

A LEGENDAGEM SOB A ÓTICA DA SEMÂNTICA DE *FRAMES*: UMA ANÁLISE DE CONTRASTE SEMÂNTICO MULTIMODAL PORTUGUÊS-FRANCÊS EM OBRA AUDIOVISUAL

SUBTITLE FROM THE PERSPECTIVE OF FRAME SEMANTICS: AN ANALYSIS OF PORTUGUESE-FRENCH MULTIMODAL SEMANTIC CONTRAST IN AN AUDIOVISUAL WORK

DOI: 10.70860/ufnt.entreletras.e19563

**Raquel Figueiredo Roza¹
Maucha Andrade Gamonal²
Adriana Silvina Pagano³**

Resumo: Este trabalho analisa a construção de sentido na legendagem de português brasileiro para francês no curta-metragem “Eu Não Quero Voltar Sozinho”, com base na Semântica de *Frames*. A pesquisa identifica a similaridade semântica entre áudio e legendas e explora como informações textuais e visuais interagem na tradução interlinguística. Utiliza-se a metodologia de anotação multimodal da FrameNet Brasil adaptada para integrar elementos visuais. Os resultados destacam como as escolhas tradutórias se expressam através de *frames* e sugerem formas de avaliação da tradução audiovisual. O estudo contribui para o *dataset* multimodal ReINventa, voltado ao desenvolvimento de sistemas de inteligência artificial assistidos por *frames* semânticos.

Palavras-chave: Semântica de *Frames*; tradução audiovisual; legendagem; multimodalidade; similaridade semântica.

Abstract: This study analyzes meaning construction in the subtitling from Brazilian Portuguese to French in the short film “*Eu Não Quero Voltar Sozinho*”, based on Frame Semantics. The research identifies the semantic similarity between audio and subtitles and explores how textual and visual information interact in interlingual translation. It adopts the multimodal annotation methodology of FrameNet Brasil, adapted to integrate visual elements. The results highlight how translational choices are expressed through frames and suggest ways to assess audiovisual translation. The study contributes to the multimodal dataset ReINventa, aimed at developing frame-based artificial intelligence systems.

Keywords: Frame Semantics; audiovisual translation; subtitle; multimodality; semantic similarity.

¹ Bacharel em Tradução com Ênfase no Português-Francês pela Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG - Brasil. E-mail: figueiredorozar@gmail.com. ORCID: <https://orcid.org/0009-0008-2306-3238>.

² Doutora em Linguística pela Universidade Federal de Juiz de Fora (UFJF), Juiz de Fora, MG – Brasil. Residente de pós-doutorado na UFJF. E-mail: maucha.andrade@visitante.ufjf.br. ORCID: <https://orcid.org/0000-0002-5810-3913>.

³ Doutora em Letras pela Universidade Federal de Minas Gerais. Professora Adjunta da Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG – Brasil. E-mail: apagano@ufmg.br. ORCID: <https://orcid.org/0000-0002-3150-3503>.

Introdução

A legendagem de conteúdos audiovisuais entre línguas tipologicamente distintas, como o português e o francês, apresenta diferentes desafios, como a adaptação de expressões idiomáticas específicas de cada cultura, a escolha de registros linguísticos adequados ao público-alvo e a necessidade de conciliar a concisão textual com a manutenção do sentido original. Por mais que se pretenda preservar o objetivo comunicativo do áudio original, as diferenças das estruturas linguísticas e dos contextos socioculturais interferem na elaboração de legendas e, conseqüentemente, na construção de sentido do conteúdo audiovisual pelo público-alvo (Cintas e Remael, 2021b). Em especial, a legendagem envolve limitações técnicas, como o tempo de exposição das legendas, o que torna necessária a concisão textual, que, por sua vez, demanda escolhas estratégicas na seleção lexical e construção das frases.

Considerando esse cenário, esta pesquisa investiga a similaridade semântica entre áudio original em português brasileiro e legendas em francês do curta-metragem “Eu Não Quero Voltar Sozinho”⁴, de Daniel Ribeiro, produzido pela Lacuna Filmes. Para tanto, relacionam-se os estudos cognitivos da linguagem por meio da teoria da Semântica de *Frames* (Fillmore, 1982) à Tradução Audiovisual via legendagem (Cintas e Remael, 2021a,b,c,d). As contribuições de Czulo (2017) são importantes, pois instituem os *frames*, representações do conhecimento, de situações e da experiência humana como modelo de tradução no Modelo de Primazia do *Frame*. Além disso, a concepção de multimodalidade de Bateman *et al.* (2017) também é explorada, uma vez que se verifica como se dá a construção de sentido na integração dos modos comunicativos textuais (áudio e legenda) e visuais (cenas do curta) seguindo a anotação semântico-lexical no modelo FrameNet Brasil.

O trabalho se justifica pela ausência de estudos que investiguem a construção de sentido em objetos multimodais por meio da Semântica de *Frames* e dos Estudos da Tradução. No contexto da FrameNet Brasil, não há pesquisas que realizem o contraste das anotações semânticas entre o português brasileiro e o francês. Além disso, considerando a recente inserção da metodologia de anotação semântica multimodal (Belcavello *et al.*, 2022), ainda não existem estudos que explorem a integração de diferentes modos comunicativos na construção de sentido em legendas para esse par linguístico.

