



# Critérios empiricamente verificáveis para a avaliação do vocabulário na produção oral do professor de língua estrangeira

## Empirically verifiable criteria to assess vocabulary accuracy in the oral production of foreign language teachers

Diego Fernando de OLIVEIRA\*

**RESUMO:** O EPPLE (Exame de Proficiência para Professores de Línguas Estrangeiras) é um instrumento avaliativo em constante aprimoramento (CONSOLO; TEIXEIRA DA SILVA, 2014; COLOMBO, 2019) que se volta para a avaliação da proficiência do professor de LE no contexto brasileiro. Atualmente, o EPPLE possui apenas uma escala holística para a avaliação do desempenho oral de candidatos, desenvolvida intuitivamente. Dessa forma, é necessário que se construa uma escala de proficiência analítica empiricamente embasada para a produção de um argumento para a validade dos critérios avaliativos que classificam os desempenhos orais no exame. A precisão do vocabulário empregado, entre outros construtos da competência linguística, é de extrema importância para a proficiência oral do professor de LE, uma vez que o professor ensina a LE utilizando a língua-alvo em contextos comunicativos (FREEMAN *et al.*, 2015). O presente estudo apresenta resultados da aplicação dos EBBs (*Empirically derived, Binary choice,*

**ABSTRACT:** The EPPLE (Proficiency Exam for Foreign Language Teachers) is an assessment instrument in constant improvement (CONSOLO; TEIXEIRA DA SILVA, 2014; COLOMBO, 2019) which aims at assessing the proficiency of the FL teacher in the Brazilian context. Currently, the EPPLE has only a holistic scale to assess the oral performance of candidates, developed intuitively. Thus, it is necessary to build an analytical and empirically based proficiency scale to produce an argument for the validity of the assessment criteria that classify the oral performances in the exam. The vocabulary accuracy, among other constructs of the linguistic competence, is extremely important for the FL language teacher oral proficiency, since the teacher teaches LE using the target language in communicative contexts (FREEMAN *et al.*, 2015). The present study presents results of the application of EBBs (*Empirically derived, Binary choice, Boundary definition scales*), an empirical methodology for the development of assessment criteria and rating scales

---

\* Mestrando no programa de Pós-Graduação em Estudos Linguísticos da UNESP/IBILCE. Orcid: <https://orcid.org/0000-0002-5121-6237>. Email: [diego.fernando@unesp.br](mailto:diego.fernando@unesp.br).

*Boundary definition scales*), metodologia empírica para o desenvolvimento de critérios avaliativos e escalas de proficiência linguística desenvolvida por Upshur e Turner (1999), em amostras do banco de dados do EPPLE. A partir da análise, construiu-se um quadro de critérios empiricamente verificáveis para a avaliação da precisão do vocabulário empregado pelos candidatos do exame, assim como desenvolveu-se uma escala de proficiência analítica que contempla a coerência do vocabulário empregado, o uso preciso de terminologia técnica e o (re)conhecimento de palavras-chave relacionadas à tarefa metalinguística.

developed by Upshur and Turner (1999), for the construction of a framework of empirically verifiable criteria to assess the vocabulary accuracy used by EPPLE candidates. From data analysis, it was possible to develop an analytical proficiency scale to assess vocabulary accuracy that integrates the coherence of the vocabulary, the accurate use of metalanguage and the knowledge of keywords related to linguistic topics.

**PALAVRAS-CHAVE:** Avaliação docente. Proficiência oral. Língua Estrangeira. Vocabulário.

**KEYWORDS:** Teacher assessment. Oral proficiency. Foreign language. Vocabulary.

## 1 Introdução

O ensino de línguas estrangeiras (doravante LE) no contexto brasileiro apresenta um histórico caracterizado por uma série de fatores desfavoráveis para o desenvolvimento da proficiência linguística entre professores e alunos. No final do século XX, Almeida Filho (1992) apresentava indícios de baixos níveis de proficiência entre professores de LE que atuavam na educação básica no contexto nacional, assim como Nicholls (2001), no início deste século, também discutia aspectos deficitários na produção oral do professor de LE no Brasil. Ainda na mesma década, Sandei (2005) discorria sobre os baixos níveis de proficiência em língua inglesa entre alunos da educação básica brasileira.

Mais recentemente, Rodrigues (2016), Consolo (2017) e Colombo (2019) discutem em seus trabalhos a baixa proficiência oral do professor de LE e questões relacionadas à sua formação nos cursos de Licenciatura em Letras. Para Consolo (2017), por exemplo, é possível observar que, no contexto brasileiro, há uma série de

fatores que atuam para a manutenção de um ciclo de baixo nível de proficiência linguística que afeta tanto o ensino básico quanto o ensino superior - nesse caso, especificamente, os cursos de formação de professores de LEs.

Pensando nesse contexto, o EPPLE (Exame de Proficiência para Professores de Línguas Estrangeiras), instrumento avaliativo específico para professores de LEs, é um exame em desenvolvimento e em constante aprimoramento (CONSOLO; TEIXEIRA DA SILVA, 2014; CONSOLO; AGUENA, 2017; CONSOLO, 2017; COLOMBO, 2019) que visa à coleta de dados acerca do desempenho oral e escrito do professor de LE. Espera-se que esses dados possam orientar futuras reelaborações nas políticas educacionais dos cursos de Licenciatura em Letras e, em um cenário otimista, contribuir para a reversão de um ciclo deficitário para o ensino de LEs no Brasil.

Um elemento importante do construto da proficiência oral do professor e, nesse caso, da competência linguística, é o vocabulário (FULCHER, 2014). A precisão do vocabulário empregado pelo professor de LE é extremamente importante para prover *feedback* aos alunos e abordar problemas linguísticos utilizando metalinguagem, assim como outras áreas funcionais específicas da produção oral docente (ELDER, 2001; FREEMAN et al., 2015). Dessa forma, é importante desenvolver um quadro de critérios válidos para a avaliação do vocabulário na produção oral docente para que se possa obter resultados confiáveis através da aplicação do EPPLE.

