




Avaliação e tradução: considerações sobre a análise computadorizada de erros

Assessment and translation: Reflections on computer-assisted error analysis

Júlia Cristina Valverde da Silva

Universidade de Brasília

Brasília, Distrito Federal, Brasil

juliavalverdesilva@outlook.com 

<https://orcid.org/0000-0002-6605-6619> 

Resumo: Perdurou por muito tempo na área da tradução métodos avaliativos pouco criteriosos, baseados em níveis de correspondência com o texto de partida e pautados por princípios genéricos e não definidos de fidelidade e equivalência. Na avaliação de traduções, no âmbito de ensino, a conferência magistral, isto é, a versão autorizada do professor permanece sendo usada como parâmetro de aceitabilidade dos textos produzidos por estudantes. Nesse contexto, torna-se imperativo o desenvolvimento de critérios e métodos avaliativos mais claros e que sirvam como retroação e como guia para práticas pedagógicas orientadas à necessidade de aprendizagem real de discentes. Com isso em mente, este artigo visa a descrever e considerar a aplicabilidade de uma metodologia componencial a processos avaliativos mais objetivos: a análise computadorizada de erros. Com a descrição de um sistema de anotação de erros, o *Translation-oriented Annotation system*, pretende-se demonstrar como a análise de erros pode ser integrada ao processo avaliativo no ensino de tradução, fornecendo subsídios para adaptações/ reelaborações de planos de ensino e de materiais didáticos que supram as dificuldades relativas à produção linguística de estudantes de tradução.

Palavras-chave: análise de erros; linguística de corpus; corpus de aprendiz; avaliação em tradução; sistema de anotação de erros.

Abstract: For many years, loosely defined assessment methods, based on levels of correspondence between source and target texts and on vague principles of fidelity and equivalence, have prevailed in the field of translation. In many translation education settings, translation assessment continues to rely on the expert's authorized version; that is, the instructor's version still serves as the



acceptability standard by which students' translations are evaluated. Within this context, the need arises to establish clearer and better-defined assessment criteria and methods that can be used both as a basis for feedback and as guidelines for needs-based pedagogical practices. This article seeks to describe and explore the applicability of a componential methodology in more objective assessment processes: computer-assisted error analysis. Through a detailed description of the Translation-oriented Annotation System, the paper demonstrates how error analysis can be integrated into the assessment process in translation education, providing resources for adapting or redesigning curricula and learning materials to address the specific challenges faced by translation students in their linguistic production.

Keywords: error analysis; error taxonomy; corpus linguistics; learner corpus; translation assessment; error annotation systems.

I. Introdução

A avaliação de traduções pode adotar diversas facetas a depender do objeto em crivo. Por muitos anos, a avaliação em tradução equivalia à crítica de traduções literárias, conduzidas predominantemente seguindo critérios estilísticos e subjetivos (Martínez Melis & Hurtado Albir, 2002). Em sua incipiência, a avaliação de traduções consistia, assim, em glosas e comentários sobre obras literárias, como romances, ensaios e poesia, os quais não observavam nenhum critério objetivo (Secară, 2005), limitando-se a discussões sobre o grau de fidelidade e equivalência com o original. Esses procedimentos conjecturais, contudo, começaram a dar lugar a critérios mais objetivos de avaliação quando a área dos Estudos da Tradução se tornou uma disciplina autônoma. Dessa maneira, advinda de práticas assistemáticas, a avaliação em tradução se mostra cada vez mais adepta ao desenvolvimento e adoção de abordagens metódicas e criteriosas para a avaliação de traduções, as quais ocorrem em três esferas principais, quais sejam:

- a) avaliação de traduções publicadas;
- b) avaliação do trabalho de tradutores profissionais;
- c) avaliação no ensino de tradução.

A necessidade de se avaliar traduções é dependente do âmbito em que ocorre. Isto é, os motivadores da tarefa avaliativa em contextos educacionais são distintos daqueles de contextos de tradução profissional ou de crítica de tradução, por exemplo. Apesar das distinções quanto às condições que ensejam o processo avaliativo, as características constitutivas de boas traduções permanecem nebulosas, fazendo com que esse campo seja permeado por perguntas seculares como: O que é uma boa tradução? Haveria um modelo de avaliação aceito universalmente? ou ainda: Quais são os critérios que devem ser satisfeitos para que uma tradução seja considerada excelente, boa ou aceitável? A essas perguntas, Sager (1989) respondeu não haver normas absolutas relativas à qualidade de tradução, propondo haver traduções mais ou menos apropriadas de acordo com os propósitos que elas devem desempenhar em um dado contexto. A subjetividade secularmente inerente às análises de traduções, contudo, pode ser substituída por uma abordagem de caráter mais objetivo, a qual precisaria se basear em procedimentos reproduzíveis, por meio de um conjunto de



parâmetros claramente definidos e de um sistema de valores gradativos em escala (Sager, 1989). Este artigo se debruçará sobre um desses parâmetros, o qual pode ser adotado durante o processo avaliativo de tradução: a análise computadorizada de erros.

Tarefa onerosa que pressupõe a criação de taxonomias de erros e sistemas de codificação, a análise de erros é, contudo, uma ferramenta que permite a identificação de áreas problemáticas no ensino-aprendizagem de tradução e contribui para a formulação de metodologias didáticas mais focadas em necessidades tangíveis. Conhecida por seu caráter mais objetivo e quantitativo (estatística de erros) e qualitativo (tipologia de erro), a análise computacional de erros em corpus de aprendiz possibilita o que Martínez Melis e Hurtado Albir (2002) denominam de “avaliação fundamentada” (*reasoned assessment*), baseada no estabelecimento de critérios concretos. Nesse contexto, as próximas seções se dedicarão a a) delinear as características da avaliação no âmbito do ensino de tradução; b) definir a concepção de erro; e c) descrever criticamente o sistema taxonômico de etiquetagem TAS (*Translation-oriented Annotation System*), desenvolvido por Granger e Lefer (2021).

De forma mais concreta, a seção 2 tratará do erro em contextos avaliativos de tradução, enquanto que a seção 3 conceituará corpus de aprendiz e a centralidade dos sistemas taxonômicos de erros para a anotação e a investigação de produções autênticas de estudantes. A seção 4 se dedicará à descrição do sistema TAS e de uma metodologia experimental para avaliação de traduções. A seção 5 indicará como a análise de erros pode ser usada no ambiente pedagógico de tradução.

2. Avaliação na tradução e a definição de erros

A consolidação dos Estudos da Tradução como área de pesquisa autônoma, e o posterior surgimento de cursos de graduação, revela a necessidade de um ensino mais sistemático que esteja em consonância com as exigências de pesquisa e mercadológicas atuais e que incorpore as tecnologias de otimização do tempo e de pesquisa terminológica para entrega de produtos tradutórios de qualidade. Tal abordagem de ensino-aprendizagem ultrapassaria o que Alekseeva (2000 citada por Askari, 2019, p. 286) chama de “abordagem tradicional” no ensino de tradução. Dentro dessa abordagem perdura ainda, em certos cursos superiores, um modelo de formação do tradutor baseado na transmissão de conhecimento vertical, em que um professor-tradutor mais experiente instrui os aprendizes. Nesse método, o tradutor mais experiente é quem faz a seleção de textos e a avaliação da tradução dos estudantes ocorre a partir da adequabilidade textual quando comparada ao texto do professor. Isto é, nesse método de ensino, conhecido como método autoritário-criativo, o professor se apoia quase que exclusivamente em seu próprio conhecimento. Na mesma linha, Al-Hadithy (2015) aponta, como as principais críticas ao ensino tradicional de tradução, a centralização na figura do professor, a rigidez e a pouca atualização e inventividade no ensino. Nesse modelo, os professores deteriam as estratégias e equivalentes de tradução, os quais disponibilizariam a estudantes que produzissem versões faltosas, dando a ideia de que as respostas do professor seriam as únicas corretas e de que textos seriam unívocos, usando, assim, critérios subjetivos de análise e de avaliação. De acordo com Martínez Melis e Hurtado Albir (2002), a



avaliação intuitiva ainda é uma tendência no ensino e na prática profissional em tradução, consistindo em análises subjetivas e baseadas em impressões que não seguem critérios claros.

A prática avaliativa orientada pela intuição se alinha à definição de avaliação holística de Verplaetse (2022) já que esse método considera o todo e se baseia em impressões sobre o alcance, por exemplo, da função da tradução na comunidade de recepção, tal como preconizada pela teoria do *skopos*, e é conduzida com base nas percepções, sentimentos e idiossincrasias do avaliador (Han, 2020). Nesses contextos, a análise magistral, isto é, a autoridade do professor, enquanto avaliador e tradutor mais experiente, se torna o método de avaliação inquestionável (Darwish, 2001).

Urge, assim, investigar a existência e eficácia de outros métodos avaliativos alternativos que orientem, sobretudo, o desenvolvimento de práticas pedagógicas direcionadas às verdadeiras necessidades de aprendizagem dos estudantes e que levem à explicitação das ferramentas de avaliação usadas e de como elas interagem com as práticas de ensino.

Nesse sentido, Martínez Melis e Hurtado Albir (2002) já haviam chamado atenção para o fato de que a maioria das pesquisas acerca de avaliação no ensino de tradução se concentrava em escalas de avaliação e na classificação e ocorrência de erros. A análise de erros se insere nas abordagens analíticas de avaliação de tradução, as quais identificam desvios específicos nas produções, aplicando critérios majoritariamente linguísticos para se ponderar a qualidade de um texto a partir do estabelecimento de taxonomias de erros (ver seção 3 e 4).

A existência de erros percebidos, contudo, durante a correção de um texto por um professor ocasiona indagações já feitas por Nord (1996), isto é: O que é um erro de tradução? Todos os erros identificados são propriamente “erros de tradução”? Como diferenciar os erros mais graves dos menos graves? Como partir do diagnóstico de erros ao tratamento e também a uma prevenção? Como fazer com que a laboriosa tarefa de avaliação de tradução contribua para o melhoramento das habilidades dos estudantes?

A partir de uma abordagem funcionalista, Nord (1996) destaca que a definição e avaliação de erros invariavelmente depende da fixação de um objetivo de aprendizagem claramente definido como ponto de referência. Para a autora, o principal objetivo do ensino de tradução é fazer com que os discentes desenvolvam as habilidades necessárias para a produção de textos que funcionem na cultura receptora e que correspondam às necessidades comunicativas de um dado projeto tradutório. Nessa perspectiva os erros de tradução corresponderiam, então, ao não cumprimento do objetivo tradutório, estabelecido no *briefing* de tradução, por exemplo (Nord, 1996). Dessa maneira, a avaliação de traduções, sob um viés funcionalista tomaria em conta a função do texto de chegada, a coerência e tipologia textual, bem como as convenções de gênero, linguísticas e culturais. Nesse bojo, a perspectiva funcionalista da avaliação revela-se problemática, nas palavras de Hönig (1997), por não fornecer regras absolutas já que desconsidera, precipuamente, a existência de traduções incondicionalmente corretas. Contudo, um problema atrelado a essa abordagem é: quem tem a autoridade para definir a função de uma tradução (Hönig, 1997)?

Quando se adota uma visão linguística, contudo, os erros podem ser definidos como um desvio a um sistema de normas e regras (Hönig, 1997), ou como um defeito ou falha na língua de chegada ocorrido durante a tentativa de reescrever o material textual de partida (Igalada, 2016), os quais podem ser quantificados, identificados, descritos e remediados. Esses erros, também denominados de erros linguísticos impactam, segundo Igalada (2016), a aceitabilidade

microlinguística, já que afetam aspectos gramaticais, lexicais, semânticos, morfológicos e sintáticos. Essa inadequação a nível linguístico, contudo, inevitavelmente afeta a aceitabilidade macrolinguística de um texto, comprometendo aspectos pragmáticos e estilísticos.