⁴ O curta-metragem foi legendado para diversas línguas, seu enredo aborda questões de inclusão e diversidade, tendo mais tarde dado origem ao longa-metragem “Hoje Eu quero Voltar Sozinho” dos mesmos produtores. O acesso ao curta é gratuito via youtube pelo canal da produtora de cinema Lacuna Filmes. Disponível em: <https://www.youtube.com/@HojeEuQueroVoltarSozinhoBR>. Acesso em: 05 de jan. 2025.

O estudo pormenorizado da construção de sentido de legendas de obras audiovisuais traz conhecimento das reformulações necessárias na tradução interlinguística e permite que se observem tanto as traduções lexicais como dinâmicas culturais e contextuais de cada idioma. No caso específico da legendagem português-francês associada ao estudo semântico das imagens do curta, uma análise contrastiva baseada em *frames* permite identificar como conceitos e experiências são estruturados em cada língua e em que medida a compreensão está ancorada nas imagens. No fazer da área, isso pode contribuir, por exemplo, para elaboração de práticas de legendagem mais eficazes e culturalmente coerentes.

Para realizar esse trabalho, a similaridade semântica é identificada por meio de uma métrica (cf. Samagaio, 2023) que avalia as anotações semânticas das sentenças advindas da transcrição do áudio e das legendas de um recorte do curta. Além das anotações semânticas dos textos (áudio transcrito e legenda), realiza-se a anotação semântica das imagens do curta que coincidam com o tempo de aparecimento dos áudios e das legendas. O interesse está em analisar como os *frames* evocados no texto se relacionam com os evocados nas imagens do curta, daí o trabalho estar vinculado ao domínio da anotação semântica multimodal. Com isso feito, as escolhas tradutórias para a elaboração das legendas em francês são avaliadas.

A metodologia de anotação semântica segue as diretrizes da FrameNet Brasil (Salomão, 2009; Torrent & Ellsworth, 2013), rede semântica estruturada por meio de corpus linguístico a partir de *frames*, unidades lexicais, elementos de *frames* e relações entre *frames*, conceitos que serão apresentados ao longo deste texto. No contexto da pesquisa, a metodologia é adaptada para anotações semânticas de imagens, permitindo a análise integrada dos *frames* evocados tanto pelos textos quanto pelos elementos visuais presentes no curta-metragem. Tanto as anotações linguísticas como as anotações de imagens (cf. Torrent *et al.*, 2022; Belcavello *et al.*, 2022) atuam como modelo de implementação da teoria Semântica de *Frames* de Fillmore (1982).

Os resultados desta pesquisa compõem os conjuntos de dados (*datasets*) multimodais da rede ReINventa (*Research and Innovation Network for Visual and Textual Analysis of Multimodal Objects*), iniciativa de pesquisa de anotação semântica de modos comunicativos diversos (texto verbal escrito e falado, imagens estáticas e vídeo, além de gestos). A proposta da rede é o desenvolvimento de tarefas de aprendizagem de máquinas e a criação de ferramentas de inteligência artificial assistidas por *frames* semânticos para diversos modos comunicativos. Além de ser útil ao propósito da rede, os resultados deste trabalho podem ser utilizados em

outras pesquisas que explorem a tradução audiovisual e o reconhecimento de padrões semânticos multimodais, por exemplo.

Quanto à estrutura do artigo, apresentam-se o referencial teórico (seção 1), que dialoga com várias áreas do conhecimento, a metodologia (seção 2), a descrição e análise dos resultados dos *frames* evocados na anotação semântica multimodal e da métrica de similaridade semântica (seção 3). Por fim, partilham-se a discussão (seção 4) e as considerações finais. Espera-se que, com este texto em mãos, o leitor tenha um panorama do trabalho realizado e possa, caso seja de interesse, contribuir com esta pesquisa, aprofundando-a ou desenvolvendo pesquisas correlatas a esta.

1 Referencial Teórico

Nesta seção, apresentam-se os pressupostos teóricos que embasaram a pesquisa: (i) a relação entre Estudos da Tradução e Linguística Cognitiva (LC) por meio dos conceitos postulados por Holmes (2000) e Rojo & Ibarretxe-Antuñano (2013); (ii) os estudos sobre legendagem desenvolvidos por Cintas e Remael (2021a,b,c,d); (iii) as reflexões de Bateman *et al.* (2017) sobre multimodalidade; (iv) a Semântica de *Frames* (Fillmore, 2009); e (v) o Modelo de Primazia do *Frame* (Czulo, 2017).

1.1 Estudos da Tradução

Estudos da Tradução é um termo criado e exposto por Holmes (2000) a fim de escolher uma nomenclatura que definisse a área de interesse. No artigo “*Nature of Translation Studies*”, Holmes propõe categorias que organizam o campo disciplinar em Puros (*Pure*) e Aplicados (*Applied*). Seguindo a classificação proposta, este trabalho se enquadra no campo dos estudos Puros Descritivos Orientados ao Produto, que analisa traduções existentes, e nos Estudos Aplicados para Assistência à Tradução, que se propõe a produzir pesquisas que possam ser aplicadas, como, por exemplo, ferramentas que auxiliam o processo tradutório.