Dessa forma, este artigo pretende contribuir com o aprimoramento do construto do EPPLE, sugerindo uma escala de proficiência analítica voltada para a avaliação da precisão do vocabulário empregado na tarefa metalinguística do exame. Além disso, objetiva-se verificar questões levantadas por Colombo (2019) acerca de inconsistências presentes nos descritores da escala de proficiência holística utilizada para a classificação dos desempenhos orais dos candidatos do exame. Por fim, o presente trabalho pretende contribuir com o desenvolvimento de um argumento para a

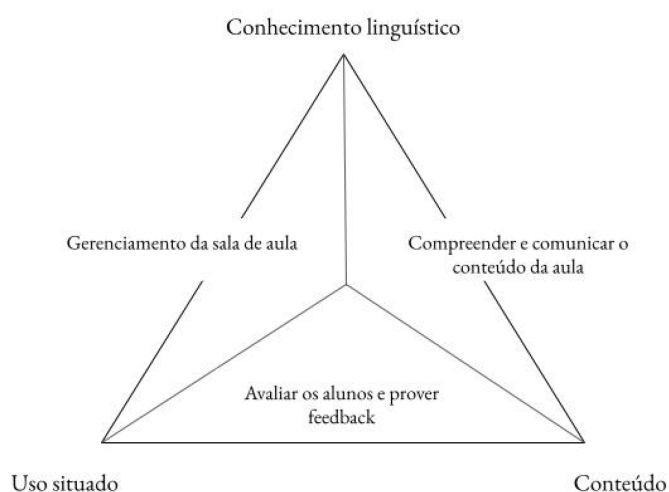
validade do construto do EPPLE, uma vez que visa apresentar um quadro de critérios passível de verificação empírica.

## 2 O EPPLE e a avaliação da proficiência oral do professor de LE

O EPPLE é um exame de proficiência em desenvolvimento que objetiva a avaliação da proficiência oral e escrita de professores de LEs em formação e em atuação (COLOMBO, 2019). Em sua versão oral, ele é dividido em cinco tarefas, sendo a tarefa número cinco responsável por estimular a produção oral para a avaliação do conhecimento específico do candidato acerca de um tópico linguístico, assim como do emprego de metalinguagem, ambas áreas funcionais da proficiência oral intrínsecas à atividade docente (ELDER, 2001; FREEMAN *et al.*, 2015).

A Figura 1, a seguir, apresenta o diagrama desenvolvido por Freeman *et al.* (2015) que ilustra o construto da proficiência oral do professor de LE, assim como as áreas funcionais da produção oral docente.

Figura 1 - Diagrama do construto da proficiência do professor de LE e suas áreas funcionais, baseado em Freeman *et al.* (2015).



Fonte: o autor.

No diagrama, é possível verificar que o construto do professor de LE é composto por um tripé: o conhecimento linguístico necessário para o ensino da LE, o

uso situado que se relaciona ao contexto específico de ensino da LE na língua-alvo e, por fim, o conteúdo que se refere a tópicos linguísticos a serem abordados durante a aula. Dessa forma, é possível verificar que o gerenciamento da sala, a compreensão e comunicação do conteúdo da aula, assim como avaliação e provimento de *feedback* resultam da intersecção das três diferentes dimensões do construto da proficiência oral docente.

Dessa maneira, atentando-se às especificidades da produção oral do professor de LE, a tarefa número cinco do EPPLE apresenta tópicos linguísticos em forma de dúvidas apresentadas por alunos, a fim de que o candidato possa prover explicações, assim como possa empregar metalinguagem para abordar as situações propostas. O Excerto 1, a seguir, apresenta os tópicos linguísticos propostos aos candidatos pelo EPPLE em sua versão 3.1.

Excerto 1 - Tópicos linguísticos propostos na versão 3.1 do EPPLE.

### Situação A

I don't know if these sentences are correct. The first is "If I happen to run into her, I would tell her where you live". The second is "If antibiotic drugs like peniciline were available during World War I, many more lives could be saved". Could you help me?

### Situação B

These sentences are really confusing in this exercise. Which alternative best completes the sentence: "I've written to Maria but she hasn't replied?" Letter A: "She might be away. You never know with her". Letter B: "She must be in Salvador. You never know". How can you explain that?

Fonte: Colombo (2019).

Para Colombo (2019), os critérios que orientam a avaliação do vocabulário empregado pelos candidatos em suas explicações no EPPLE podem ter sua validade questionada, uma vez que permitem a classificação de candidatos que expressam desconhecimento sobre os tópicos linguísticos nas faixas superiores apenas por

empregarem terminologia técnica. Dessa maneira, para a autora, é necessário que se revisem os descritores da escala de proficiência holística do exame, a fim de que a precisão do vocabulário empregado no provimento de explicações e no uso da metalinguagem seja contemplada.

### 3 Critérios para a avaliação do vocabulário

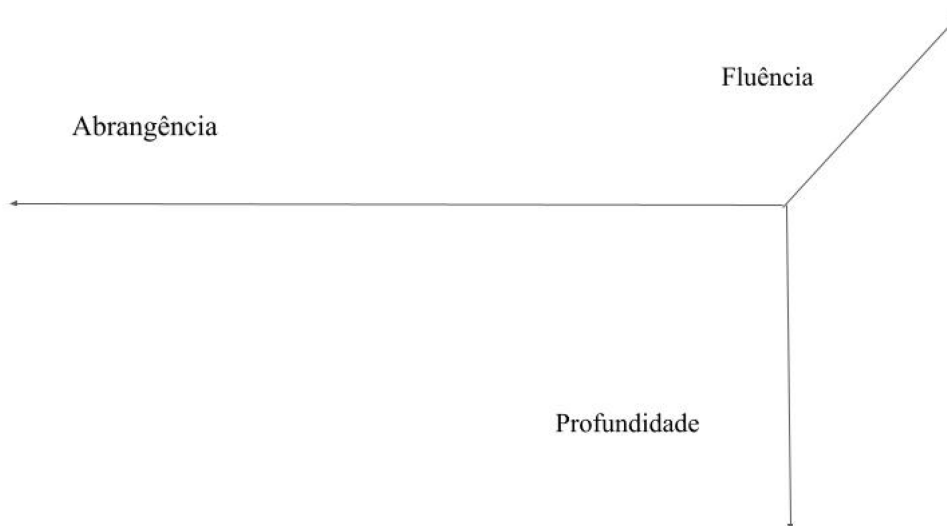
Para a avaliação do vocabulário na produção oral, é necessário delimitar uma área de interesse que melhor defina os aspectos relevantes para o construto que nos interessa avaliar. Primeiramente, é possível dividir vocabulário em dois âmbitos diferentes: abrangência e profundidade<sup>1</sup>. Milton (2009) define abrangência como o número total de palavras que um aprendiz de LE sabe e profundidade como o que o aprendiz sabe sobre essas palavras, ou seja, sobre seu significado, sobre sua adequação e sobre sua ocorrência típica em enunciados. Além disso, segundo o autor, a avaliação do vocabulário pode considerar outra convenção: dividir o conhecimento do vocabulário em conhecimento passivo (ou receptivo) e produtivo (ou ativo) (MILTON, 2009).