Tendo em vista que a aceitabilidade de um texto está diretamente relacionada à qualidade do produto final, se faz necessário, assim, estipular parâmetros para a avaliação de traduções no contexto de ensino que ofereçam subsídios pedagógicos e que orientem o processo de ensino-aprendizagem, já que, como aponta Hönig (1997), as práticas de avaliação da qualidade de textos traduzidos em cursos superiores são marcadas por sua falta de homogeneidade. Ademais, estudantes de tradução, indubitavelmente, não apenas se beneficiam, mas necessitam da avaliação de seus textos para que possam melhorar a qualidade de seus trabalhos (Hönig, 1997).

Há que se considerar, nesse sentido, a existência de múltiplos tipos de avaliação de tradução no âmbito instrucional e que visam a desempenhar funções particulares (Martínez Melis & Hurtado Albir, 2002), quais sejam:

- a) *avaliação diagnóstica*: possibilita a obtenção de um prognóstico do potencial do estudante, isto é, revela as habilidades e dificuldades dos estudantes antes do início do processo de aprendizado, servindo também como uma espécie de teste de nivelamento;
- b) *avaliação somativa*: visa a determinar os resultados finais e verificar se dado conhecimento foi adquirido e se os objetivos de aprendizagem foram atingidos. Geralmente, é de natureza compulsória e é realizada ao fim do processo de aprendizagem;
- c) *avaliação formativa*: objetiva coletar informações para fins de formação, tendo como foco principal o estudante, que se torna ativamente envolvido no próprio processo de aprendizado. Esse tipo de avaliação se preocupa tanto com o processo, quanto com o resultado da aprendizagem e os dados obtidos servem como ferramenta de retroação e correção de problemas.

A função das diferentes avaliações, contudo, não oferece efetivamente subsídios de como elas devem ser realizadas ou que faceta da tradução (processo/produto) deve ser objeto de investigação. Por isso a necessidade de implementar uma tipologia de erros, a qual delinea as áreas do texto a serem observadas e permite a identificação dos níveis em que um determinado desvio ocorre. O reconhecimento de erros por meio de um sistema de classificação fornece maior direcionalidade e foco ao processo avaliativo já que possibilita, segundo Martínez Melis e Hurtado Albir (2002), estabelecer:

- I. a diferença entre erros relativos ao texto de partida (sentido incorreto ou oposto, adição ou omissão) e ao texto de chegada (ortografia, vocabulário, sintaxe, coerência e coesão);
- II. a distinção entre erros individuais, sistemáticos e aleatórios na população investigada.

Depreende-se, assim, que a identificação do erro não apenas permite a adoção de medidas corretivas e um acompanhamento individualizado dos estudantes, mas também: I) fomenta a

autoavaliação e conscientização sobre os próprios erros, por parte dos discentes; e 2) evidencia a natureza múltipla dos erros e de como eles podem ser tratados.

A utilização de erros para fins de avaliação pressupõe, preliminarmente, algumas etapas, quais sejam, sua identificação em um corpus de aprendiz, localização, descrição, etiquetagem e codificação, as quais só podem ocorrer após o estabelecimento de categorias de erro, assuntos que serão objeto da seção seguinte.

3. Corpus de aprendiz, taxonomia de erros e sistemas de etiquetagem

A anotação e análise de erros enquanto etapas que compõem o processo de avaliação de traduções são orientadas pelo produto e têm como objeto as produções textuais de estudantes. Para que seja possível obter dados acerca das produções tradutórias de estudantes, e de possíveis pontos de melhoria, faz-se necessária a coleta de um corpus de aprendizes de tradução, que consiste em um conjunto de traduções produzidas por estudantes de tradução e cuja principal função é fornecer apontamentos acerca dos traços linguísticos significativos dos textos, os quais, por sua vez, orientarão a prática pedagógica em tradução (Castagnoli et al., 2011). Largamente usados em estudos de Aquisição de Segunda Língua, os corpora eletrônicos de aprendizes são definidos por Granger (2002, p. 7, tradução nossa¹) como “[...] coleções eletrônicas de dados textuais autênticos coletados a partir de critérios explícitos e para fins de estudo sobre aquisição de segunda língua ou de língua estrangeira. Esses dados são codificados e uniformizados homogeneamente e documentados quanto a sua origem e fonte”.

A coleta de textos de estudantes, entretanto, não basta para a análise de possíveis erros, sendo premente a criação prévia de uma taxonomia ou tipologia de erros, que será atribuída a produções faltosas por meio de códigos, permitindo a anotação manual e posterior análise por ferramentas da Linguística de Corpus². A anotação, por sua vez, pode ser definida como a prática de se adicionar informações interpretativas a um corpus por meio do uso de códigos preliminarmente definidos. Um exemplo de anotação é a etiquetagem de classes morfológicas e a etiquetagem de erros (Granger, 2002). A etiquetagem de erros é especialmente criada para tratar da natureza anômala da produção linguística de estudantes e pode rotular os erros a partir da natureza (gramatical, lexical) ou origem (interlingual ou intralingual) do desvio, sendo que a identificação de erros a partir da natureza é preferível, segundo Granger (2002), por envolver menor nível de subjetividade e maior consistência.

Os sistemas de etiquetagem de um texto consistem no uso de taxonomias de erros baseadas em categorias linguístico-estruturais (Callies, 2015). A taxonomia de erros, conforme propõem Díaz-Negrillo e Domínguez (2006), deve ser baseada na descrição de dados observáveis e incluir

¹ No texto fonte: “[...] *computer learner corpora are electronic collections of authentic FL/SL textual data assembled according to explicit design criteria for a particular SLA/FLT purpose. They are encoded in a standardised and homogeneous way and documented as to their origin and provenance*” (Granger, 2002, p. 7).

² A Linguística de Corpus pode ser definida como uma abordagem metodológica fundamentada no uso de conjuntos de textos em formato eletrônico (corpora) reunidos a partir de um critério específico com vistas a investigar um dado problema de pesquisa. A principal informação fornecida pelo estudo de corpora eletrônicos, por meio de ferramentas da Linguística de Corpus, diz respeito à frequência de determinados aspectos linguísticos. Essa abordagem, pautada no uso do computador, se mostra bastante propícia para estudos de caráter quantitativo (Granger, 2002).

categorias linguísticas bem delineadas a fim de reduzir ao máximo a subjetividade no processo de diagnóstico e classificação de erro no corpus. A partir do estabelecimento de categorias de erros, os textos são etiquetados com vistas a identificar comportamentos desviantes nas escritas dos estudantes, quando comparados a corpora de referência ou de língua geral, por exemplo. Embora uma taxonomia de erros possa ser criada por um grupo de pesquisa para ser usado em um projeto específico, Granger (2003) defende que, para ser eficaz, um sistema de anotação de erros deve ser consistente, informativo, flexível e reutilizável em uma variedade de línguas.

Conforme explicitam Díaz-Negrillo e Domínguez (2006), a taxonomia de erro pode incluir a descrição da produção textual em diferentes dimensões, quais sejam:

1. classificação de categoria linguística, que abarca as características linguísticas do erro relativos ao léxico, tempo verbal e regência, por exemplo;
2. classificação quanto à modificação da língua quando comparada a produções típicas em contextos normativos da língua de chegada, em termos de omissão, adição e ordenação.

Granger e Lefer (2020), por sua vez, dividem a tipologia de erros em duas principais categorias, a saber:

1. erros de transferência de conteúdo, que têm ligação com o processo da tradução, como erros de omissão ou distorção;
2. erros linguísticos que ocorrem independentemente do processo de tradução, como aqueles relacionados à concordância ou tempo verbal.

Tendo tratado da importância de se definir uma tipologia de erros para conduzir a análise computadorizada de erros, a seção seguinte descreverá o *Translation-oriented annotation system*, um sistema de anotação de traduções que abarca as categorias de Granger e Lefer (2020) para a descrição de traduções de aprendizes, a qual busca salientar erros de transferência, erros linguísticos e boas soluções tradutórias.

4. Translation-oriented annotation system: características e formas de uso

O *Translation-oriented annotation system* é um sistema de descrição e etiquetagem de traduções desenvolvido no âmbito das pesquisas do projeto MUST (*Multilingual Student Translation*), o qual visa à criação de um ambiente *online* de armazenamento de traduções anotadas e alinhadas de estudantes para o desenvolvimento de materiais didático, formação de tradutores, lexicografia pedagógica, bem como para fins de pesquisa. O ambiente *online* aqui referenciado é a plataforma *Hypal4MUST*, a qual permite a coleta e anotação de textos no âmbito do projeto, garantindo a:

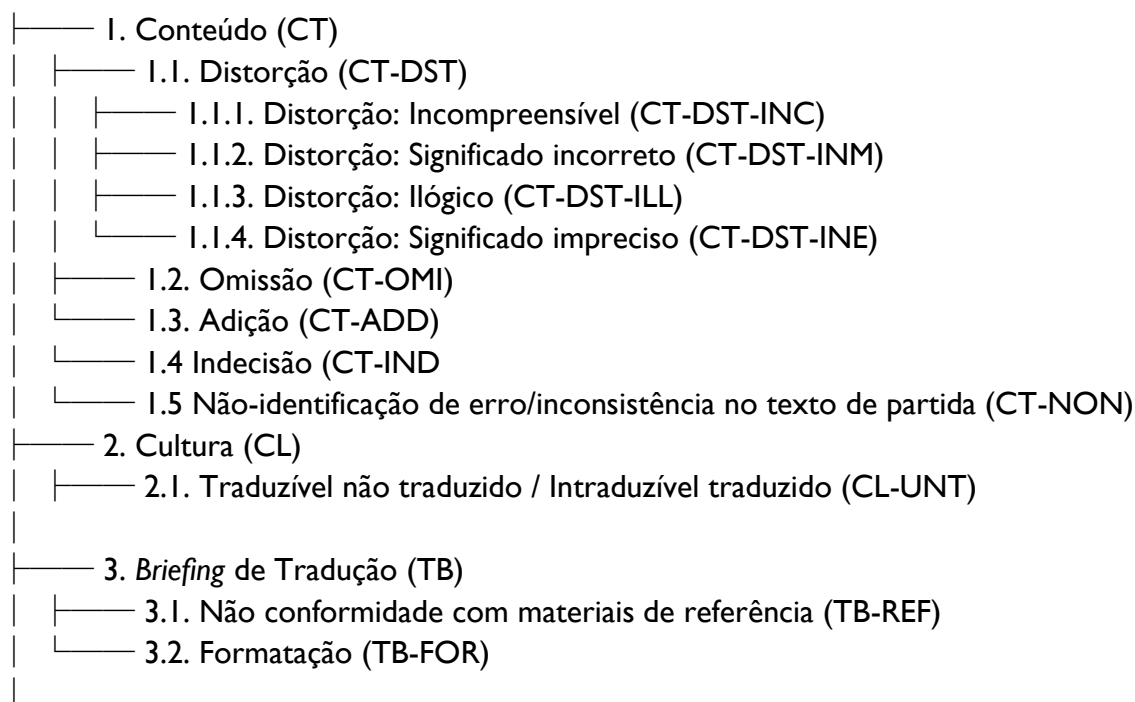
- armazenagem dos textos de partida, traduções e metadados;
- alinhamento automático dos corpora a nível de parágrafo e oração;
- lematização e etiquetagem de erros dos corpora;
- retroação aos estudantes;
- anotação das traduções a partir de um sistema taxonômico padrão.



As desenvolvedoras (Granger & Lefer, 2020) do TAS optaram por não incluir em sua denominação a expressão “sistema de anotação de erros de tradução” pois a taxonomia concebida inclui não apenas erros de tradução, abarcando também as categorias referentes a “boas soluções tradutórias” e aquelas que possibilitam a identificação do procedimento de tradução utilizado na solução de um problema³ de tradução (e.g. normalização, empréstimo). Nesse sentido, o TAS apresenta uma faceta dedicada à descrição de erros de natureza linguística, especialmente úteis para a descrição de versões, isto é, da tradução feita para a língua estrangeira, e de uma faceta destinada às soluções decorrentes da operação linguística sendo realizada, isto é, a tradução. Cumpre, assim, listar os aspectos mais representativos do sistema em crivo, os quais configuram também características básicas de sistemas de anotação confiáveis:

- a) Número limitado de etiquetas;
- b) Categorias bem delimitadas e sem duplicidade;
- c) Estrutura hierárquica para anotação de textos em diferentes níveis (e.g. erro gramatical, na classe verbal, no aspecto “tempo”);
- d) Documentação detalhada para reutilização do sistema.