Segundo Rojo e Ibarretxe-Antuñano (2013), os Estudos da Tradução se relacionam aos estudos cognitivos da linguagem. A relação se baseia principalmente na necessidade de estudos que explicam e descrevem os fenômenos tradutórios e na utilidade de materiais linguísticos contrastivos para as pesquisas linguísticas. Para a LC, a linguagem é parte da cognição humana e deve ser estudada em conjunto com outras faculdades mentais, como a memória, a atenção ou o raciocínio.

Para os Estudos da Tradução, o tradutor desempenha um papel fundamental, sendo os fatores que influenciam o produto tradutório:

o nível de domínio linguístico do tradutor e como isso está relacionado à (ii) direção da tradução (para dentro ou fora da língua dominante), (iii) à trajetória de desenvolvimento (ordem de aquisição das línguas e estágio de proficiência) e (iv) ao(s) item(s) ou tarefa(s) linguística(s) específica(s) em questão (Halverson, 2013, pág. 41, tradução nossa)

Para Rojo e Ibarretxe-Antuñano (2013), o tradutor é um mediador intercultural entre o texto de partida e o texto de chegada, e não apenas um especialista em duas línguas. O ato tradutório é compreendido como a seleção de conhecimentos através do processo cognitivo do tradutor. Os autores afirmam que, a partir do século XXI, houve o enfoque na cultura, na necessidade de dados empíricos, na importância do texto e nas metodologias experimentais.

A partir dessa mudança, os estudiosos da tradução passaram para uma análise baseada em *corpus*, a fim de investigar o processo tradutório, questões teóricas e até mesmo o uso para ensino. Rojo e Ibarretxe-Antuñano (2013) afirmam que tecnologias como o *software Translog*⁵ foram amplamente utilizadas para auxiliar as investigações relacionadas ao processo tradutório e concluem que a entrada no novo século trouxe conceitos e noções fundamentais para os estudos, como a noção de equivalência tradutória e a importância das questões culturais. No entanto, há noções que precisam ser redefinidas a fim de lidar com o dinamismo e a complexidade da tradução.

1.2 Legendagem

A legenda, segundo Cintas e Remael (2021a), é um texto que aparece na parte inferior da tela cujo objetivo é narrar o diálogo e outras informações sonoras e/ou visuais presentes em um produto audiovisual. A prática da legendagem é um processo complexo, uma vez que o objeto a ser traduzido se relaciona com outros elementos, como o som e a imagem presentes no produto audiovisual.

Segundo os autores, “as legendas não devem contradizer o que os personagens fazem ou dizem na tela, e a mensagem traduzida deve coincidir com a do discurso original” (Cintas; Remael, 2021a, p.4, tradução nossa). No entanto, os autores citam questões técnicas, definidas

⁵ Ferramenta desenvolvida para capturar e analisar as ações do tradutor durante a execução de uma tarefa de tradução. Originalmente elaborada para mapear a tarefa de tradução por meio do uso das teclas e do *mouse*, o *software* se aperfeiçoou com o passar dos anos e se tornou mais complexo com a inserção de rastreamento ocular. (Cf. Jakobsen, 2018)

pelos produtores, que interferem no produto tradutório. Algumas delas são a limitação de duas linhas para cada legenda, a limitação de caracteres, uma média de 37 caracteres com espaços incluídos, e o limite de seis segundos em tela para cada legenda, para, assim, evitar a tendência de releitura do texto. A sincronização é tida como parâmetro essencial, uma vez que a falta de sincronia pode incomodar o telespectador, que pode ter dificuldade para associar as legendas à fala das personagens.

Diante das limitações encontradas durante o processo de legendagem, Cintas e Remael (2021d) classificaram os processos linguísticos que caracterizam esse tipo de mídia, que são, principalmente, a reformulação do texto. Samagaio (2023) esquematiza os processos de reformulação, reproduzido no Quadro 1.

Quadro 1: Métodos de facilitação da legendagem - Adaptado de Cintas e Remael (2020d)

Processo	Descrição
Simplificação	Usar formas verbais mais simples do que as usadas no discurso original para economizar espaço e poder condensar mais informação.
Generalização de enumerações	Normalmente, generalizar enumerações pode remover algum traço estilístico do falante, mas economizar espaço (ex.: o falante pode dizer “Cães, lobos e raposas”, e as legendas podem dizer “canídeos”).
Sinônimos	Substituir uma palavra por um sinônimo com menos caracteres pode economizar espaço, mesmo que seja um consenso entre linguistas que não existem sinônimos perfeitos.
Tempos verbais simples	Usar tempos verbais simples em vez dos tempos verbais compostos pode economizar a quantidade de caracteres em uma linha de legenda.
Alteração no nível da sentença	Mudar sentenças negativas para sentenças positivas, transformar perguntas indiretas em diretas.
Simplificar modalizadores	Simplificar sentenças que usam modalizadores como “possivelmente” ou “provavelmente”, trocando-os por “é possível” ou “é provável”.
Alteração de discurso	Mudar o discurso direto para indireto.
Alteração de Sujeito	Alterar a estrutura da sentença, uma sentença como “isso pode ser visto por qualquer um.” Pode ser alterada por “É visível.”
Manipulação do Tema e do Rema	Inverter ou reformular a ordem das informações pode ser econômico para o legendador. Normalmente, o discurso falado insere o rema no início da sentença e o