A Figura 2, a seguir, ilustra a relação entre a abrangência do vocabulário e a sua respectiva profundidade no desempenho oral. Tal relação também é referida na literatura como espaço lexical<sup>2</sup> (DALLER *et al.*, 2007). Dessa maneira, é possível observar na figura que, quanto maior a abrangência e a profundidade do vocabulário, maior tende a ser a fluência da produção oral em LE.

---

<sup>1</sup> No original: *Breadth e Depth*, Milton (2009).

<sup>2</sup> No original: *Lexical space*, Daller *et al.* (2007).

Figura 2 - Dimensões do conhecimento de vocabulário, baseado em Daller *et al.* (2007).

Fonte: o autor.

Na produção oral, é através do conhecimento receptivo que as palavras são reconhecidas quando ouvidas e através de seu conhecimento produtivo é que palavras podem ser mentalmente evocadas e utilizadas em enunciados. Além de tais delimitações teóricas para se avaliar e investigar o vocabulário dentro do construto da proficiência oral, é preciso considerar, ainda, a definição de palavra em si - principalmente nos casos em que se deseja investigar a abrangência do vocabulário. Para Milton (2009),

Não há uma única, simples definição de uma palavra que é utilizada na criação de testes para mensurar conhecimento de vocabulário e aprendizagem. Investigadores podem mudar a unidade de mensuração de acordo com as circunstâncias e os aprendizes que eles estão investigando<sup>3</sup> (MILTON, 2009, p. 12).

---

<sup>3</sup> No original: There is no single, simple definition of a word that is used in the creation of tests to measure vocabulary knowledge and learning. Investigators can change the unit of measurement according to the circumstances and the learners they are investigating (MILTON, 2009, p. 12).

Considerando a necessidade de se estabelecer recortes específicos os quais devem contemplar qual aspecto do construto da proficiência oral e do conhecimento do vocabulário se tem o interesse de investigar, o Quadro 1, a seguir, desenvolvido por Nation (2001), apresenta possíveis definições operacionais para a avaliação do vocabulário na produção oral, considerando tanto o conhecimento quanto o conhecimento passivo.

No Quadro 1, divide-se o construto do conhecimento de vocabulário em três diferentes áreas de interesse: (i) Forma, que se relaciona à estrutura do vocabulário em si, como ele se parece em forma escrita ou como ele soa; (ii) Significado, que se relaciona à identificação de uma forma e sua associação a um ou mais significados; (iii) Uso, que se define pela relação que uma palavra estabelece com outras em um enunciado. Tais divisões possuem subcategorias para a língua escrita e língua falada e, por fim, são fracionadas em conhecimento receptivo (R) e conhecimento produtivo (P).

Quadro 1 - Possíveis critérios para a avaliação do vocabulário, segundo Nation (2001).

|             |                        |   |  |
|-------------|------------------------|---|--|
| Forma       | Falado                 | R | Como a palavra soa?  |
|             |                        | P | Como a palavra é pronunciada?  |
|             | Escrito                | R | Como a palavra se parece?  |
|             |                        | P | Como a palavra é escrita?  |
|             | Partes das palavras    | R | Quais partes são reconhecíveis nessa palavra?                          |
|             |                        | P | Quais partes das palavras são necessárias para expressar significado?  |
| Significado | Forma e significado    | R | Qual o significado indicado / sugerido pela forma desta palavra?       |
|             |                        | P | Qual forma lexical pode ser utilizada para expressar esse significado? |
|             | Conceitos e referentes | R | O que é incluído nesse conceito?                                       |



|     |                     |   |  |
|-----|---------------------|---|--|
|     | Associações         | P | A quais itens o conceito pode se referir?                            |
|     |                     | R | Em quais outras palavras essa palavra nos faz pensar?                |
|     |                     | P | Quais outras palavras podemos usar ao invés dessa?                   |
| Uso | Funções gramaticais | R | Em quais padrões a palavra ocorre?                                   |
|     |                     | P | Em quais padrões nós devemos usar essa palavra?                      |
|     | Colocações          | R | Quais palavras ou tipos de palavras ocorrem com essa?                |
|     |                     | P | Quais palavras ou tipos de palavras nós devemos usar com essa?       |
|     | Restrições de uso   | R | Onde, quando e com que frequência deveríamos encontrar essa palavra? |
|     |                     | P | Onde, quando e com que frequência podemos usar essa palavra?         |

Fonte: Nation (2001, p. 27).

Dessa forma, Nation (2001) apresenta um quadro de possíveis definições operacionais que auxilia avaliadores, pesquisadores e construtores de testes a delimitar o foco investigativo e os aspectos-chave para a avaliação do conhecimento de vocabulário na produção oral. Além de tais especificações, Read (2000) destaca três dimensões da avaliação do vocabulário que devem ser levadas em consideração, uma vez que são importantes para a definição operacional de um teste ou exame, assim como para a validade de seus critérios avaliativos. O Quadro 2, a seguir, apresenta as três dimensões da avaliação do vocabulário, segundo Read (2000).

Quadro 2 - Dimensões da avaliação do vocabulário.

|  |  |
|--|--|
| <p><b>Discreto</b><br/>Mensura o conhecimento ou uso do vocabulário como um construto independente.</p>  | <p><b>Integrado</b><br/>Mensura o vocabulário como um componente de um construto maior.</p>  |
| <p><b>Seletivo</b><br/>Mensura itens específicos do vocabulário, os quais são o foco da avaliação.</p>   | <p><b>Compreensivo</b><br/>Mensura o vocabulário considerando todo o conteúdo presente no <i>input</i> (tarefas de leitura/compreensão auditiva) ou na resposta do candidato (escrever/falar).</p> |
| <p><b>Independente do contexto</b><br/>Uma mensuração do vocabulário que considera a resposta do candidato que não se refere a um contexto específico.</p> | <p><b>Dependente do contexto</b><br/>Uma mensuração do vocabulário que avalia a habilidade do candidato de considerar informação contextual para a produção da resposta.</p>                       |

Fonte: Read (2000, p. 9).