O TAS (anexo I), baseado no *Louvain error tagging*⁴, é composto por 12 categorias e 48 subcategorias, representadas por códigos hierárquicos, estruturadas da seguinte forma:



³ O problema de tradução é definido por Nord (1991) como uma tarefa objetiva de transferência a qual todo tradutor deve resolver durante o processo tradutório. Haveria quatro categorias de problemas, quais sejam: a) problemas de ordem pragmática, resultantes do contraste entre a situação de uso do texto de partida e a de produção, do texto de chegada; b) problemas de convenção, como as diferenças existentes entre sistemas de medida ou convenções do sistema literário; c) problemas linguísticos, causado pelas distinções estruturais entre as línguas envolvidas; e d) problemas específicos do texto, devido a características próprias do material de partida, como a existência de jogos de palavras.

⁴ Sistema de etiquetagem de erros desenvolvido inicialmente para etiquetar a primeira versão do *International Corpus of Learner English*.

- 4. Gramática e Sintaxe (GR)
 - 4.1. Morfologia flexional (GR-INF)
 - 4.2. Tempo e aspecto verbal (GR-TNS)
 - 4.3. Voz verbal (GR-VOI)
 - 4.4. Ordem das palavras (GR-WORD)
 - 4.5. Confusão de categoria de palavras (GR-WCO)
 - 4.6. Determinante (GR-DET)
 - 4.7. Preposição (GR-PRE)
 - 4.8. Pronome (GR-PRO)
 - 4.9. Concordância (GR-CCD)
 - 4.10. Conjunção (GR-CJX)
 - 4.11. Complementação (GR-COM)
 - 4.11.1. Adjetivo (GR-COM-ADJ)
 - 4.11.2. Substantivo (GR-COM-NOU)
 - 4.11.3. Verbo (GR-COM-VRB)
 - 4.11.4 Advérbio (GR-COM-ADV)
- 5. Léxico e Terminologia (LT)
 - 5.1. Palavra (LT-WRD)
 - 5.1.1 Palavra única (LT-WRD-SWO)
 - 5.1.2. Derivação (LT-WRD-DER)
 - 5.1.3. Cognato (LT-WRD-COG)
 - 5.1.4 Unidades de multipalavras (LT-WRD-MWO)
 - 5.1.5 Expressões compostas (LT-WRD-MWO-COP)
 - 5.1.6 Colocação (LT-WRD-MWO-COL)
 - 5.1.7. Expressões idiomáticas (LT-WRD-MWO-IDI)
 - 5.2. Termos (LT-TRM)
 - 5.2.1. Termo traduzido por não-termo (LT-TRM-TNT)
 - 5.2.2. Inconsistência terminológica (LT-TRM-TIN)
- 6. Discurso e Pragmática (DP)
 - 6.1. Coesão (DP-COH)
 - 6.1.1. Coesão: Uso de conectivos (DP-COH-LIN)
 - 6.1.2. Coesão: Referência pronominal (DP-COH-PRF)
 - 6.2. Estruturação da Informação (DP-STR)
 - 6.2.1. Progressão temática (DP-STR-THE)
 - 6.2.2. Frases divididas ou fundidas (DP-STR-SPL)
 - 6.2.3. Preferências sintáticas (DP-STR-SYN)
- 7. Mecânica (ME)
 - 7.1. Ortografia (ME-SPE)
 - 7.2. Pontuação (ME-PUN)
 - 7.3. Unidades, datas e números (ME-UDN)

- |
- | — 8. Estilo e Registro (REG)
- | — 8.1. Grau de formalidade (REG-FML)
- | — 8.2. Estilo truncado (REG-HEA)
- | — 8.3. Redundância (REG-RED)
- 9. PLUS (PLUS)
- 10. Interferência da língua de partida (SLI)
- 11. Distorção semântica (DIS)
- 12. Repetição de Erro (REP)

Do esquema acima, pode-se inferir que as categorias de 1 a 8 dizem respeito a desvios identificados nos campos linguísticos, discursivos e de gênero, da produção escrita, enquanto que as categorias de 9 a 12 relacionam-se a ocorrências decorrentes do próprio processo tradutório, o qual pode ensejar a tomada de boas decisões⁵ (PLUS), a interferência da língua de partida (SLI), a distorção semântica causada por qualquer tipo de erro (DIS) e a repetição consecutiva de erros da mesma natureza (REP). De caráter descritivo, esse sistema de anotação de traduções possibilita a detecção e etiquetagem de produções a serem avaliadas. Produzido de forma flexível e reutilizável, conforme preconizava Granger (2003), ele pode ser usado em uma plethora de línguas e é organizado hierarquicamente, com a atribuição de códigos a partir dos níveis mais gerais aos mais específicos.

Tendo estabelecido, assim, as características fundamentais do TAS, a seção seguinte se dedicará a mostrar possíveis formas de utilização e implementação do sistema enquanto componente de processos avaliativos de tradução no âmbito educacional.

4.1 Anotação de corpus de aprendiz a partir do sistema TAS com um editor de erros

O TAS é um sistema de anotação de traduções e, como tal, se constitui apenas como taxonomia, fornecendo os subsídios tipológicos para a descrição e não os instrumentais. Com isso em mente, propõe-se aqui a utilização de um *software* para realizar as anotações do corpus de aprendiz de forma mais célere, o *UCLouvain Error Editor (UCLEEv2)*⁶. Esse editor de erros facilita a inserção de etiquetas e também a de sugestões de correções nos textos dos aprendizes, além de conter funcionalidades como:

- Menu com seleção de etiquetas;
- Concordanciador;
- Gerador de frequências;
- Gerador de exercícios automáticos.

Ao descarregar e instalar o *software* na máquina, o sistema de etiquetagem padrão é o do *Louvain Error Tagging Version 2.0* (Granger et al., 2023b), contudo, a flexibilidade do sistema permite

⁵ A inclusão dessa categoria se alinha ao que Kussmaul (1995) defende quando diz que a motivação é essencial aos alunos de tradução e a avaliação positiva de boas estratégias tradutórias contribui nesse aspecto.

⁶ Esse *software* foi desenvolvido pelo *Centre for English Corpus Linguistics (CECL)* da Universidade de Louvain, na Bélgica. Mais informações podem ser consultadas em Granger et al. (2023a).

que se faça a inclusão de outras estruturas de anotação, com base no Guia de usuário do *UCLouvain Error Editor* (Granger et al., 2023c).

A criação de um novo arquivo, com códigos do sistema de anotação *TAS*, por exemplo, pode ser feita em um documento *.TXT*, respeitando os níveis hierárquicos entre categorias e subcategorias, conforme demonstrado na Figura 1.

Figura 1: Modelo de criação de arquivo do sistema de anotação de traduções (*TAS*) para uso no *Louvain Error editor*

```
[CT
    CT-DST
    CT-DST-INC
    CT-DST-INM
    CT-DST-ILL
    CT-DST-INE
    CT-OMI
    CT-ADD
    CT-IND
    CT-NON
]
[CL
    CL-UNT
]
[TB
    TB-REF
    TB-FOR
]
[GR
    GR-INF
    GR-TNS
    GR-VOI
    GR-WOR
    GR-WCO
    GR-DET
    GR-PRE
    GR-PRO
    GR-CCD
    GR-CJX
    GR-COM
    GR-COM-ADJ
    GR-COM-NOU
    GR-COM-VRB
    GR-COM-ADV
]
```

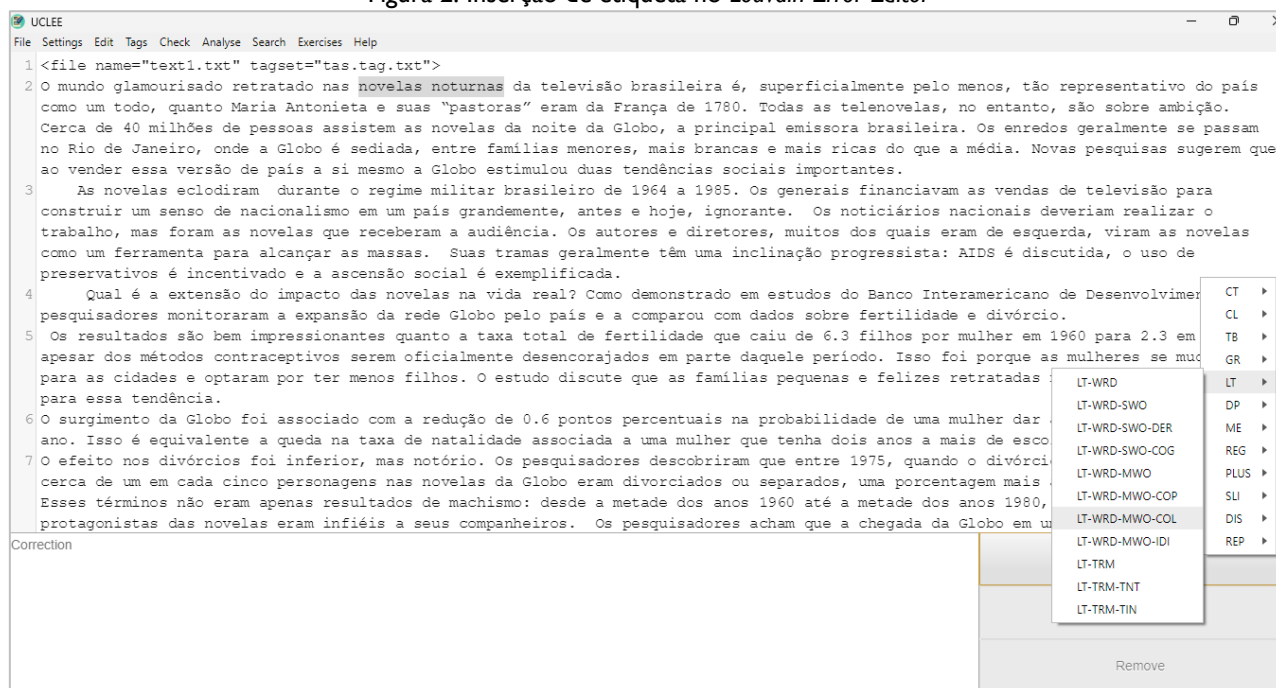
Fonte: Autora (2025)

[Descrição] Lista de categorias e subcategorias de análise do *TAS* organizadas na vertical e hierarquizadas, em que uma categoria mais abrangente é deslocada para a esquerda e seus componentes são dispostos a partir da linha seguinte com uma indentação [Fim da descrição].

Na figura acima, o código “CT” representa a categoria “Conteúdo”, sendo sucedida pelas subcategorias abarcadas e relacionadas aos desvios da mesma natureza concernentes a produções incompletas (CT-DST-INC), ilógicas (CT-DST-ILL), incorretas (CT-DST-INM), inexatas (CT-DST-INE) ou que apresentem omissões (CT-OMI), adições (CT-ADD) e afins. Após a criação do arquivo, que deve ser inserido na pasta de etiquetas do *software*, é possível usá-lo para anotar o corpus de aprendiz, o qual também deve estar em formato *.TXT*.

A anotação é feita em uma interface amigável com a seleção do trecho problemático e a escolha dos códigos para descrição do desvio, processo esse ilustrado na Figura 2:

Figura 2: Inserção de etiqueta no *Louvain Error Editor*

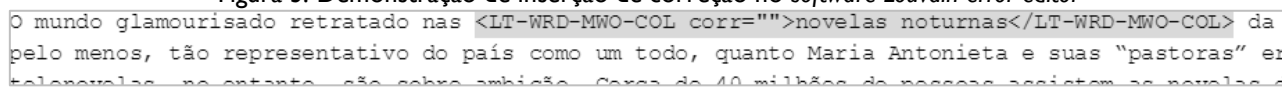


Fonte: Autora (2025)

[Descrição] Interface do *Louvain Error Editor* que contém as janelas *file*, *settings*, *edit*, *tags*, *check*, *analyse*, *search*, *exercises* e *help*. O texto a ser anotado toma toda a tela e um trecho, em que se lê “novelas noturnas” está em destaque. No canto inferior direito há um menu de seleção em que estão listadas as categorias de erro, entre as quais o analista deve escolher para atribuir à parte selecionada [Fim da descrição].