	tema em seguida, porque querem dar mais relevância para as informações novas.
Transformar sentenças compostas em simples	Transformar uma sentença longa e composta que tomara várias legendas em diversas sentenças menores que aparecem como informações ininterruptas em várias legendas menores.
Alteração de voz verbal	Transformar uma sentença ativa em passiva, ou vice-versa pode ser outra estratégia interessante.
Uso de Dêiticos	Substituir uma sentença inteira por um pronome demonstrativo ou um dêitico.
Junção de sentenças	É o oposto do processo de transformação de sentenças compostas em simples, nesse caso, o legendador simplesmente junta duas sentenças pequenas em uma grande para que se possa aproveitar melhor o tempo das legendas.
Omissões	Na legenda, as omissões, ou apagamentos são inevitáveis. Muitas vezes, a repetição de uma palavra pode ser considerada redundante para o texto escrito, em outros momentos, as imagens podem apresentar as informações omitidas nas legendas.

Fonte: Samagaio (2023, p. 44-45)

Os autores ressaltam que o uso das reformulações textuais depende de cada produção e das necessidades encontradas em cada conteúdo, pois uma legenda curta para um personagem que fala demasiado ou uma legenda longa para um personagem que fala pouco pode causar estranheza no telespectador. Portanto, o legendador deve se tornar um especialista capaz de identificar o que é ou não necessário ao traduzir um conteúdo audiovisual.

1.3 Multimodalidade

Bateman *et al.* (2017) definem a multimodalidade como a relação entre diferentes modos comunicativos, que são “um recurso semiótico socialmente moldado e culturalmente dado para criar significado” (Kress, 2010, p. 79, tradução nossa). Kress (2010) cita a música, o gesto, a fala, a escritura e a imagem como alguns exemplos de modos utilizados na comunicação e representação. Desta forma, enquadram-se as produções audiovisuais como textos multissemióticos e multimodais.

Bateman *et al.* (2017) afirmam que a multimodalidade deve ser entendida como algo complexo, não apenas como a junção de formas comunicativas, mas como a inter-relação entre as formas. Portanto, ao realizar a tradução, o legendador trabalha com um dos modos comunicativos, o texto, mas deve considerar os elementos cinematográficos presentes, pois eles

fazem parte de um conjunto e, quando separados, tornam-se menos compreensíveis. A coesão intersemiótica está na associação entre fala e gesto.

Levando isso em consideração, Belcavello *et al.* (2020) argumentam que os elementos visuais podem evocar *frames* ou podem auxiliar o modo comunicativo verbal, construindo e contribuindo para a compreensão de significado. Desta forma, os autores frisam a importância da multimodalidade nas análises semânticas realizadas pela FrameNet Brasil.

1.4 Semântica de *Frames*

A Semântica de *Frames*, desenvolvida por Charles Fillmore (2009), busca compreender como o significado é construído a partir das palavras em uso, considerando os contextos nos quais são empregadas. Nesse sentido, a teoria parte do pressuposto de que as palavras evocam categorias cognitivas que refletem experiências vividas e compartilhadas em contextos socioculturais específicos. Fillmore compara o papel do linguista ao do antropólogo, interessado em investigar as categorias de conhecimento que estruturam a linguagem natural.

Fillmore (2009) exemplifica a importância das categorias na construção do sentido. A palavra “órfão” é um exemplo. Ela categoriza o grupo de crianças/adolescentes cujos pais faleceram. Esse entendimento está cercado por questões culturais, em que os pais são responsáveis pela criação dos filhos. No entanto, apesar de a palavra não possuir informação relativa à idade, supõe-se que uma pessoa com idade suficiente para se sustentar não deveria ser chamada de órfã. Além disso, a categoria inclui ser humano que seja passível de compaixão e cuidado. Desse modo, o autor afirma que a sociedade não admite como órfãos aqueles que causaram a morte dos próprios pais, e, por isso, deixam de ser percebidos como dignos de compaixão.

Nesse sentido, para o autor, o significado é moldado via *frames*, que são conceitos interligados em que, para entender um deles, é necessário compreender a estrutura na qual ele se enquadra. Um *frame* é ativado, quando evocado por uma Unidade Lexical (UL), que é o pareamento de um lema a um sentido específico. Quando o *frame* é evocado, os elementos da estrutura da qual ele faz parte são disponibilizados. Por exemplo, o *frame* que especifica uma transação comercial, necessariamente, envolve um comprador e uma mercadoria, além disso, pode haver material linguístico (ou visual no caso da anotação semântica multimodal, como será discutido adiante) que delimite outros elementos, como, por exemplo, dinheiro e finalidade.