Na primeira dimensão, o vocabulário pode ser considerado parte de um construto maior ou ser avaliado como um construto independente, separado de outras competências linguísticas. Na segunda dimensão, leva-se em conta a abrangência do número de palavras contempladas na avaliação do vocabulário, ou seja, se há delimitação de algumas palavras de interesse do avaliador ou se todas as palavras produzidas serão consideradas na avaliação. Por fim, na terceira dimensão, o vocabulário pode ser avaliado como independente ou dependente de um contexto avaliativo específico.

Por fim, é importante mencionar os critérios presentes no CEFR (*Common European Framework*) para a avaliação do vocabulário. Divididas em diferentes categorias, critérios para a avaliação da abrangência do vocabulário e de sua precisão podem ser encontradas nas escalas *Linguistic Range*, *Vocabulary Range* e *Vocabulary Control*. O Quadro 3, a seguir, apresenta um resumo dos critérios utilizados para a classificação dos desempenhos satisfatórios e insatisfatórios nas escalas de proficiência linguística mencionadas.

Quadro 3 - Critérios para a avaliação do vocabulário no CEFR.

| Escala de proficiência | Critérios   |
|------------------------|---|
| Linguistic Range       | Satisfatórios: O candidato possui amplo vocabulário empregado precisamente, assim como pode diferenciar e desambiguar palavras. Não demonstra ter limitações de vocabulário.  |
|                        | Insatisfatórios: O candidato possui vocabulário suficiente para se expressar em contextos específicos. Demonstra limitações de vocabulário, apresenta repetição de palavras ao tratar de temas comuns, assim como busca mentalmente por palavras, expressões básicas do dia a dia etc.                |
| Vocabulary Range       | Satisfatório: O candidato possui amplo repertório, expressões idiomáticas, coloquialismo, conotação. Supre falta de vocabulário suprido por paráfrases. Faz poucas buscas mentais por palavras. Conhece sinônimo de palavras pouco comuns. Usa amplo vocabulário técnico de sua área de conhecimento. |
|                        | Insatisfatório: Evita repetições frequentes, há hesitação e paráfrase na falta de uma palavra. Tem problemas com termos técnicos fora de sua área. Vocabulário necessário para tratar de temas do dia a dia.  |
| Vocabulary Control     | Satisfatório: Uso apropriado de palavras menos comuns (às vezes com erros não significativos), ocorrem escolhas incorretas sem que a compreensão seja afetada.  |
|                        | Insatisfatório: Comete erros mais graves ao falar de temas pouco familiares, mais complexos. Controla vocabulário simples que se refere ao dia a dia.   |

Fonte: o autor.

É importante ressaltar que o quadro fornece apenas um resumo dos critérios presentes nas escalas do *CEFR*, os quais podem ser consultados integralmente no *website* do Conselho Europeu<sup>4</sup>. Além disso, foram categorizados como satisfatórios os

<sup>4</sup> Disponível em: <https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4>

critérios presentes nas faixas mais altas das escalas de proficiência, assim como foram categorizados como insatisfatórios os critérios presentes nas faixas mais baixas das escalas mencionadas. Dessa forma, é possível ter uma visão mais global dos critérios utilizados para a avaliação do vocabulário na produção oral em exames de proficiência, uma vez que o *CEFR* é uma referência importante para a avaliação de proficiência oral e escrita em LE.

#### 4 Procedimentos metodológicos

Para a análise dos dados, uma coleção de vinte e oito amostras de desempenho oral oriundas do banco de dados do EPPLE foi utilizada, compreendendo amostras coletadas nos anos de 2015 e 2017 em universidades públicas do estado de São Paulo e Minas Gerais<sup>5</sup>. As amostras consistem em gravações de produções orais de licenciandos em Letras que realizaram o EPPLE. Além disso, para este trabalho, foram consideradas apenas as respostas providas à tarefa metalinguística do exame, uma vez que a metodologia empregada produz critérios estritamente relacionados a uma única tarefa (FULCHER *et al.*, 2011).

É importante destacar que as amostras de desempenho oral foram transcritas com a utilização do *software* CLAN<sup>6</sup>. O CLAN permite a transcrição de gravações em formato de diálogos e possui códigos próprios para codificar diferentes fenômenos da produção oral. O Anexo 1, ao fim deste artigo, apresenta os códigos empregados nas transcrições dos desempenhos orais e seus respectivos significados.

Os *EBBs* (*Empirically derived, Binary choice, Boundary definition scales*) é uma metodologia empírica para o desenvolvimento de critérios avaliativos e escalas de proficiência linguística desenvolvida por Upshur e Turner (1999) que consiste no

---

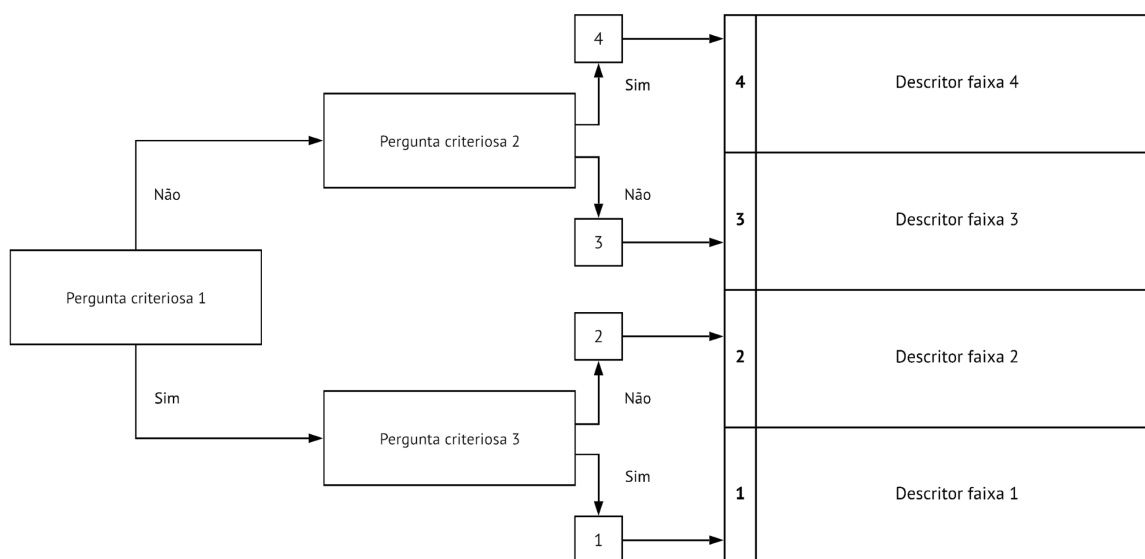
<sup>5</sup> Número do protocolo do processo do Comitê de Ética da UNESP/IBILCE: 0007.0.229.325-11.