Como depreendido, as etiquetas, que indicam a tipologia do erro, são marcadas entre os símbolos de menor e maior que (<>), por meio dos códigos que fazem referência aos diferentes níveis de anotação. Além da identificação, rotulação e descrição dos erros, o sistema permite, ainda, a inclusão de uma correção (entre aspas), marcação essa que se mostra útil no processo de retroação aos estudantes e em sua própria autoavaliação, conforme demonstrado na figura a seguir:

Figura 3: Demonstração de inserção de correção no software *Louvain error editor*



Fonte: Autora (2025)

[Descrição] Excerto do corpus em que se lê: “o mundo glamourizado retratado nas novelas noturnas da pelo menos, tão representativo do país como um todo, quanto Maria Antonieta e suas “pastoras”. A expressão “novelas noturnas”, que aparece em destaque, foi etiquetada com o código <LT-WRD-MWO COL corr=>. O “corr” permite a inserção de uma sugestão de tradução alternativa pelo analista etiquetador [Fim da descrição].

Ao finalizar a etiquetagem de todos os textos que compõem o corpus de aprendiz, as funcionalidades adicionais do *Louvain Error Editor* possibilitam a obtenção de informações que servem de orientação para reformulações pedagógicas, ponto esse que será discutido na subseção seguinte.

4.1.1 Investigando um corpus de aprendizes de tradução para fins pedagógicos: uma metodologia experimental

Esta seção mostrará, de forma experimental, uma maneira de operacionalizar a análise dos dados de caráter quantitativo e qualitativo com o uso do *Louvain error editor* e com ferramentas presentes na maioria dos *softwares* de Linguística de Corpus.

Uma possível abordagem à análise de corpus de aprendizes de tradução pode se inspirar na metodologia sugerida por De Sutter et al. (2017), em seu trabalho sobre a avaliação de traduções com base em análises estatísticas de corpus, e ser composta dos seguintes estágios:

1. Compilação do corpus de aprendiz;
2. Processamento do corpus—com a etiquetação de erros;
3. Extração de características relevantes do *texto de partida* por meio de ferramentas como *keyword*⁷, *n-grams*⁸, *collocate*⁹;
4. Criação de um corpus paralelo, com o texto de partida e as traduções dos estudantes, para verificar como as características relevantes foram vertidas;
5. Análise estatística;
6. Interpretação dos dados.

A etapa dois, a qual foi demonstrada na seção anterior, e a cinco podem ser colocadas em prática por meio da própria plataforma do *Louvain error editor*, em cuja aba *Analyse* realiza-se uma série de processamentos que permitem quantificar, organizar e analisar os erros identificados (Granger et al., 2023b). Os dados fornecidos por meio dessas operações já permitem um vislumbre dos erros mais recorrentes e sua distribuição no corpus, isto é, se se tratam de desvios mais pontuais e individuais ou mais sistemáticos, que perpassam toda a população, como o erro LT-TRM-TNT, o qual ocorre igualmente (uma vez) nos textos que compõem o corpus ilustrativo da figura 4.

Os dados de natureza estatística produzidos pelo editor de erros em questão podem ser visualizados de três formas principais:

1. frequência absoluta (*absolute frequency*) — que indica o número de ocorrências por etiqueta;
2. distribuição relativa por etiqueta (*by tag*) — que exhibe a porcentagem de erros de uma determinada categoria nos textos do corpus; e
3. distribuição por arquivo (*by file*) — que apresenta, em porcentagem, a representação de cada erro em arquivos individuais, organizados horizontalmente.

⁷ Lista das palavras mais estatisticamente mais representativas de um corpus quando comparadas com as de um corpus de referência (Bernardini, 2022).

⁸ Unidades lexicais compostas por mais de uma palavra (e.g. *on the one hand*), incluindo conectivos, sintagmas e idiomatismos (Rees, 2022).

⁹ Exibe as palavras que co-ocorrem mais frequentemente com a palavra buscada em uma dada janela de palavras.

Figura 4: Interface do Louvain error editor com indicação da frequência de erros

The screenshot shows the 'UCLEE - Analyse' window. It has a 'Files' table at the top left, a 'Search' panel on the top right, and a large 'Frequencies' table at the bottom. The 'Files' table lists four files with their respective token and tag counts. The 'Search' panel includes input fields for 'Tag', 'Error', and 'Correction', a 'Sort by' dropdown set to 'Tag', and a 'Find & Sort' button. The 'Frequencies' table has columns for 'File', 'Total', and various error categories like CT-DST, CT-DS, CT-DS..., CT-OMI, CT-ADD, GR-TNS, GR-PRE, GR-CCD, GR-CO..., LT-WRD, LT-WR..., LT-TRM-TNT, LT-TR..., DP-COH, DP-CO..., ME-SPE, and ME-F. A summary row at the bottom shows the percentage of each error type relative to the total number of tags.

File	Total	CT-DST	CT-DS	CT-DS...	CT-DS...	CT-OMI	CT-ADD	GR-TNS	GR-PRE	GR-CCD	GR-CO...	LT-WRD	LT-WR...	LT-TRM-TNT	LT-TR...	DP-COH	DP-CO...	ME-SPE	ME-F
C:\Users\Júlia Valverde\Desktop\...\8	8										1			1			1	2	
C:\Users\Júlia Valverde\Desktop\...\19	19	1	1	2	1	4			1			2	1	1		1		2	
C:\Users\Júlia Valverde\Desktop\...\16	16	1	1			2		1	1	1				1				3	2
C:\Users\Júlia Valverde\Desktop\...\11	11	1			1		1	1						1	1			1	
% out of total number of tags	100%	6%	4%	4%	4%	11%	2%	4%	4%	2%	2%	4%	2%	7%	2%	2%	2%	15%	4%
Total	54	3	2	2	2	6	1	2	2	1	1	2	1	4	1	1	1	8	2

Fonte: Autora (2025)

[Descrição] A interface está dividida em várias seções principais. Na parte superior esquerda, os textos que compõem o corpus aparecem listados em tabela. À direita dessa tabela, há os respectivos números de *tags* e *tokens* de cada texto. Na parte superior direita, há uma barra de pesquisa e filtros. Essa barra permite buscar e ordenar erros específicos por *tags*, tipo, correção ou outros critérios. Na metade inferior da interface, são exibidos dados sobre a frequência dos erros. Todas as categorias de erro presentes são organizadas verticalmente, em tabela, e o número de desvios é indicado em relação aos documentos em que foram identificados. Há várias formas de se analisar os dados quantitativos. A tabela mostra, para cada arquivo, a quantidade de ocorrências de cada *tag* e o percentual de cada tipo de erro em relação ao total de erros. Na parte inferior da tela, uma linha final resume as porcentagens totais de cada tipo de erro para todos os arquivos carregados [Fim da descrição].

Figura 5: Distribuição relativa de etiquetas nos arquivos que compõem o corpus

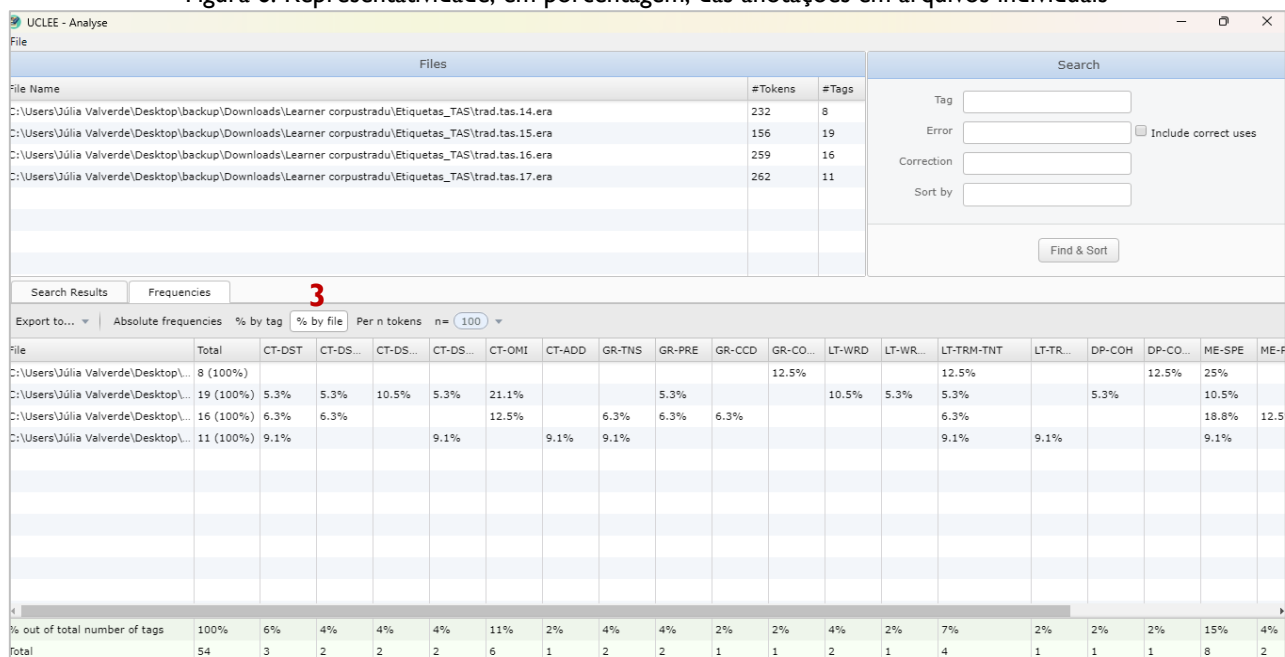
This screenshot shows the 'Frequencies' tab in the Louvain error editor. It displays a table where each row represents a file and each column represents a specific error tag. The values are percentages indicating the relative frequency of each tag within that file. A summary row at the bottom provides the overall percentage distribution for all tags across the entire corpus.

File	Total	CT-DST	CT-DS	CT-DS...	CT-DS...	CT-OMI	CT-ADD	GR-TNS	GR-PRE	GR-CCD	GR-CO...	LT-WRD	LT-WR...	LT-TRM-TNT	LT-TR...	DP-COH	DP-CO...	ME-SPE	ME-F
C:\Users\Júlia Valverde\Desktop\...\8	8										100%			25%			100%	25%	
C:\Users\Júlia Valverde\Desktop\...\19	19	33.3%	50%	100%	50%	66.7%			50%			100%	100%	25%		100%		25%	
C:\Users\Júlia Valverde\Desktop\...\16	16	33.3%	50%			33.3%		50%	50%	100%				25%				37.5%	100%
C:\Users\Júlia Valverde\Desktop\...\11	11	33.3%			50%		100%	50%						25%	100%			12.5%	
Total %		100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
% out of total number of tags	100%	6%	4%	4%	4%	11%	2%	4%	4%	2%	2%	4%	2%	7%	2%	2%	2%	15%	4%
Total	54	3	2	2	2	6	1	2	2	1	1	2	1	4	1	1	1	8	2

Fonte: Autora (2025)

[Descrição] Recorte da parte inferior da interface do Louvain error editor. Nessa parte, os erros são organizados a partir da distribuição *by tags*, isto é, distribuição relativa de etiquetas individuais no corpus em porcentagem. Esses dados são apresentados em tabela, em que cada erro aparece verticalmente com os respectivos números de desvios correspondentes aos textos, os quais estão listados à esquerda [Fim da descrição].