Esses elementos são chamados de Elementos de *Frame* (EF) e conceptualizam o *frame*. Por meio da teoria, Fillmore (2009) desenvolve uma rede semântica destinada à pesquisa e

anotação semântica mediante *frames*, chamada FrameNet (FN)⁶. No Brasil, a FrameNet Brasil⁷ se expande para a anotação multimodal pelo projeto ReINVenTA.

1.5 Modelo de Primazia do *Frame*

O Modelo de Primazia do *Frame* foi apresentado no artigo “*Aspects of a Primacy of Frame Model of Translation*”, escrito por Czulo (2017). O autor propõe que o principal objetivo da tradução seja o de transferir significado. A transferência seria compreendida como a passagem de um texto na língua fonte para a língua alvo, e o significado, interpretado, de forma generalista, de duas maneiras, como aquilo apresentado por uma expressão, sua semântica, e/ou os efeitos de como é apresentada, sua função. No entanto, o significado de um texto pode sofrer alterações ao ser traduzido, e o autor se propõe a mapear essas alterações.

Para formular a teoria, Czulo (2017) utiliza a Gramática de Construções para explicar os aspectos relativos à forma, e a Semântica de *Frames*, para as questões relativas à semântica. A importância de ambas se dá pelo fato de presumirem que o contexto de cultura das línguas envolvidas em uma tradução influencia o processo e o produto tradutório. O *frame* é considerado pelo modelo como unidade de tradução, uma vez que se encontra em um nível abstrato da língua, em que são compartilhados por uma comunidade de falantes. Portanto, o tradutor deve encontrar na língua alvo o *frame* maximamente comparável àquele da língua fonte, ou seja, um *frame* que seja comparável ou próximo àquele da língua fonte. No entanto, a comparatividade máxima entre os *frames* considera que *frames* com o mesmo nome ou tidos como equivalentes podem ter conceptualizações diferentes a depender da cultura.

Parte-se da hipótese de coativação do *frame* (*co-activation hypothesis*), que compreende que outros *frames* são ativados quando um *frame* é evocado, havendo, portanto, uma relação entre eles. Essa relação entre os *frames* influencia a decisão do falante, que opta por selecionar *frames* diferentes daqueles disponibilizados originalmente, por exemplo, quando um recado dado é repassado a alguém. Segundo o autor, este caminho entre os *frames*, feito no momento da seleção, pode ser calculado em uma escala bidimensional em que a nulidade, quando não há semelhança ou proximidade entre os *frames*, ocorre por questões tipológicas, contrastivas e/ou culturais. Para calcular este caminho, foi utilizada a Métrica de Similaridade Semântica, apresentada a seguir.

⁶ Rede semântica FrameNet. Disponível em: <https://framenet.icsi.berkeley.edu>. Acesso em: 02 de jan. 2024.

⁷ Rede semântica FrameNet Brasil. Disponível em: <https://webtool.frame.net.br>. Acesso em 02 de jan. 2024.