<sup>6</sup> Disponível em <https://talkbank.org/software/>

desenvolvimento de Árvores de Decisão de Desempenho<sup>7</sup>. Através de uma série de perguntas binárias (sim/não), desenvolve-se uma Árvore de Decisão de Desempenho que irá compor, ao final do processo, uma escala de proficiência linguística empiricamente verificável que estabelece limites precisos entre as diferentes faixas de proficiência produzidas.

A Figura 3, a seguir, ilustra o desenvolvimento de uma Árvore de Decisão de Desempenho e de uma escala de proficiência linguística de quatro níveis a partir da aplicação dos *EBBs*.

Figura 3 - Esquema de desenvolvimento de Árvores de Decisão de Desempenho e escalas de proficiência linguística através dos *EBBs*.



Fonte: o autor.

Dessa forma, como é possível observar, a Pergunta Criteriosa 1 estabelece os limites entre os dois níveis mais altos e mais baixos da escala - ou seja, classifica os desempenhos mais e menos satisfatórios observados nos dados. A Pergunta Criteriosa 2 divide os desempenhos satisfatórios, caracterizando os desempenhos mais proficientes entre as amostras analisadas. O mesmo acontece com a Pergunta

<sup>7</sup> No original: Performance Decision Trees (FULCHER *et al.*, 2011).

Críteriosa 3, a qual distingue os desempenhos menos satisfatórios e caracteriza o desempenho menos proficiente observado nas amostras analisadas.

## 5 Análise de dados

A partir da análise das amostras de desempenho oral oriundas do banco de dados do EPPLE, foi possível desenvolver um quadro de critérios empiricamente verificáveis que se relacionam tanto ao conhecimento ativo quanto ao conhecimento passivo do vocabulário. Além disso, os critérios visam, especificamente, a avaliação da profundidade do vocabulário utilizado pelos candidatos ao tratar do problema proposto pela tarefa metalinguística do exame. As delimitações realizadas seguem orientações de Milton (2009) para o estabelecimento de um foco avaliativo. Dessa forma, objetivou-se desenvolver uma escala que avalia a precisão do vocabulário utilizado pelos candidatos em sua produção oral.

Além disso, considerando as dimensões da avaliação do vocabulário de Read (2000), é possível dizer que o construto da precisão do vocabulário é integrado a um construto maior da proficiência oral do professor de LE - a expressão de conhecimento linguístico, uma vez que a escolha das palavras realizada pelo candidato reflete a sua habilidade de explicar a LE na língua-alvo (FREEMAN *et al.*, 2015). Além disso, o construto operacionalizado é compreensivo, pois considera o *input* presente na tarefa metalinguística, assim como é dependente do contexto, uma vez que a produção oral se relaciona a uma tarefa que simula uma situação real de interação entre professor e aluno.

Para a classificação dos desempenhos mais e menos proficientes, foi tomado como critério principal o estabelecimento de relações coerentes entre terminologia técnica e/ou palavras referentes ao problema linguístico presentes na produção oral dos candidatos. Conforme apontado por Colombo (2019), parte das amostras de desempenho não apresentam emprego de metalinguagem, porém contém uma série

de palavras e referentes os quais se relacionam ao problema linguístico proposto pela tarefa, sendo possível, dessa forma, verificar a coerência da explicação por meio da análise da precisão das palavras utilizadas pelos candidatos.

O Excerto 2, a seguir, apresenta a transcrição da amostra de desempenho oral do candidato TAC, em que se podem observar associações incoerentes entre a terminologia técnica empregada e as demais palavras escolhidas pelo candidato para abordar o problema linguístico proposto pela tarefa. Nota-se que, dessa forma, a precisão do vocabulário reflete um construto maior: o conhecimento do candidato acerca do tópico linguístico abordado. Dessa forma, evita-se que candidatos que expressam desconhecimento linguístico, mas empregam terminologia técnica, sejam classificados nas faixas mais altas da escala - uma possível solução para a questão levantada por Colombo (2019).

Excerto 2 - Desempenho oral do candidato TAC.

- \*RIC: let's go .
- \*RIC: on [: in] [\*] situation B (.) the best alternative is letter B .
- \*RIC: you can n(o)t (.) use the alternative A (.) because (.) the (.) verb tenses in the second sentence is wrong .
- \*RIC: if you (.) use "she might be away" (.) you have to change the tense "you: never know xxx her" to: another thing to [/] to say .
- \*RIC: this is completely unsense [//] (.) nonsense .
- \*RIC: but (.) alternative B is correct .
- \*RIC: if you (.) really wanted to use the alternative A (.) &yo you could (.) say: .
- \*RIC: she might be away (.) you [/] you never know about that .

Fonte: o autor.

À vista disso, por se tratar de uma tarefa metalinguística, o emprego preciso de terminologia técnica foi tomado como critério para delimitar as faixas superiores da escala de proficiência linguística. Uma vez que a tarefa número cinco do EPPLA aborda um problema linguístico e também orienta o candidato a empregar metalinguagem, o uso preciso de terminologia técnica sugere que o candidato possui níveis de

proficiência mais altos. Além disso, incorporar o emprego de metalinguagem no quadro de critérios avaliativos também reflete o construto do conhecimento específico necessário ao professor de LE para ensinar a língua-alvo (ELDER, 2001; FREEMAN et al., 2015).

O Excerto 3, a seguir, apresenta a transcrição do desempenho oral do candidato LAC, que emprega terminologia técnica de maneira precisa, assim como estabelece uma relação coerente entre os termos técnicos utilizados e seus referentes para abordar o problema linguístico proposto pela tarefa metalinguística.