Figura 6: Representatividade, em porcentagem, das anotações em arquivos individuais

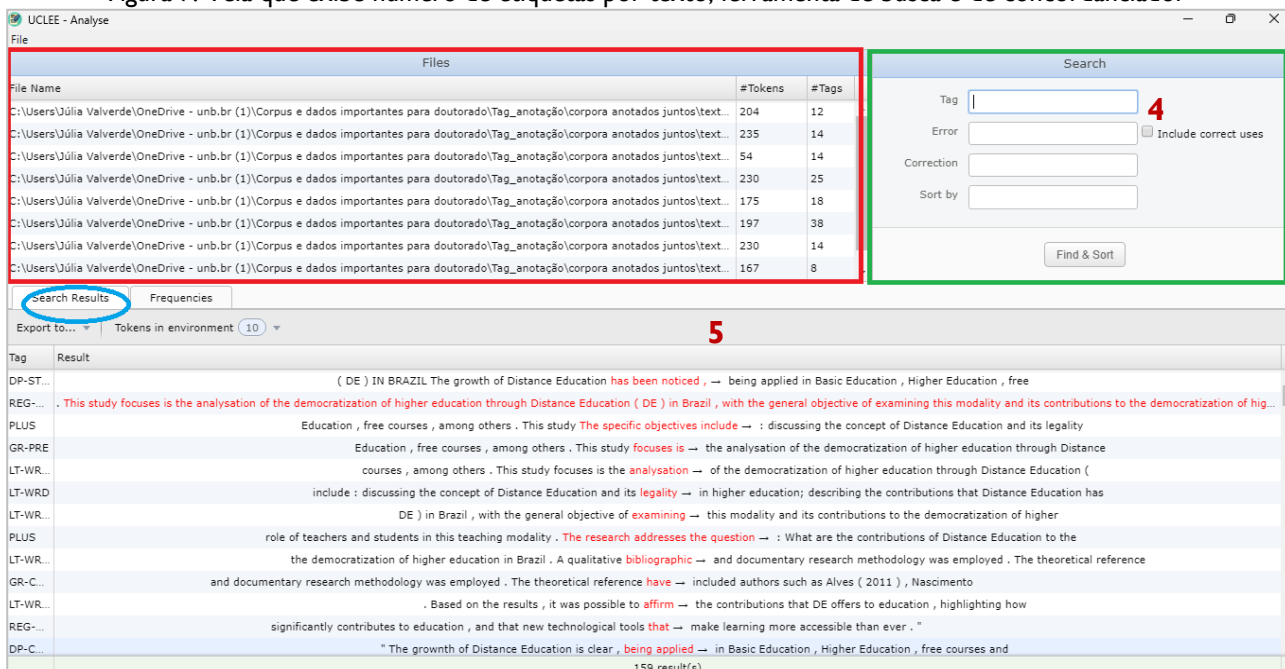


Fonte: Autora (2025)

[Descrição] mesma interface do *Louvain error editor* descrita anteriormente., contudo com novo critério de exibição dos dados quantitativos. Foi selecionada a opção de distribuição de anotações por arquivo (*by file*) que mostra a distribuição de etiquetas em arquivos individuais, em porcentagem [Fim da descrição].

Convém mencionar, ainda, que os trechos anotados podem ser pesquisados por meio dos códigos das etiquetas (5) e apresentados em típicas linhas de concordância¹⁰ (6).

Figura 7: Tela que exibe número de etiquetas por texto, ferramenta de busca e de concordanciador



Fonte: Autora (2025)

¹⁰ Ferramenta que encontra todas as ocorrências de um item de busca e as organiza, apresentando contextos à esquerda e à direita (Anthony, 2022).

[Descrição] Interface do *Louvain error editor*. Na parte superior esquerda há um quadrado em vermelho que destaca os arquivos, organizados em tabela, e, em duas colunas paralelas o número de anotações e *tokens* respectivos. Na parte superior direita, um quadrado verde sinaliza as quatro seções de busca: *tag*, *error*, *correction* e *sort by*. Na metade inferior há linhas de concordância com textos em inglês. Nos excertos há partes em vermelho—que marcam a parte anotada. À esquerda estão listadas as etiquetas. Há uma linha ao final da página que indica o número de resultados (159) encontrados [Fim da descrição].

Um corpus de aprendiz de tradução etiquetado a partir do sistema taxonômico do *TAS* torna possível, dessa maneira, categorizar uma dada população de estudantes em relação aos erros de maior ocorrência e suas respectivas proporções. Ademais, os dados quantitativos fornecidos pelo editor de erros fornecem uma visão estatística conferida pela Linguística de Corpus. Isto é, com a quantificação das etiquetas é possível saber a média de erros por texto, a distribuição de erros em um único texto ou em todo o corpus, a variância¹¹ dos dados, medidas essas que se tornam pontos de partida na implementação de processos corretivos.

As atividades relacionadas à etapa três — Extração de características relevantes do texto de partida por meio de ferramentas como *keyword*, *n-grams*, *collocate* — visam a identificar características representativas do texto de partida, os quais se tornam pontos de interesse para verificação no corpus de aprendiz. Essa verificação se fundamentaria em uma abordagem proposta por Marchi (2010) e advogada por Ancarno (2020) como um dos métodos de pesquisa em análise de discurso baseado em corpus, e consiste na investigação de:

1. Lista de palavras/ palavras-chave para identificar domínios semânticos centrais, isto é para estabelecer a temática principal; 2. Lista de colocados para explorar o comportamento textual dos termos-chave. Ou seja, para identificar padrões ou começar a identificar termos a partir dos resultados do corpus (lista de palavras-chave e lista de colocados); 3. Lista do concordanciador para explorar padrões prevalentes em contexto (Ancarno, 2020, p. 177, tradução nossa¹²).

Essa abordagem permite fazer um levantamento das características linguísticas representativas do texto que ensejou a tradução e tomar como ponto de partida seus traços salientes, buscando nas traduções de aprendizes padrões tradutórios e estratégias predominantes. Vejamos, na subseção seguinte, uma breve aplicação dessa metodologia experimental com um corpus de aprendiz composto de 9 *abstracts* traduzidos por estudantes do curso de graduação em Letras-Tradução (Inglês), da Universidade de Brasília, na disciplina “Prática de tradução de textos gerais (Português-Inglês)”.

4.1.2 Aplicação do método

Observando as orientações da subseção 4.1.1, tem-se:

¹¹ Indica a dispersão dos dados em relação à média. Uma alta variância indica que há diferença considerável entre os dados, o que sugere que há textos que apresentam mais erros que outros.

¹² No texto fonte: “1. *word lists/keyword lists to identify key semantic domains, that is, to establish central themes*; 2. *collocation lists to explore the textual behavior of key terms, that is, to identify patterns or start classifying terms from the corpus outputs (keyword and collocate lists)*; 3. *concordance lists to explore further dominant patterns in context*” (Ancarno, 2020, p. 177).

I. *Compilação do corpus de aprendiz*: Nesta etapa, foram coletados 9 *abstracts* traduzidos por estudantes do curso de graduação em Letras-Tradução (Inglês), da Universidade de Brasília, na disciplina de “Prática de tradução de textos gerais (Português-Inglês)”, no dia 20 de janeiro de 2025. O texto original, intitulado “A democratização do ensino superior por meio da educação à distância (EAD) no Brasil”, foi publicado pela revista Humanidades e Inovação, pertence ao gênero *abstract* e deveria ser traduzido conforme o *briefing* a seguir:

Quadro 1: *Briefing* para tarefa de tradução

Este *briefing* foi elaborado pela Assessoria Internacional da Universidade de Brasília e visa fornecer ao tradutor em potencial informações essenciais sobre o teste de tradução a ser realizado e garantir que o produto se alinhe com as expectativas do projeto.

Nome do projeto: Internacionalização da pesquisa brasileira

Língua de partida: Português do Brasil

Língua de chegada: Inglês

Número aproximado de palavras: 200-350

Prazo para entrega: 60 minutos após recebimento da tarefa

Objetivo: Selecione um *abstract* e traduza-o para a língua inglesa, se atentando à correção vocabular e terminológica, bem como às convenções de gênero e discursivas que caracterizam o *abstract*.

Público-alvo: Leitores proficientes de língua inglesa e pesquisadores da área de Humanidades interessados em tópicos como *ensino superior*, *educação à distância*, *educação musical*, *educação infantil*, *uso de ferramentas tecnológicas no ensino* e *educação na América Latina*.

Exigências: Abster-se de usar ferramentas de assistência à tradução automática ou de Inteligência Artificial.

Fonte: Autora (2025)

II. *Processamento do corpus com etiquetagem dos erros*: Os textos foram anotados conforme o manual *TAS* no ambiente do *Louvain error editor*, conforme procedimentos descritos na seção 4.1.

III. *Extração de características relevantes do texto de partida*: Essa etapa consiste em estabelecer os princípios funcionais e comunicacionais do texto de partida e as formas pelas quais eles se manifestam linguisticamente. Consiste também em compreender a estruturação retórica e discursiva do gênero a fim de considerar as melhores estratégias para que o texto circule e seja aceito na comunidade discursiva antevista. Isto é, as características linguístico-funcionais e estilísticas utilizadas no texto de partida podem ser replicadas na tradução? Busca-se, assim, entender como o texto funciona em seu contexto de partida, quais os públicos-alvo previstos no ambiente receptor, com vistas a averiguar, conseqüentemente, se há intersecção de atributos entre os gêneros, em termos interculturais.

No presente caso ilustrativo, entre as características terminológicas relevantes do texto de partida destacam-se:

- Educação à distância;
- Educação básica;
- Educação superior;
- democratização do ensino;
- modalidade de ensino;

Entre as características retórico-discursivas relevantes salientam-se:



superior por meio da Educação a Distância (EaD) no Brasil, destacando seu crescimento e o objetivo do estudo. Colunas B a E: Versões traduzidas para o inglês do texto original, seguindo a mesma segmentação. As traduções variam em estilo e precisão, e alguns segmentos contêm etiquetas de anotação do sistema TAS [Fim da descrição].

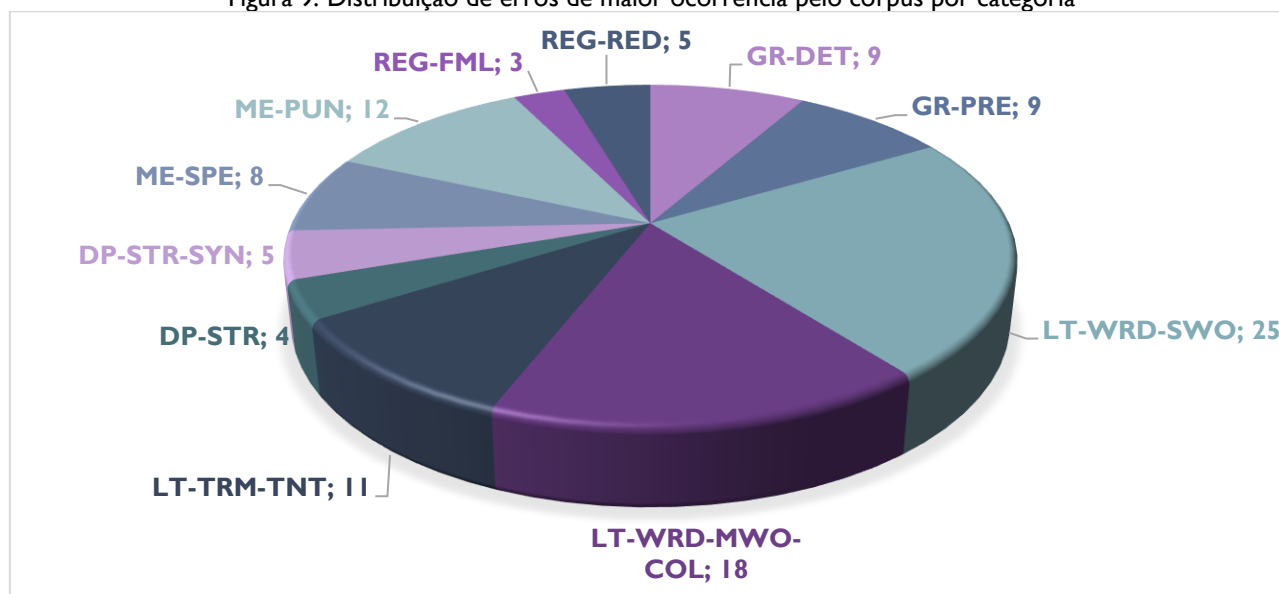
V. *Análise estatística*: O próprio ambiente do *Louvain Error editor* gerencia as estatísticas obtidas a partir das anotações em termos de anotações e desvios mais recorrentes e afins, conforme descrito na subseção 4.1.1.

