'un côté à l'autre
6. traverser

1. bras
2. circuit
3. détermination 2 tour
4. match 6 principe et règle
5. réception 1 une partie du corps
6. théorie

1. bâtiment
2. consultation
3. essence 2 enquête
4. habitude 1 construction
5. leçon 4 la façon normale de faire
6. prestation

1. brûler
2. distinguer
3. examiner 5 imaginer
4. mentionner 2 remarquer
5. rêver 1 détruire par le feu
6. supprimer

1. fondamental
2. global
3. moderne 2 complet
4. prudent 1 qui est la base
5. récent 4 qui ne prend pas de risques
6. traditionnel

1. attaque
2. contribution
3. dommage 4 institution
4. église 1 action violente
5. incident 6 ensemble de pièces
6. mécanisme

1. actif
2. inutile
3. fier 1 occupé
4. majeur 2 qui ne sert à rien
5. puissant 5 qui a un grand pouvoir
6. scolaire

1. dégager
2. élire
3. mériter 2 choisir
4. persuader 4 convaincre
5. résumer 5 raconter de façon brève
6. soulever

La troisième tranche de mille mots

1. ambassadeur
2. enfance
3. portrait 5 conflit
4. rayon 2 première partie de la vie
5. trouble 1 représentant du gouvernement
6. vœu

1. aube
2. docteur
3. issue 1 lever du soleil
4. législation 4 science des lois
5. préparation 6 article d'un journaliste
6. reportage

1. amateur
2. cellule
3. expansion 2 petite chambre
4. profil 4 un visage vu de côté
5. sondage 5 des questions et des réponses
6. terrorisme

1. brutal
2. formidable
3. impressionnant 1 dur
4. mobile 6 juste
5. obligatoire 5 nécessaire
6. raisonnable

1. automobile
2. barre
3. dominant 5 espèce
4. paquet 2 objet long et étroit
5. race 6 activité d'une personne qui voyage
6. tourisme

1. chéri
2. décisif
3. magnifique 4 strict
4. rigoureux 5 terrible
5. tragique 1 celui qui est très aimé
6. vital

1. douleur
2. minorité
3. permanence 6 passage
4. rédaction 4 texte écrit
5. sagesse 5 bon sens et connaissance
6. transition

1. annuler
2. cultiver
3. défaire 2 travailler la terre
4. mentir 4 ne pas dire la vérité
5. plaider 5 prendre la défense d'une cause
6. siéger

1. commenter
2. engendrer
3. promouvoir 2 faire naître
4. remporter 4 gagner un jeu
5. songer 3 élever à un rang supérieur
6. téléphoner

1. avertir
2. blesser
3. caractériser 3 définir
4. déclencher 1 signaler
5. serrer 4 provoquer un phénomène
6. trahir

La cinquième tranche de mille mots

1. cote
2. fourniture
3. goutte
4. hépatite
5. juridiction
6. moyenne

__6__ le milieu entre deux extrêmes
 __3__ très petite quantité d'un liquide
 __2__ des objets utilisés en salle de classe

1. anormal
2. convaincant
3. éminent
4. insensible
5. mensuel
6. séparatiste

__4__ froid
 __1__ bizarre
 __5__ qui arrive une fois par mois

1. affection
2. chancelier
3. fouet
4. golfe
5. lance
6. qualification

__2__ premier ministre
 __1__ un type de sentiment
 __6__ compétence et expérience

1. amusant
2. boursier
3. irrégulier
4. malin
5. naïf
6. solidaire

__4__ mauvais
 __3__ qui n'est pas uniforme
 __2__ qui concerne le marché financier

1. aviser
2. capter
3. desservir
4. empresser
5. inverser
6. stipuler

__1__ informer
 __5__ changer le sens
 __6__ formuler des conditions précises

1. analogue
2. barbare
3. échéant
4. intact
5. philosophique
6. stupéfiant

__1__ pareil
 __4__ qui n'a pas été touché
 __6__ qui cause une grande surprise

1. brouillard
2. coïncidence
3. farce
4. instituteur
5. pneu
6. soumission

__3__ une histoire qui fait rire
 __1__ ce qui empêche de voir loin
 __4__ un professionnel de l'éducation

1. coffre
2. convoi
3. duc
4. inondation
5. nuance
6. tente

__6__ abri
 __1__ caisse où l'on met de l'argent
 __5__ différence entre deux choses semblables

1. cage
2. panneau
3. repli
4. sou
5. traite
6. vice

__6__ défaut
 __1__ prison
 __4__ monnaie

1. confiner
2. énumérer
3. imputer
4. octroyer
5. pendre
6. venger

__1__ isoler
 __5__ être suspendu
 __2__ dresser une liste

La dixième tranche de mille mots

1. amas
2. dortoir
3. loque
4. huissier
5. nain
6. syntaxe

__6__ mots, phrases, règles
 __2__ chambre à plusieurs lits
 __5__ personne de très petite taille

1. armure
2. charpentier
3. épinette
4. granit
5. hypothèque
6. torsade

__3__ espèce d'arbre
 __1__ vêtement en métal
 __2__ celui qui construit des immeubles

1. cafard
2. étang
3. fenouil
4. ouïe
5. répit
6. sangle

__2__ lac
 __3__ plante
 __5__ pause

1. buée
2. dicton
3. fût
4. garniture
5. maillet
6. piéton

__5__ ce que l'on utilise pour frapper
 __3__ destiné à contenir des liquides
 __6__ homme qui marche dans les rues

1. avortement
2. matrice
3. persévérance
4. ravitaillement
5. recrue
6. tronçon

__5__ jeune soldat
 __6__ morceau coupé d'un objet
 __3__ qualité de celui qui a de la patience

1. astiquer
2. foisonner
3. incarcérer
4. marteler
5. railler
6. tirailler

__3__ enfermer quelqu'un
 __2__ avoir en grande quantité
 __6__ faire hésiter entre deux choses

1. bourrer
2. crouler
3. émanciper
4. mijoter
5. patauger
6. teinter

__1__ remplir
 __2__ détruire ou tomber
 __6__ revêtir d'une couleur

1. culbuter
2. décupler
3. grelotter
4. poncer
5. pulvériser
6. scier

__3__ souffrir du froid
 __6__ couper un arbre
 __5__ briser ou réduire en morceaux

1. âpre
2. fulgurant
3. hébété
4. rusé
5. trapu
6. utopique

__2__ violent ou rapide
 __5__ fort, gros et solide
 __1__ qui est désagréable au toucher

1. ahuri
2. dru
3. nutritif
4. opaque
5. pondéré
6. trépidant

__1__ étonné
 __4__ qui n'est pas transparent
 __6__ très animé ou mouvementé