Excerto 3 - Desempenho oral do candidato LAC.

- \*LAC: so (.) & i the sentence you would like to correct is .
- \*LAC: "if I happen to run into her I would tell her where you live" .
- \*LAC: right ?
- \*LAC: the problem you (a)re having here (.) is: basically just (.) using the correct verb form .
- \*LAC: so if you say "if I happen" (.) you (a)re using the: present form .
- \*LAC: and it refers to future .
- \*LAC: so you have to say "I will tell her where you live" .
- \*LAC: becau:se you can n(o)t (.) mix "if I happen" with "would" .
- \*LAC: if you prefer to use would conversely you can also say .
- \*LAC: "if I happened to run into her I would tell her where you live" .
- \*LAC: but then there is a slight difference (.) in that .
- \*LAC: when you say "if I happen" <it (i)s &m> [//] it (i)s something that is more likely to occur .
- \*LAC: and if you say "if I happened" then: the likelihood is a bit lower .
- \*LAC: so that (i)s the basic difference .

Fonte: o autor.

Para os níveis inferiores, além da incoerência entre as palavras escolhidas para abordar o tópico linguístico proposto, a habilidade de reconhecer palavras-chave importantes relacionadas ao problema gramatical foi tomada como critério limitador entre as faixas mais baixas da escala. Foi possível observar que um número expressivo de candidatos não demonstrou conhecimento acerca das palavras presentes nos



tópicos propostos, assim como desconheciam *phrasal verbs* empregados como distratores, provendo uma explicação incorreta e pouco satisfatória em vista do objetivo da tarefa.

O Excerto 4, a seguir, apresenta a transcrição do desempenho oral do candidato TAC, em que o examinando emprega corretamente terminologia técnica, assim como identifica palavras-chave necessárias para abordar o problema linguístico proposto, porém expressa desconhecimento acerca do tópico, provendo uma explicação imprecisa e, como consequência, teve seu desempenho classificado como insatisfatório.

Excerto 4 - Desempenho oral do candidato TAC.

- \*TAC: so <the &st> [//] &-um student pay attention to situation A .
- \*TAC: “if I happen to run into her I will tell her where you live” .
- \*TAC: pay attention that <we use> [///] it is a conditional sentence .
- \*TAC: &-um <and we &ha> [///] it is a second conditional in which we [//] you have to use +...
- \*TAC: &-um in the first part you have used &-um (.) present .
- \*TAC: but you do need to use what ?
- \*TAC: past .
- \*TAC: right ?
- \*TAC: so “if I happened to run into her I will tell her where you live” .
- \*TAC: sometimes it (i)s difficult to &-um remember all the rules .
- \*TAC: but do n(o)t forget this .
- \*TAC: whenever 0you use “would”
- \*TAC: right ?
- \*TAC: &-um it is (.) the action trying to understand what the action is .
- \*TAC: and &-um (.) you see that it is a possibility .
- \*TAC: okay ?
- \*TAC: so it (i)s not unreal .
- \*TAC: okay ?
- \*TAC: so the: [/] the: [/] the: verb +/.

Fonte: o autor.

Os candidatos que expressaram desconhecimento de vocabulário em sua produção oral (conhecimento ativo), como também na interpretação da tarefa (conhecimento passivo), foram classificados na faixa mais baixa da escala, por apresentarem um desempenho extremamente insatisfatório. O Excerto 5, a seguir, apresenta a transcrição do desempenho oral do candidato LID que, ao reconhecer erroneamente a sentença *you never know with her* como agramatical, desenvolve uma explicação incoerente, imprecisa e que expressa desconhecimento em relação ao tópico abordado pela tarefa metalinguística.

Excerto 5 - Desempenho oral do candidato LID.

\*LID: so: (.) I chose [: choose] [\*] situation B .  
 \*LID: &-um what do you have to do here ?  
 \*LID: read letter A and then letter B .  
 \*LID: "she might be away (.) you never know with her" .  
 \*LID: in this situation you are using the wrong preposition .  
 \*LID: so "with" and "her" is [: are] [\*] not in the context .  
 \*LID: you never know (.) about her .  
 \*LID: you would never know about her .  
 \*LID: you should use (.) another expression (.) to: (.) mean what you  
 (.) want to day .  
 \*LID: letter B is better .  
 \*LID: "she must be in Salvador (.) you never know" .  
 \*LID: "you never know" because you do not know where (.) this person is  
 where Mary is .  
 \*LID: so: (.) in letter A (.) "you never know with her" (.) the  
 preposition (.) is being used wrongly .  
 \*LID: while in (.) letter B (.) "you never know" (.) is better .  
 \*LID: because you do not have a preposition .  
 \*LID: and you do not have (.) an undirect [: indirect]  
 [\*] object in your way .

Fonte: o autor.

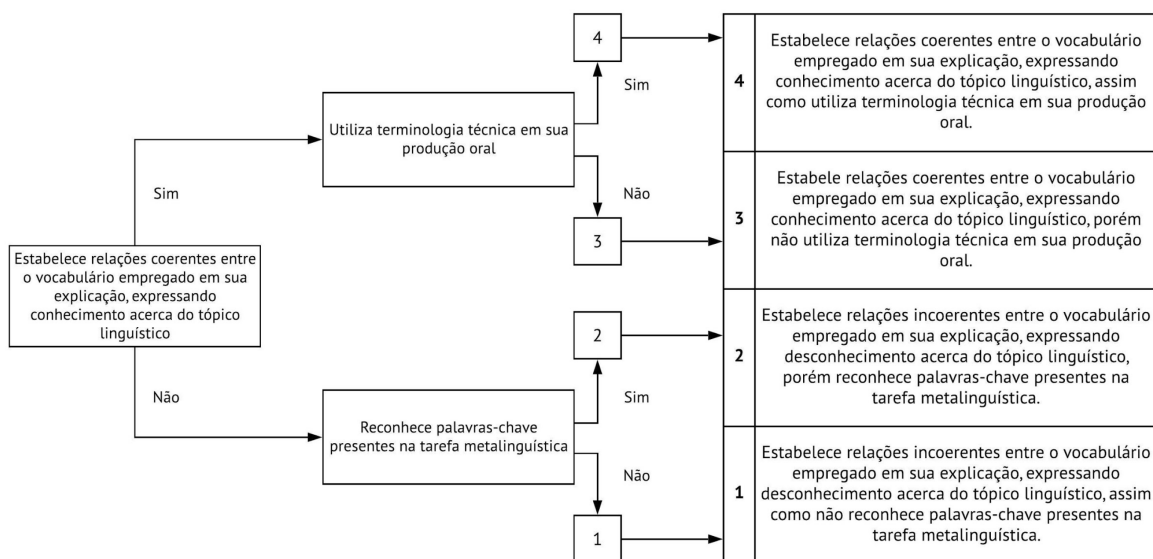
Dessa maneira, é possível dizer que o estabelecimento de uma relação de coerência entre as palavras empregadas na explicação dos candidatos é o que determina, em primeiro lugar, o seu grau de precisão. Além disso, os desempenhos