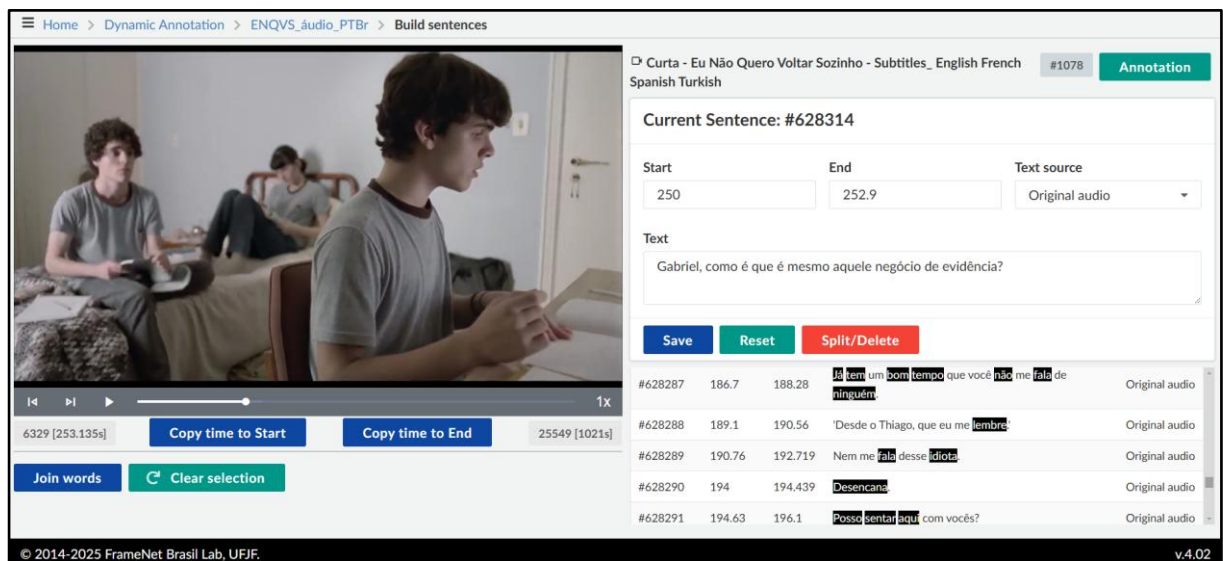
2 Metodologia

O material de análise utilizado faz parte do *corpus Audition*, composto por curtas-metragens animados e em *live action*. O *corpus* pertence ao *dataset* de anotação multimodal do projeto ReINventa. No recorte analisado, foram selecionados os cinco primeiros minutos e cinquenta segundos do curta-metragem brasileiro “Eu Não Quero Voltar Sozinho”, de Daniel Ribeiro, que tem dezessete minutos. Ao todo, realizou-se a anotação de 100 sentenças relativas ao francês e 109 sentenças relativas ao português, que geraram 107 sentenças alinhadas. A diferença numérica entre as anotações e o alinhamento ocorreu, pois o tradutor optou, em alguns casos, por agrupar duas ou mais sentenças do português em uma sentença no francês. Assim, durante o processo de alinhamento, algumas sentenças foram duplicadas para incluir aquelas a que faziam referência. No escopo do recorte proposto, foram definidas 155 *bounding boxes* na anotação de imagem dinâmica.

O curta se caracteriza pelo romance adolescente, que possui três personagens principais: Leonardo, protagonista portador de deficiência visual, Gabriel, o novo estudante, e Giovana, melhor amiga de Leonardo. A narrativa se desenvolve à medida que as personagens constroem uma relação de amizade, e, nesse processo, Leonardo e Gabriel se apaixonam. O curta explora os sentimentos dos personagens e suas inseguranças.

Com a escolha e delimitação do material de análise, foi realizada a anotação estrutural do curta, que consiste na sincronização das sentenças conforme as falas das personagens e a correção de possíveis erros gramaticais e/ou ortográficos gerados pela transcrição automática. Após isso, adicionam-se os lemas e *lexemas*, presentes nas sentenças extraídas do português e do francês, que não compunham a base de dados, e, então, passa-se à anotação de texto corrido e de imagem dinâmica. Por último, foi feito o cálculo da Métrica de Similaridade Semântica com as anotações de texto corrido.

Figura 6: Captura de Tela - Anotação Estrutural (Criação e Edição de Sentença)



Current Sentence: #628314

Start: 250 End: 252.9 Text source: Original audio

Text: Gabriel, como é que é mesmo aquele negócio de evidência?

Save Reset Split/Delete

ID	Start	End	Text	Source
#628287	186.7	188.28	Is tem um bom tempo que você não me fala de ninguém.	Original audio
#628288	189.1	190.56	Desde o Thiago, que eu me lembre.	Original audio
#628289	190.76	192.719	Nem me fala desse idiota.	Original audio
#628290	194	194.439	Desencana.	Original audio
#628291	194.63	196.1	Posso sentar aqui com vocês?	Original audio

© 2014-2025 FrameNet Brasil Lab, UFJF. v.4.02

Fonte: Captura de tela da Webtool 4.02⁸

Na Figura 6, há uma tela na qual o vídeo é exposto, e, ao lado, há espaço para o registro das sentenças e demarcação do tempo inicial e final. Há um recurso que detecta a pronúncia das palavras e as transforma em texto, isso se dá a partir de “*speech-to-text*”. O anotador pode fazer a seleção do conjunto que corresponde à sentença pronunciada pela personagem, que aparecerá no campo ao lado do vídeo (*text*). Neste campo, há espaço destinado à edição da sentença, classificação e marcador de início e fim da pronúncia ou, para as legendas, o tempo sobre a tela de cada frase.

O enfoque do *corpus Audition*, ao qual este curta está vinculado, é a anotação multimodal para Tradução Audiovisual Acessível. A fonte do texto pode advir de “*original audio*” (áudio original), “*subtitle*” (legenda), “*text overlay*” (texto sobreposto à imagem) e “*audio description*” (audiodescrição). Logo abaixo da edição das sentenças, há o agrupamento daquelas já editadas para serem anotadas. O passo seguinte é a anotação linguística, que pode ser lexicográfica e de texto corrido. Enquanto na primeira, escolhe-se um item lexical, e fazem-se as análises de várias sentenças sobre esse mesmo item, na anotação de texto corrido, todos os *frames* evocados em determinado texto são foco da análise linguística do anotador. Neste trabalho, é feita a anotação de texto corrido, que é realizada pelo *software* de anotação Webtool. A anotação linguística envolve, necessariamente, a aplicabilidade prática de três conceitos: *frames*, Unidades Lexicais (UL) e Elementos de *Frames* (EF).

⁸ Disponível em: <https://webtool.frame.net.br>. Acesso em: 02 de jan. 2024.

Os *frames* são evocados por Unidades Lexicais (ULs), que são o pareamento de um lema e um significado. Eles se estruturam por meio dos Elementos de *Frame* (EF), elementos que definem seus escopos. Podem ser Nucleares, necessários para a compreensão de um *frame*, e Não Nucleares, complementares à compreensão. Quando não são expressos na sentença, os EF Nucleares são registrados na forma de instanciação nula, que se subdivide em Definida, Indefinida e Construcional. A Instanciação Nula Definida ocorre quando o EF não está expresso na sentença, mas pode ser recuperável pelo leitor, ao contrário da Instanciação Nula Indefinida, na qual não é possível recuperar contextualmente o EF. Já a Instanciação Nula Construcional ocorre quando a construção da frase causa a falta do EF, como em frases no modo imperativo. Outra maneira pela qual o EF pode se apresentar é pela Incorporação, quando o EF está incorporado à Unidade Lexical.