Assim, depreendeu-se, por exemplo, que uma das anotações mais significativas, no caso em crivo, foi o desvio referente a LT-WRD-SWO, que diz respeito a usos inapropriados de itens lexicais únicos. Isto é, uma dada palavra, embora referencialmente correta, não seria a mais apropriada em termos expressivos, de registro, gênero ou discurso. Citam-se como exemplos:

- “Legality” em vez de “legal framework” → *Discussing the concept of Remote Learning and it’s [sic] legality in Higher Education;*
- “Function” em vez de “role” → *Analyze [sic] the function of professors and students in this mode of learning.*

Ademais, por meio dos dados obtidos na ferramenta, é possível gerar arquivos de saída que, se devidamente processados, geram gráficos informativos, que permitem ver os desvios mais frequentemente identificados.

Figura 9: Distribuição de erros de maior ocorrência pelo corpus por categoria



Fonte: Autora (2025)

[Descrição] Gráfico circular (pizza) dividido em 11 fatias coloridas, cada uma representando um tipo de erro na tradução. As fatias são organizadas em ordem decrescente de tamanho, com cores distintas e porcentagens correspondentes ao número de ocorrências [Fim da descrição].

Outrossim, um breve aprofundamento estatístico permite obter dados relevantes no que diz respeito à predominância dos erros na população e identificar a ocorrência isolada ou sistemática de desvios, por meio de dados como os da Tabela 1:

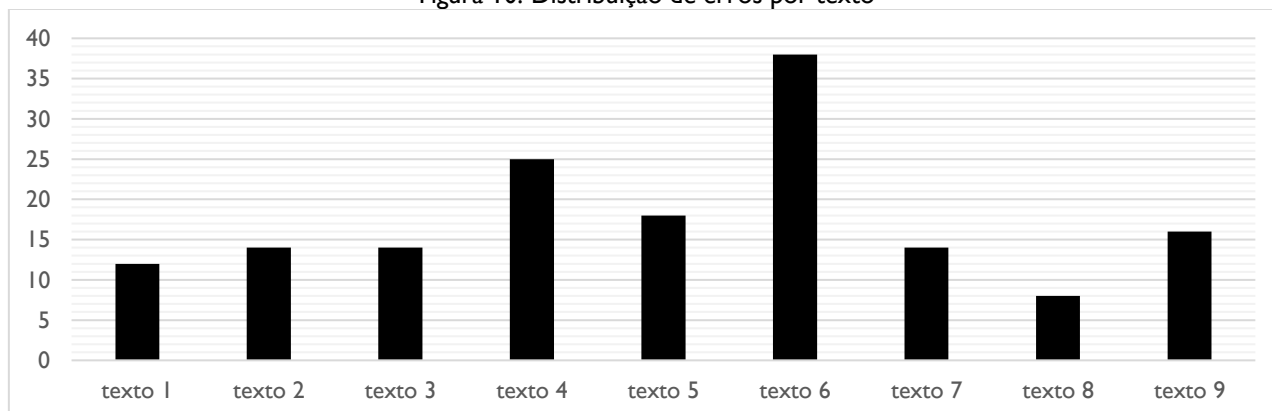
Tabela 1: Resumo de dados estatísticos relevantes

Total de anotações	159
Média de anotações por texto	17,67
Maior número de anotações em um texto	38
Menor número de anotações em um texto	8
Variância de erros por texto	70.67
Variância de erros por categoria	28.74

Fonte: Autora (2025)

Considerando que a média é uma medida influenciada por extremos, ela não mostra com acurácia se há predomínio de anotações em um dado texto. Identificou-se, por exemplo, que o texto seis apresenta 38 anotações, enquanto o texto oito, apenas 8 anotações. Isso indica, por exemplo que, embora haja necessidades pedagógicas gerais, há também demandas de ensino mais individualizadas. Outras medidas estatísticas também podem oferecer informações adicionais, como a variância, a qual indica a dispersão dos dados em relação à média. Uma variância elevada sugeriria, assim, que há textos que apresentam mais erros que outros, havendo, portanto, bastante heterogeneidade de anotações e, conseqüentemente, na qualidade dos textos, conforme vemos no gráfico a seguir:

Figura 10: Distribuição de erros por texto



Fonte: Autora (2025)

[Descrição] Gráfico de distribuição composto por nove barras verticais pretas, alinhadas da esquerda para a direita. No eixo vertical, os números variam de 0 a 40, indicando o número de anotações/erro. A barra referente ao texto 6 apresenta grande quantidade de erros, próximo a marca de 40. As demais são relativamente mais baixas [Fim da descrição].

A variância de erros por categoria, por sua vez, sugere que há uma dispersão moderada de erros entre as categorias, com erros mais recorrentes que outros, como é o caso do LT-WRD-SWO.

VI. Interpretação dos dados: Os dados obtidos nos procedimentos anteriores tornam-se subsídio para processos decisórios no âmbito do ensino-aprendizagem, já que os fatos explicitados pelas operações podem revelar necessidades de ensino antes não abrangidas pela ementa, exigindo, assim, uma recalibração dos objetivos de ensino e dos materiais utilizados. Ademais, a natureza dos desvios predominantes sugerirá as medidas remediadoras apropriadas. Ilustrativamente, foi constatado muitos erros lexicais, o que aponta para uma inobservância de questões como nuance,

expressividade e registro nas escolhas, chamando atenção para exigências de remediação sobre a natureza gradativa e contextual do significado. Findada essa abordagem preliminar, a seção seguinte tratará de algumas das principais aplicabilidades da análise computadorizada de erros de aprendiz e sua implicação do ensino de tradução.

5. Aplicações e discussão

A coleta e anotação de traduções de estudantes podem se tornar ponto de partida para muitas e variadas iniciativas na área de ensino e pesquisa em tradução (Granger & Lefer, 2020). A primeira aplicabilidade se refere ao uso das traduções anotadas para identificação de áreas de dificuldade em potencial e a consequente produção de exercícios para corrigi-las e /ou para a adoção de outras práticas pedagógicas pautadas em necessidades reais. Ademais, com o corpus de aprendiz é possível detectar a distribuição de um erro em uma dada população e verificar sua pontualidade ou sistematicidade o que, por sua vez, orienta a prática docente para uma recalibragem no programa e nos materiais e no percurso de ensino. Antevê-se, assim, uma multiplicidade de fins aos quais a análise computadorizada de erros pode ser aplicada, já que serve para compor avaliações diagnósticas, formativas—fornecendo retroação aos estudantes—, e somativas, ao fim de um período, com vias a informar o alcance de certos objetivos educacionais (Han, 2020).

Figura 11: Exemplos de formatos de exercícios produzidos no *software* Louvain error editor

1. Detailed tags

The urgent need for imagination [LS: During] _____ the last decade of the twentieth century, [GPP: it] _____ is [LS: easy] _____ to exemplify how science, industrialisation and technology have pervaded our everyday life.

Some will claim that the increasing number of factories has spoiled the [GNC: landscape's [LS: attraction] _____].

Others [LS: assert] _____ that imagination and dreams have definitely vanished from our modern society.

0, in the twentieth century, fears for the future of mankind [LS: have predominated] _____ the optimism about inevitable progress which was more typical of the nineteenth century.

All these examples point out that artists (that is to say people who give shape to their dreams and imagination through works of art) either play the role of catalysts for technological progress, or that of a [LS: railing] _____ against the uses and abuses of science and industrialisation.

1. Simplified tags

The urgent need for imagination [L: During] _____ the last decade of the twentieth century, [G: it] _____ is [L: easy] _____ to exemplify how science, industrialisation and technology have pervaded our everyday life.

Some will claim that the increasing number of factories has spoiled the [G: landscape's [L: attraction] _____].

Others [L: assert] _____ that imagination and dreams have definitely vanished from our modern society.

0, in the twentieth century, fears for the future of mankind [L: have predominated] _____ the optimism about inevitable progress which was more typical of the nineteenth century.

All these examples point out that artists (that is to say people who give shape to their dreams and imagination through works of art) either play the role of catalysts for technological progress, or that of a [L: railing] _____ against the uses and abuses of science and industrialisation.

2. No tags

The urgent need for imagination [SDuring] _____ the last decade of the twentieth century, [Sit] _____ is [Seasy] _____ to exemplify how science, industrialisation and technology have pervaded our everyday life.

Some will claim that the increasing number of factories has spoiled the [SLandscape's [Sattraction] _____].

Others [Sassert] _____ that imagination and dreams have definitely vanished from our modern society.

0, in the twentieth century, fears for the future of mankind [Shave predominated] _____ the optimism about inevitable progress which was more typical of the nineteenth century.

All these examples point out that artists (that is to say people who give shape to their dreams and imagination through works of art) either play the role of catalysts for technological progress, or that of a [Srailling] _____ against the uses and abuses of science and industrialisation.

Fonte: Granger et al. (2023b, p. 27)

[Descrição] Há 3 títulos que indicam os tipos de exercícios produzidos pelo editor de erros. Os títulos (*detailed tags*, *simplified tags*, *no tags*) estão em negrito. Abaixo de cada título os mesmos excertos em inglês são apresentados e há espaços em branco a serem preenchidos por alunos. No primeiro, os espaços em branco são antecedidos pela *tag*, que indica o erro, e o excerto errôneo, em vermelho ou laranja. No segundo título, o excerto também apresenta espaços em branco, os quais são antecedidos pela etiqueta simplificada e pela respectiva anotação. No último título, não há etiquetas e os espaços em branco são antecedidos apenas pelo excerto anotado antecedido do símbolo \$. Os excertos que apresentam erros leem-se: The urgent need for imagination _____ the last decade of the twentieth century, _____ is _____ to exemplify how science, industrialisation and technology have pervaded our everyday life. Some will claim that the increasing number of factories has spoiled the _____. Others



_____ that imagination and dreams have definitely vanished from our modern society. O, in the twentieth century, fears for the future of mankind _____ the optimism about inevitable progress which was more typical of the nineteenth century. All these examples point out that artists (that is to say people who give shape to their dreams and imagination through works of art) either play the role of catalysts for technological progress, or that of a _____ against the uses and abuses of science and industrialization [Fim da descrição].

A criação de materiais didáticos com base em um corpus de aprendiz pode adotar uma abordagem *corpus-driven* ou *corpus-based*. Na primeira, o conteúdo de aprendizagem tomará como base as dificuldades avaliadas e o desenvolvimento de materiais levará em consideração os erros produzidos por alunos; enquanto que na segunda, os recursos produzidos contêm exemplos do próprio corpus de aprendiz, com listas de concordância para identificação do desvio e para a sugestão de soluções alternativas. O *Louvain error editor*, por exemplo, contém a funcionalidade de produzir atividades baseadas em corpus de aprendiz (*corpus-based*) com exercícios de preenchimento de lacuna que orientam a atenção de estudantes aos tipos de erros cometidos para que possam corrigi-los.

A adoção de um sistema de anotação de erros como parte do processo avaliativo autoriza também uma visão mais informada acerca do estágio de competência dos estudantes, já que, embora a avaliação de competência tradutória não seja equivalente à avaliação da qualidade de tradução, esta é um reflexo daquela (Pokorn et al., 2019). Tendo em vista que sistemas de anotação como o *TAS*, por exemplo, abarcam uma série de categorias linguísticas, de gênero e aquelas relacionadas aos procedimentos tradutórios, fornecer retroação quanto à qualidade do texto também é facilitada já que uma tradução de alta qualidade é aquela em que:

A mensagem do texto fonte é transferida completamente ao texto-alvo, inclusive a denotação, conotação, nuances e estilo, em que o texto-alvo é escrito na língua-alvo com uso de gramática e ordem correta das palavras para a produção de um texto culturalmente apropriado e que, na maioria dos casos, é lido como se tivesse sido originalmente escrito por um falante nativo da língua-alvo para leitores da cultura-alvo (Koby et al., 2014, p. 416-417 citados por Verplaetse, 2022, p. 2009, tradução nossa¹³).