observados e classificados como mais precisos também apresentaram emprego adequado de metalinguagem, o que caracteriza o desempenho mais proficiente entre as amostras analisadas, por serem extremamente precisos quanto ao uso de terminologia técnica e expressarem conhecimento de vocabulário.

Nos desempenhos menos satisfatórios, ou seja, aqueles que apresentam incoerências entre as palavras empregadas para o provimento da explicação, o reconhecimento de palavras-chave para a resolução do problema gramatical foi tomado como critério que limita as faixas mais baixas da escala e que caracteriza o desempenho menos proficiente entre as amostras analisadas, por serem extremamente imprecisos quanto ao uso de terminologia técnica e expressarem desconhecimento de vocabulário.

A Figura 4, a seguir, ilustra a Árvore de Decisão de Desempenho e a escala de proficiência linguística desenvolvida para a avaliação da precisão do vocabulário com base em amostras de desempenho oral de candidatos do EPPLE.

Figura 4 - Árvore de Decisão de Desempenho e escala de proficiência linguística para a avaliação da precisão do vocabulário.



Fonte: o autor.

## 6 Considerações finais

A partir da aplicação dos *EBBs*, foi possível desenvolver uma escala de proficiência linguística analítica para a avaliação da precisão do vocabulário empregado pelos candidatos do EPPL, assim como a construção de um quadro de critérios empiricamente verificáveis os quais constituem um argumento para a validade do construto operacionalizado. Além disso, procurou-se contemplar os questionamentos levantados por Colombo (2019) em relação à escala de proficiência holística existente no EPPL e apontamentos de Consolo e Teixeira da Silva (2014) acerca de aprimoramentos necessários ao exame, como, por exemplo, o desenvolvimento de uma escala analítica empiricamente embasada.

A escala de proficiência linguística produzida para a avaliação do vocabulário foca a profundidade do vocabulário do candidato (MILTON, 2009) e operacionaliza a precisão do vocabulário como dependente do contexto, integrado ao construto do conhecimento linguístico do professor de LE. Além disso, a escala caracteriza-se como compreensiva, uma vez que o desempenho dos candidatos é estritamente relacionado à tarefa metalinguística. Dessa forma, considerando as categorias avaliativas propostas por Read (2001), os critérios avaliativos contemplam o conhecimento de vocabulário ativo e passivo, os quais estão ligados tanto à compreensão do significado das palavras quanto ao seu uso na produção oral.

Foi possível observar que o estabelecimento de relações coerentes entre as palavras empregadas pelos candidatos em sua produção oral é o que caracteriza os desempenhos mais e menos proficientes. Para limitar os desempenhos mais proficientes, o emprego de terminologia técnica associada à expressão de conhecimento linguístico foi o critério tomado como definidor entre as fronteiras dos níveis superiores da escala. Para limitar os desempenhos menos proficientes, o desconhecimento acerca do significado de palavras-chave para abordar os tópicos

linguísticos propostos pelo exame foi selecionado como critério limitador das faixas mais baixas da escala.

Sugere-se que investigações sejam realizadas para a compreensão de outras dimensões avaliativas do vocabulário empregado pelo professor de LE, como a sua abrangência, por exemplo. Além disso, é preciso que outros avaliadores/especialistas em avaliação empreguem a escala de proficiência produzida por este estudo em amostras de desempenho oral para que correlações possam ser estabelecidas entre os resultados obtidos e possíveis discrepâncias sejam identificadas. Tais procedimentos metodológicos assegurariam a confiabilidade dos resultados alcançados através desta pesquisa, assim como contribuiriam como argumento para a validade do construto operacionalizado na escala de proficiência analítica.

## Referências

ALMEIDA FILHO, J. C. O professor de língua estrangeira sabe a língua que ensina? A questão da instrumentalização lingüística. **Contexturas**. Ensino Crítico de Língua Inglesa. UNICAMP, 1992.

COLOMBO, C. S. **Avaliação de Proficiência Oral em Língua Estrangeira: Uma Proposta de Abordagem Avaliativa em um Exame para Professores de Línguas com base em Métodos de Estruturação de Problemas e em Métodos Multicritério de Análise da Decisão**. 2019. Tese (Doutorado em Estudos Linguísticos) - Universidade Estadual Paulista "Júlio de Mesquita Filho". São José do Rio Preto/SP, 2019.

CONSOLO, D. A.; TEIXEIRA DA SILVA, V. L. Em defesa de uma formação linguística de qualidade para professores de línguas estrangeiras: o exame EPPL. **Revista Horizontes de Linguística Aplicada**, v. 13, 2014. DOI <https://doi.org/10.26512/rhla.v13i1.1334>

CONSOLO, D. A; AGUENA, D. M. Pré-testagem ao Exame de Proficiência para professores de língua estrangeira (EPPL): desenho e desafios tecnológicos para avaliação de proficiência oral em dispositivos tecnológicos móveis. **Revista Polifonia**, v. 24 n. 32/5, 2017.

CONSOLO, D. A. Assessing EFL teachers' proficiency: on the development of teacher programmes and testing policies in Brazil. *In: DOUGLAS A. C. ET AL (org). **Perspectivas em avaliação no ensino e na aprendizagem de línguas: pesquisas e encaminhamentos na formação e na prática docente.** Pontes Editores, Campinas, 2017.*

COUNCIL OF EUROPE. **Common European Framework of Reference for Languages, Learning, Teaching and Assessment.** Strasbourg: Council of Europe Publishing. 2020. Disponível em [www.coe.int/lang-cefr](http://www.coe.int/lang-cefr), acesso em 17/02/2021.