A anotação registra tanto as informações semânticas quanto sintáticas do comportamento linguístico da unidade predicadora, diz-se ser uma anotação em três camadas: Elementos de *Frame*, Função Gramatical e Tipo Sintagmático. Para este trabalho, foi anotada exclusivamente a camada de Elementos de *Frame*, uma vez que o interesse é verificar a medida de similaridade semântica entre áudio e legenda.

O mapeamento das relações entre *frames* é outro registro importante feito pela FrameNet, por meio delas, verificam-se várias relações. São elas Herança (*Inheritance*), Uso (*Using*), *Subframe*, Precedência (*Precedes*), Perspectiva (*Perspective_on*), Causativo_de/Incoativo_de (*Causative_of/ Inchoative_of*). A relação de Herança se caracteriza pelo vínculo hierárquico entre um *frame* mãe e um *frame* filho. A relação de Uso pressupõe que, para compreender um *frame*, é necessário ter outro como plano de fundo. Na relação de *Subframe*, há um *frame* mãe e um *frame* filho em que o filho é um subevento do evento mais complexo representado pela mãe.

Assim como na relação de *Subframe*, a relação de Precedência representa a relação entre um *frame* mãe e um *frame* filho, no entanto, o *frame* filho é categorizado conforme o modo temporal do evento. Na Perspectiva, os *frames* compartilham o mesmo evento sob pontos de vista diferentes. Por fim, na relação Causativo_de, o sujeito é agente e causador de uma ação, ao contrário de Incoativo_de, em que o sujeito sofre a ação, ou seja, ele é o experienciador.

Destaca-se que a métrica de similaridade semântica usada recupera as informações das relações *frame a frame*⁹ para gerar o nível de similaridade semântica entre as sentenças. Na

⁹ Para mais informações sobre as relações *frame a frame*, consultar Ruppenhofer *et al.* (2016) e Samagaio (2023).

FrameNet Brasil, além das mencionadas, novas relações foram inseridas com o intuito de proporcionar refinamento semântico da rede e garantir que outros fenômenos linguísticos também fossem estudados, é o caso da relação metonímia e das relações qualia ternárias (Torrent *et al.*, 2022).

A Figura 7 mostra a sentença “A biblioteca eu já encontrei” com as ULs destacadas em cor preta para anotação linguística: biblioteca.n, já.adv, encontrar.v, já os Elementos de *Frame* são destacados com etiquetas coloridas.

Figura 7: Anotação da sentença #534407 do curta “Eu não Quero Voltar Sozinho”



Fonte: Captura de tela da WebTool 4.02¹⁰

Os EFs são disponibilizados conforme a definição do *frame* selecionado. Todas as sentenças do recorte escolhido para o francês e para o português foram analisadas dessa maneira. A Figura 8 exibe a sentença em francês correspondente “*Je sais où est la bibliothèque*”, em que as ULs anotadas são *savoir.v*, *où.adv*, *bibliothèque.n*.

Figura 8: Anotação da sentença #534288 do curta “Eu não Quero Voltar Sozinho”

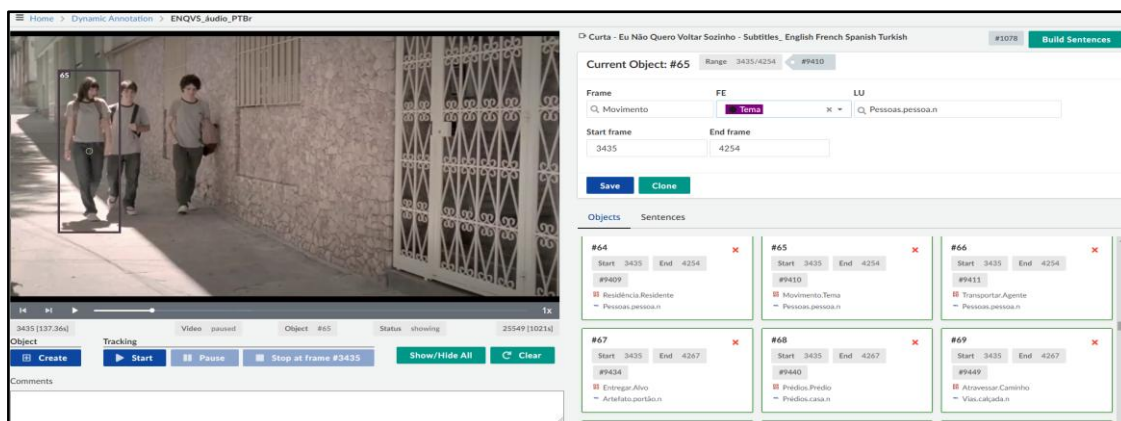
¹⁰Disponível em: <https://webtool.framenetbr.ufjf.br/index.php/webtool/annotation/main/annotation/628286/0/c>. Acesso em: 10 dez 2023.