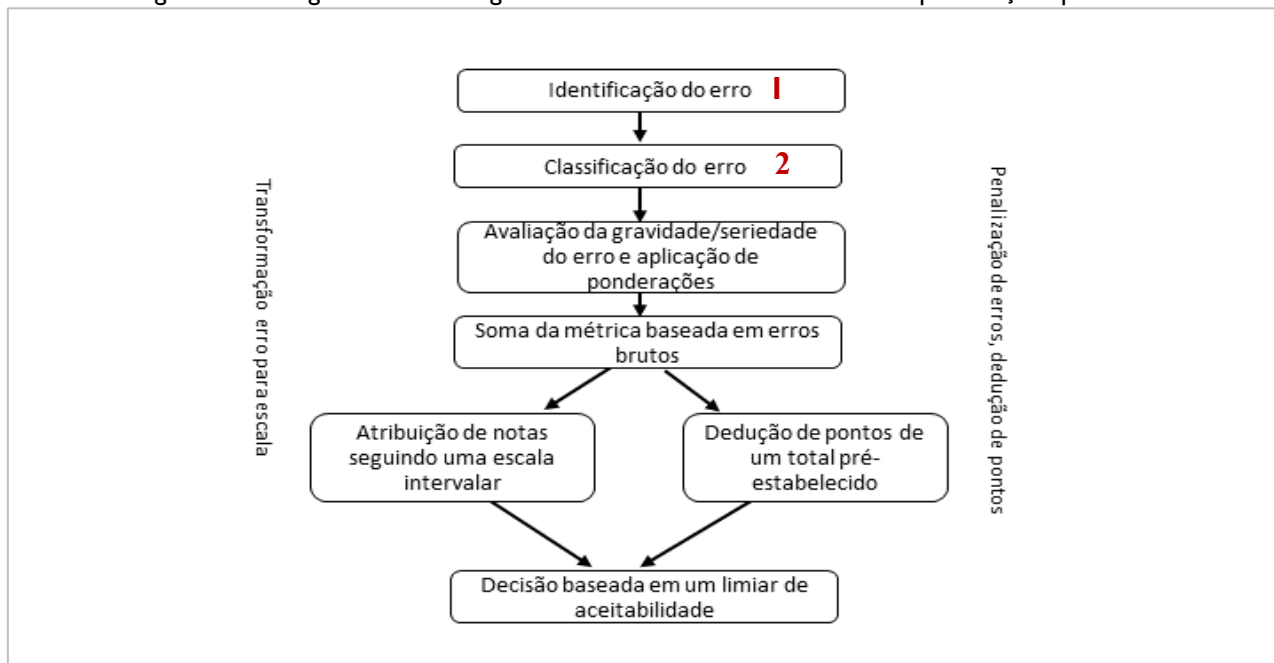
Nesse sentido, os parâmetros para delimitar o que constitui uma tradução de qualidade se tornam mais claros já que a identificação de erros, como distorção ou omissão não justificada, pode apontar para desvios que afetam esse construto. Ademais, quando associados a outros tipos de informações, como a de metadados, ou a ambientes diversos de avaliação, um corpus de aprendiz etiquetado para erros permite que se identifique, como apontam Castagnoli et al. (2011), além de desvios relacionados a palavras ou sequências lexicais específicas:

- a) Quais erros são mais comuns no corpus e as principais tendências tradutórias;
- b) Se há correlação entre frequência de erros e nível de conhecimento teórico sobre os Estudos da Tradução;
- c) Se erros terminológicos tendem a desaparecer em traduções cujos estudantes têm acesso a recursos como dicionários ou a *web*.

¹³ No texto fonte: “[...] the message embodied in the source text is transferred completely into the target text, including denotation, connotation, nuance and style, and the target text is written in the target language using correct grammar and word order, to produce a culturally appropriate text that, in most cases, reads as if originally written by a native speaker of the target language for readers in the target culture” (Koby et al., 2014, p. 416-417 citados por Verplaetse, 2022, p. 2009).

Importa ressaltar, contudo, que a análise de erros compõe apenas uma etapa do processo avaliativo, a qual pode ser adicionada outras abordagens a depender do objetivo avaliativo em questão. Isto é, se houver a exigência de que o desempenho de alunos seja quantificado, a análise de erros pode ser combinada com a aplicação de pesos(ponderações) ou dedução de pontos a depender da seriedade dos desvios e de suas respectivas frequências, e do alcance de um de limiar de aceitabilidade, que delimitaria aprovação ou reprovação, conforme esquematizado pelo modelo de Han (2020), adaptado abaixo, em que a anotação computadorizada de erros aqui advogada corresponde aos números um e dois:

Figura 12: Fluxograma de abordagens ao erro baseadas em escalas e em penalizações por erro



Fonte: A autora (2025), adaptado de Han (2020)

[Descrição] O fluxograma começa no topo com a etapa: Identificação do erro (1), que aponta para a próxima etapa: Classificação do erro (2). Em seguida, uma seta conduz à terceira etapa: Avaliação da gravidade/seriedade do erro e aplicação de ponderações. A partir dessa etapa, uma seta leva a: Soma da métrica baseada em erros brutos. Dessa etapa, o processo se divide em dois caminhos: “Atribuição de notas seguindo uma escala intervalar” e “Dedução de pontos de um total pré-estabelecido”. Ambos os caminhos convergem para a etapa final: Decisão baseada em um limiar de aceitabilidade. À esquerda do fluxograma, uma anotação vertical indica “Transformação do erro para escala”, referindo-se ao caminho de atribuição de notas. À direita, outra anotação vertical indica “Penalização de erros: dedução de pontos”, referindo-se ao caminho de dedução de pontos [Fim da descrição].

Esse tipo de combinação avaliativa também é possível ao se associar outras abordagens da Linguística de Corpus, como o uso de *corpora* de língua geral ou de especialização para verificar como as produções dos estudantes se comparam a textos efetivamente escritos em uma determinada língua e gêneros (veja Bowker, 2001), viabilizando a realização de estudos contrastivos. Nesse sentido, a relevância da implementação de formas e padrões de avaliação fica clara, principalmente se considerarmos que, se não implementados, “[...] a avaliação de tradutores e da garantia da qualidade de traduções continuará sendo arbitrária, sujeita a preferências pessoais, aos

caprichos do avaliador ou de arcabouços interpretativos, e às perspectivas burocráticas e medidas draconianas de docentes e avaliadores” (Darwish, 2001, p. 2, tradução nossa¹⁴).

Urge salientar, contudo, que a aplicação da análise de erros como componente da avaliação de traduções não é despida de barreiras, tendo em vista que a aplicação de novos métodos avaliativos pode encontrar resistência. Entre os dificultadores citam-se o ceticismo em relação à objetividade e aos métodos quantitativos, a onerosidade da prática da metodologia—a qual demanda familiarização com a taxonomia e certa literacia digital. Ademais, o foco excessivo no erro pode ser, de fato, um fator de desmotivação tendo em vista que, apesar da existência de uma tipologia, o fazer avaliativo tem, eminentemente, caráter subjetivo (Bittner, 2020), principalmente se se considerar que a anotação manual sofre com a baixa concordância entre avaliadores.

Considerações finais

Considerando que estudantes de tradução podem ser descritos como um tipo de aprendiz de língua, a coleta de suas produções escritas pode, como apontam Bowker e Bennison (2002, p. 504 citadas por Granger & Lefer, 2020) fornecer subsídios pedagógicos e de pesquisa valiosos.

A análise de erros enquanto parte do processo de avaliação de traduções, embora observada com receio por alguns por conta da negatividade com que “erros” geralmente são vistos, é, na verdade, um aspecto crucial para a identificação de pontos de melhoria no ensino, sendo essencial dentro de um quadro pedagógico de tradução. As informações obtidas por meio dos corpora de aprendizes podem orientar não apenas a prática docente, mas também a produção de materiais, fornecendo entendimentos acerca de estágios de aprendizagem e de pontos de dificuldade sistemáticos e individuais. Diferente das práticas passadistas de análises de erros descontextualizados e com uso de tipologias de erros pouco padronizadas, atualmente os desvios são investigados de forma contextualizada já que as ocorrências errôneas de um dado item lexical podem ser visualizadas no nível da sentença, do parágrafo e do texto (Granger, 2002) graças aos procedimentos da Linguística de Corpus, que dão primazia à uniformização e consistência de sistemas de etiquetagem.

Como pretendeu-se mostrar, os dados relativos à frequência, fornecidos pelas ferramentas de Linguística de Corpus, se mostram úteis para a identificação de áreas sistemáticas de dificuldade, as quais podem ser usados para a reestruturação de currículos e de percursos de ensino (Granger, 2002). Se usada como instrumento de avaliação diagnóstica, a análise de erros em corpus de aprendizes de tradução pode inclusive orientar o desenho do curso e tornar o processo de ensino-aprendizagem mais transparente e consistente. Ademais, se integrado a uma abordagem longitudinal, com o acúmulo e estudo de traduções realizadas em um dado período, é possível estabelecer parâmetros comparativos de evolução e propor retroações mais completas aos discentes.

Outrossim, a análise de erros possibilita a obtenção de informações sobre os construtos internos de estudantes (López, 2009) e sobre seu conhecimento teórico de tradução e de língua. Similarmente, a identificação de erros como o da “omissão” de excertos cuja supressão não é

¹⁴ No texto fonte: “[...] *translator assessment and translation quality assurance will always be haphazard and subject to the personal preferences and whims of the individual assessor or the interpretive frameworks, bureaucratic perspectives and draconian measures of educators and evaluators alike*” (Darwish, 2001, p. 2).

justificada, por exemplo, aponta para uma evitação dos problemas tradutórios e indica que certo progresso linguístico, transferencial, enciclopédico ainda precisa ocorrer.

Ao dar acesso à produção autêntica de estudantes, a Linguística de Corpus e a análise de erros se tornam abordagens cruciais para estudar como ocorre a progressão do ensino-aprendizagem, fornecendo ao pesquisador uma metodologia e arcabouço para o estudo da língua de aprendizes. Por apontar áreas de dificuldade dos estudantes, como gramática, léxico, discurso, a análise de erros se torna uma ferramenta de desenho pedagógico, permitindo o delineamento de ementas, planos de ensino e de materiais didáticos mais focados na remediação das dificuldades encontradas (López, 2009). Entretanto, para que os erros não sejam usados como o único componente em um processo avaliativo, a adoção de critérios mais funcionalistas sobre a adequabilidade de traduções e o respeito a um *briefing* de tradução, que simula o ambiente profissional em sala de aula, tal como sugerido por Nord (1996), se mostra um passo adequado rumo ao desenvolvimento de um ensino de tradução mais estruturado e orientado tanto pelas necessidades de aprendizagem, como mercadológicas.

Por fim, uma série de benefícios advém do estabelecimento de critérios e abordagens avaliativas claramente delineados: maior credibilidade do professor, maior aceitação dos estudantes— que perceberão que avaliação não se iguala a críticas desconstrutivas—, e percepção acerca das principais estratégias e processos de tradução em jogo.

Referências

- Al-Hadithy, T. M. (2015). The traditional vs. the modern translation classroom: A need for new directions in the UAE undergraduate translation programs. *Procedia: Social and Behavioral Sciences*, 192, 180–187. <https://doi.org/10.1016/j.sbspro>
- Ancarno, C. (2020). Corpus-Assisted Discourse Studies. In A. De Fina & A. Georgakopoulou (Eds.), *The Cambridge Handbook of Discourse Studies* (pp. 165–185). Cambridge University Press. <https://doi.org/10.1017/9781108348195.009>
- Anthony, L. (2022). What can corpus software do? In A. O'Keeffe & M. J. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (2nd ed., pp. 103–125). Routledge.
- Askari, I. M. (2018). Teaching methods in translation courses: Quality, relevance, and resources. *Research in English Language Pedagogy*, 7(2), 283–312. <https://doi.org/10.30486/relp.2019.665885>
- Bernardini, S. (2022). How to use corpora for translation. In A. O'Keeffe & M. J. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (2nd ed., pp. 485–498). Routledge.
- Bittner, H. (2020). *Evaluating the evaluator: A novel perspective on translation quality assessment* (1st ed.). Routledge.
- Bowker, L. (2001). Towards a methodology for a corpus-based approach to translation evaluation. *Meta*, 46(2), 345–364. <https://doi.org/10.7202/002135ar>
- Callies, M. (2015). Learner corpus methodology. In S. Granger, G. Gilquin & F. Meunier (Eds.), *The Cambridge Handbook of Learner Corpus Research* (pp. 35–56). Cambridge University Press.