DALLER, H.; MILTON, J.; TREFFERS-DALLER, J. Conventions and terminology. *In: DALLER, H.; MILTON, J.; TREFFERS-DALLER, J. (ed.). **Modelling and Assessing Vocabulary Knowledge.** Cambridge: Cambridge Applied Linguistics. 2007. DOI <https://doi.org/10.1017/CBO9780511667268>*

ELDER, C. Assessing the language proficiency of teachers: are there any border controls? **Language Testing**, 2001, p. 149-170. DOI <https://doi.org/10.1177/026553220101800203>

FREEMAN, D.; KATZ, A.; GOMEZ; P. G.; BURNS, A. English-for-teaching: rethinking teacher proficiency in the classroom. **Oxford: ELT Journal**, Oxford University Press, 2015. DOI <https://doi.org/10.1093/elt/ccu074>

FULCHER, G.; DAVIDSON F.; KEMP, J. Effective rating scale development for speaking tests: Performance decision trees. **Language Testing**, 2011. DOI <https://doi.org/10.1177/0265532209359514>

FULCHER, G. Defining the construct. *In: **Testing second language speaking.** New York: Routledge, 2014. DOI <https://doi.org/10.4324/9781315837376>*

MILTON, J. **Measuring Second Language Vocabulary Acquisition.** Dublin: Trinity College. 2009. DOI <https://doi.org/10.21832/9781847692092>

NATION, I. S. P. **Learning vocabulary in another language.** Cambridge: Cambridge University Press, 2001. DOI <https://doi.org/10.1017/CBO9781139524759>

NICHOLLS, S. M. **Aspectos pedagógicos e metodológicos do ensino de inglês.** Maceió: EdUFAL, 2001.

READ, J. **Assessing vocabulary.** Cambridge: Cambridge University Press, 2000. DOI <https://doi.org/10.1017/CBO9780511732942>

RODRIGUES, L. C. B. A formação do professor de língua estrangeira no século XXI: Entre as antigas pressões e os novos desafios. **Revista Signum**. 2016. DOI <https://doi.org/10.5433/2237-4876.2016v19n2p13>

SANDEI, M. L. R. **Prospecções sobre a abordagem/senso de plausibilidade na ação de ensinar de um professor de língua estrangeira no ensino médio**. 2005. Dissertação (Mestrado em Linguística Aplicada) – Universidade Estadual de Campinas, Campinas.

UPSHUR J. A; TURNER, C. E. Systematic effects in the rating of second-language speaking ability: test method and learner discourse. **Language Testing**. 1999. DOI <https://doi.org/10.1177/026553229901600105>

#### Anexo 1 - Códigos empregados na transcrição das amostras de desempenho oral.

| CÓDIGO               | CARACTERÍSTICAS DE FALA  |
|----------------------|--|
| *ABC: bla            | Fala do interagente  |
| bla .                | Declaração completa  |
| bla +...             | Declaração incompleta (sem interrupção)  |
| bla +/.              | Declaração interrompida  |
| bla !                | Exclamação completa  |
| bla ?                | Pergunta completa  |
| bla +...?            | Pergunta incompleta (sem interrupção)  |
| bla +/?              | Pergunta interrompida  |
| + , bla              | Auto-preenchimento (após interrupção)  |
| ++ bla               | Preenchimento do enunciado de outro falante  |
| "bla"                | Citação de pequenos trechos  |
| bla +/?.             | Citação longa iniciada na próxima trilha ( <i>tier</i> )   |
| +* bla               | Início de citação longa  |
| goed [ ] [: went]    | Erro estrutural seguido de correção pela transcrição   |
| piecee [: peach] [ ] | Erro de escolha lexical seguido de correção pela transcrição   |
| <bla> [ ] bla        | Repetição de palavras ou trechos de fala   |
| <bla> [ ] bla bla    | Reconstituição de palavras ou trechos de fala  |
| <bla> [ ] bla bla    | Falso início sem reconstituição  |
| <bla> [ ] cha        | Reformulação completa da mensagem  |
| bla [ ] [ ] bla      | Erro no trecho de fala a ser reparado  |
| bla [ ] [ ] bla      | Erro no trecho de fala de reparo   |
| 0bla                 | Palavra omitida  |
| bl(a)                | Encurtamento fonológico  |
| bla:bla              | Alongamento de som silábico  |
| blabla:              | Alongamento de palavra   |
| &bl                  | Fragmentos fonológicos   |
| &-um                 | Pausas preenchidas (preenchedores não lexicais)  |
| bla*bla              | Pausa dentro de uma palavra  |
| (.)                  | Pausa silenciosa (sem duração)   |
| (0.6)                | Pausa silenciosa (com duração em segundos)   |
| xxx                  | Discurso incompreensível   |
| bla [=  tosse]       | Vocalizações não verbais (ocorrência paralinguística) com ocorrência linguística na trilha ( <i>tier</i> ) |
| 0 [=  tosse]         | Vocalizações não verbais (ocorrência paralinguística) sem ocorrência linguística na trilha ( <i>tier</i> ) |
| bla ↑                | Entonação ascendente   |
| bla ↓                | Entonação descendente  |
| bla &*ABC: bla bla   | Sobreposição de discurso (comentários curtos)  |
| <bla> >]             | Sobreposição de discurso (falas simultâneas)   |
| <bla> [ <            | Sobreposição de discurso (falas simultâneas)   |
| bla@s:por            | Palavra em língua portuguesa   |
| bla@q                | Uso de metalinguagem   |
| %p:ho: bla           | Desvio de pronúncia  |
| bla'bla              | Sílabas enfatizadas  |
| bla <bla bla> [ ]    | Trecho enfatizado  |
| %com: bla            | Comentários  |
| bla ‡ bla            | Agrupamento de marcadores conversacionais (posição inicial): vocativos                                     |
| bla .. bla           | Agrupamento de marcadores conversacionais (posição final): vocativos, perguntas curtas                     |
| bla [= bla]          | Identidade dêitica   |
| bla +.               | Pausa na transcrição   |
| bla [^c]             | Delimita orações dentro de sentenças complexas   |

Fonte: Colombo (2019).

Artigo recebido em: 23.02.2021

Artigo aprovado em: 04.03.2021