- Castagnoli, S., Ciobanu, D., Kunz, K., Volanschi, A., & Kübler, N. (2011). Designing a learner translator corpus for training purposes. In N. Kübler (Ed.), *Corpora, language, teaching and resources: From theory to practice* (pp. 221–248). Peter Lang.
- Darwish, A. (2001). *Transmetrics: A formative approach to translator competence assessment and translation quality evaluation for the New Millennium*. RMIT University.
- De Sutter, G., Cappelle, B., De Clercq, O., Looock, R., & Plevoets, K. (2017). Towards a corpus-based, statistical approach to translation quality: Measuring and visualizing linguistic deviance in student translation. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 16, 25–39. <https://doi.org/10.52034/lanstts.v16i0.440>
- Díaz-Negrillo, A., & Domínguez, J. F. (2006). Error tagging systems for learner corpora. *Revista Española de Lingüística Aplicada*, 19, 83–102.
- Granger, S. (2002). A bird's-eye view of learner corpus research. In S. Granger, S. Petch-Tyson & J. Hung (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 3–33). <https://doi.org/10.1075/llt.6.04gra>
- Granger, S. (2003). Error-tagged learner corpora and CALL: A promising synergy. *CALICO Journal*, 20(3), 465–480. <https://doi.org/10.1558/cj.v20i3.465-480>
- Granger, S., & Lefer, M.-A. (2020). The multilingual student translation corpus: A resource for translation teaching and research. *Language Resources and Evaluation*, 54(4), 1183–1199. <https://doi.org/10.1007/s10579-020-09485-6>
- Granger, S., & Lefer, M.-A. (2021). *Translation-oriented annotation system manual (Version 2.0)*. Centre for English Corpus Linguistics.
- Granger, S., Paquot, M., Swallow, H., & Thewissen, J. (2023a). *UCLouvain Error Editor (Version 2)* [software]. Centre for English Corpus Linguistics, Université catholique de Louvain. <http://hdl.handle.net/20.500.12279/968>
- Granger, S., Swallow, H., & Thewissen, J. (2023b). *The Louvain Error Tagging Manual (Version 2.0)*. Louvain-la-Neuve: Centre for English Corpus Linguistics, Université catholique de Louvain.
- Granger, S., Swallow, H., & Thewissen, J. (2023c). *The UCLouvain Error Editor User guide - Version 2.0*. Louvain-la-Neuve: Centre for English Corpus Linguistics, Université catholique de Louvain.
- Han, C. (2020). Translation quality assessment: A critical methodological review. *The Translator*, 26(3), 257–273. <https://doi.org/10.1080/13556509.2020.1834751>
- Hönig, H. G. (1997). Positions, power and practice: Functionalist approaches and translation quality assessment. *Current Issues in Language and Society*, 4(1), 6–34. <https://doi.org/10.1080/13520529709615477>
- Igualada, M. T. (2016). Towards a new characterization of translation errors: From pre-symptomatic to post-symptomatic errors. *International Journal of Language and Applied Linguistics*, 2(1), 78–93.
- Kussmaul, P. (1995). *Training the Translator*. John Benjamins.
- López, W. C. (2009). Error analysis in a learner corpus: What are the learners' strategies? In P. Cantos Gómez & A. Sánchez Pérez (Eds.), *A Survey of Corpus-Based Research* (pp. 675–690). Asociación Española de Lingüística del Corpus.
- Marchi, A. (2010). The moral in the story: A diachronic investigation of lexicalised morality in the UK press. *Corpora*, 5(2), 161–189. <https://doi.org/10.3366/cor.2010.0104>

- Martínez Melis, N., & Hurtado Albir, A. (2002). Assessment in translation studies: Research needs. *Meta*, 46(2), 272–287. <https://doi.org/10.7202/003624ar>
- Nord, C. (1991). *Text Analysis in Translation: Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis*. Rodopi.
- Nord, C. (1996). El error en traducción: Categorías y evaluación. In A. Hurtado Albir (Ed.), *La enseñanza de la traducción* (pp. 91–107). Servei de Comunicació i Publicacions.
- Pokorn, N. K., Blake, J., Reindl, D., & Peterlin, A. P. (2019). The influence of directionality on the quality of translation output in educational settings. *The Interpreter and Translator Trainer*, 14(1), 58–78. <https://doi.org/10.1080/1750399X.2019.1594563>
- Rees, G. (2022). Using corpora to write dictionaries. In A. O'Keeffe & M. J. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (2nd ed., pp. 387–404). Routledge.
- Sager, J. C. (1989). Quality and standards: The evaluation of translations. In C. Picken (Ed.), *The Translator's Handbook* (pp. 91–102). ASLIB.
- Secară, A. (2005). Translation evaluation: A state-of-the-art survey. In *eCoLoRe/MeLLANGE Workshop Proceedings* (pp. 39–44). Universidade do Porto.
- Verplaetse, H. (2022). Translation quality in student specialized translation: The impact of corpus use. In S. Granger & M.-A. Lefer (Eds.), *Extending the Scope of Corpus-Based Translation Studies* (pp. 209–236). Bloomsbury Academic.

Anexo – Tabela descritiva das categorias do *Translation-oriented annotation system* de Granger e Lefer (2021)

Content	CT
Content: Distortion	CT-DST
Content: Distortion: Incomprehensible	CT-DST-INC
Content: Distortion: Incorrect meaning	CT-DST-INM
Content: Distortion: Illogical	CT-DST-ILL
Content: Distortion: Inexact meaning	CT-DST-INE
Content: Omission	CT-OMI
Content: Addition	CT-ADD
Content: Indecision	CT-IND
Content: Non-identification of source-text error/inconsistency	CT-NON
Culture	CL
Culture: Translated untranslatable/untranslated translatable	CL-UNT
Translation brief	TB
Translation brief: Non-compliance with reference materials	TB-REF
Translation brief: Formatting	TB-FOR
Grammar and Syntax	GR
Grammar and Syntax: Inflectional morphology	GR-INF
Grammar and Syntax: Tense/aspect	GR-TNS
Grammar and Syntax: Voice	GR-VOI
Grammar and Syntax: Word order	GR-WOR
Grammar and Syntax: Word category confusion	GR-WCO
Grammar and Syntax: Determiner	GR-DET
Grammar and Syntax: Preposition	GR-PRE
Grammar and Syntax: Pronoun	GR-PRO
Grammar and Syntax: Concord	GR-CCD



Grammar and Syntax: Conjunction	GR-CJX
Grammar and Syntax: Complementation	GR-COM
Grammar and Syntax: Complementation: Adjective	GR-COM-ADJ
Grammar and Syntax: Complementation: Noun	GR-COM-NOU
Grammar and Syntax: Complementation: Verb	GR-COM-VRB
Grammar and Syntax: Complementation: Adverb	GR-COM-ADV
Lexis and Terminology	LT
Lexis and Terminology: Word	LT-WRD
Lexis and Terminology: Word: Single word	LT-WRD-SWO
Lexis and Terminology: Word: Single word: Derivative	LT-WRD-SWO-DER
Lexis and Terminology: Word: Single word: Cognate	LT-WRD-SWO-COG
Lexis and Terminology: Word: Multiword	LT-WRD-MWO
Lexis and Terminology: Word: Multiword: Compound	LT-WRD-MWO-COP
Lexis and Terminology: Word: Multiword: Collocation	LT-WRD-MWO-COL
Lexis and Terminology: Word: Multiword: Idiom	LT-WRD-MWO-IDI
Lexis and Terminology: Term	LT-TRM
Lexis and Terminology: Term: Term translated by non-term/non-term translated by term	LT-TRM-TNT
Lexis and Terminology: Term: Terminological inconsistency	LT-TRM-TIN
Discourse and Pragmatics	DP
Discourse and Pragmatics: Cohesion	DP-COH
Discourse and Pragmatics: Cohesion: Linkword	DP-COH-LIN
Discourse and Pragmatics: Cohesion: Pronoun reference	DP-COH-PRF
Discourse and Pragmatics: Information structuring	DP-STR
Discourse and Pragmatics: Information structuring: Thematic progression	DP-STR-THE
Discourse and Pragmatics: Information structuring: Split/merge sentences	DP-STR-SPL
Discourse and Pragmatics: Information structuring: Syntactic preferences	DP-STR-SYN
Mechanics	ME
Mechanics: Spelling	ME-SPE
Mechanics: Punctuation	ME-PUN
Mechanics: Units, dates, numbers	ME-UDN
Register and Style	REG
Register and Style: Degree of (in)formality and register mismatch	REG-FML
Register and Style: Heavy	REG-HEA
Register and Style: Redundant	REG-RED
PLUS (meta-tag)	PLUS
Source Language Intrusion (meta-tag)	SLI
Distortion (meta-tag)	DIS
Error repetition (meta-tag)	REP

Notas editoriais

Contribuição de autoria

Concepção e elaboração do manuscrito: J. C. Valverde da Silva

Coleta de dados: J. C. Valverde da Silva

Análise de dados: J. C. Valverde da Silva

Discussão dos resultados: J. C. Valverde da Silva

Escrita - revisão e aprovação: J. C. Valverde da Silva



Conjunto de dados de pesquisa

Os dados aqui utilizados fazem parte de uma pesquisa de doutorado, atualmente em curso, no âmbito do Programa de Pós-Graduação em Linguística, da Universidade de Brasília.

Financiamento

Não se aplica.

Consentimento de uso de imagem

A autora declara que foi obtido o consentimento para o uso das imagens neste trabalho

Aprovação de comitê de ética em pesquisa

A pesquisa recebeu parecer favorável do CEP/CONEP, do Instituto de Ciências Humanas e Sociais da Universidade de Brasília. N° do CAAE (Certificado de Apresentação de Apreciação Ética): 82069324.0.0000.5540. N° do parecer: 7.159.784.

Conflito de interesses

Não se aplica.

Declaração de disponibilidade dos dados da pesquisa

Os dados desta pesquisa, que não estão expressos neste trabalho, poderão ser disponibilizados pela autora mediante solicitação.

Licença de uso

A autora cede à *Cadernos de Tradução* os direitos exclusivos de primeira publicação, com o trabalho simultaneamente licenciado sob a [Licença Creative Commons Attribution](#) (CC BY) 4.0 International. Essa licença permite que terceiros remixem, adaptem e criem a partir do trabalho publicado, atribuindo o devido crédito de autoria e publicação inicial nesta revista. Os autores têm autorização para assumir contratos adicionais separadamente, para distribuição não exclusiva da versão do trabalho publicada neste periódico (por exemplo: publicar em repositório institucional, em website pessoal, em redes sociais acadêmicas, publicar uma tradução, ou, ainda, republicar o trabalho como um capítulo de livro), com reconhecimento de autoria e publicação inicial nesta revista.

Publisher

Cadernos de Tradução é uma publicação do Programa de Pós-Graduação em Estudos da Tradução, da Universidade Federal de Santa Catarina. A revista *Cadernos de Tradução* é hospedada pelo [Portal de Periódicos UFSC](#). As ideias expressadas neste artigo são de responsabilidade de seus autores, não representando, necessariamente, a opinião dos editores ou da universidade.

Editores de seção

Andréia Guerini – Willian Moura

Revisão de normas técnicas

Alice S. Rezende – Ingrid Bignardi – João G. P. Silveira – Kamila Oliveira

Histórico

Recebido em: 22-12-2024

Aprovado em: 26-07-2025

Revisado em: 02-09-2025

Publicado em: 10-2025