Fonte: Captura de tela da WebTool 4.02¹¹

Feita a anotação de texto corrido, é realizada a anotação de imagem dinâmica. Esta anotação consiste em relacionar os *frames* evocados pela fala dos personagens aos elementos presentes na cena. Uma pasta foi criada na plataforma, denominada “ENQVS_áudio_PTBr”, com o vídeo em questão. A pasta direciona para a área na qual ocorre a anotação. Através da página, visualiza-se o vídeo, criam-se as *bounding boxes* e são apresentados os objetos criados, devidamente etiquetados. Há uma área para a anotação de *frames*, EFs, LUs e a duração de cada *bounding box*.

Figura 9: Anotação de imagem #9410 do curta “Eu não Quero Voltar Sozinho”



Fonte: Captura de tela da WebTool 4.02¹²

¹¹Disponível em: <https://webtool.framenetbr.ufjf.br/index.php/webtool/annotation/main/annotation/628286/0/c>. Acesso em: 10 dez 2023.

¹²Disponível em: <https://webtool.framenetbr.ufjf.br/index.php/webtool/annotation/main/annotation/628286/0/c>. Acesso em: 10 dez 2023.

As *bounding boxes* são caixas que delimitam a imagem e a relacionam ao *frame*, ao EF e à UL selecionada. Sua criação é feita por um comando que desenha sobre a tela uma caixa que seleciona o objeto. Após esse processo, sua marcação é mantida até sua desapareição por completo. A *bounding box* na Figura 9 demarca a jovem, anotada como Elemento de *Frame* Tema no *frame* Movimento e também é destacada a UL, no caso pessoa.n. Isso feito, o último passo da metodologia foi a Métrica de Similaridade Semântica.

A Métrica de Similaridade Semântica permite a avaliação quantitativa de similaridade, ou seja, a proximidade entre duas sentenças por meio dos *frames* evocados por ambas. A Métrica é uma combinação entre o algoritmo de *spread activation*, a métrica de Cosseno Soft (CS) e a métrica de avaliação da similaridade de cossenos de Czulo (2017).

O algoritmo de *spread activation* é uma métrica que espelha as sentenças e completa a lacuna existente entre elas com um valor zero, assim, o algoritmo atribui um valor para cada nó, que são os *frames* na rede semântica. O algoritmo modela um processo iterativo de propagação de energia de um ou mais nós para outros nós em um gráfico em três estágios: (i) pré-ajuste, (ii) espalhamento e (iii) pós-ajuste (Crestani, 1997). Antes da etapa de espalhamento, o valor energético de cada nó foi calculado durante a etapa de pré-ajuste. O decaimento de energia foi calculado para o valor do nó para que este valor fique dentro do intervalo [0,1]. O valor calculado é enviado para os nós vizinhos. Não foi utilizado pós-ajuste, pois o gráfico da FN é acíclico, e as hierarquias não abrangem diversos níveis (Viridiano *et al.*, 2022, p. 4, tradução nossa).

De acordo com Viridiano *et al.* (2020), o algoritmo de Cosseno Soft ocorre em três etapas: a formulação de uma tabela associada através de *frames*, criação de matrizes que associam as anotações de *frames* e a atribuição de valores para os níveis de similaridade entre cossenos das matrizes. O resultado é escalado em um intervalo de [0,1]. Segundo Samagaio (2023), os objetos que necessitam de maior energia para saltarem de um nó para outro recebem valores de similaridade mais baixos, ao contrário dos objetos que gastam menos energia, que recebem valores mais altos. Portanto, a proximidade entre os *frames* é calculada através da energia gasta nos saltos entre nós na rede semântica.

As sentenças do áudio em português e das legendas em francês foram alinhadas, somando 117 pares, que foram enviados à equipe de especialistas da computação da FrameNet Brasil a fim de extrair os dados referentes à métrica.

3 Descrição e Análise dos Resultados

As análises foram divididas em dois eixos: análise dos *frames* evocados pelas anotações e análise dos resultados referentes à Métrica de Similaridade Semântica. A divisão foi realizada porque a anotação de imagens dinâmicas não integrou a métrica de similaridade semântica, dado que sua metodologia difere da adotada para a anotação de texto corrido.

3.1 Resultado dos *Frames* Evocados nas Anotações

Ao todo, foram encontrados 635 *frames* evocados pelas anotações de texto corrido e de imagem dinâmica. Na Tabela 1, há a relação numérica da anotação.

Tabela 1: Relação dos *Frames* na anotação de texto e imagem

Anotação de Texto Corrido		Anotação de Imagem	Total
Português	Francês		
235	245	155	635
37%	38,5%	24,4%	100%

Fonte: Elaborado pelas autoras, 2025.

A Tabela 1 mostra que as anotações de texto possuem mais *frames* evocados que as anotações de imagem dinâmica; foram anotadas no texto corrido 235 *frames* para o português e 245 *frames* no francês. Nas anotações de imagem dinâmica, foram anotados 155 *frames*. A anotação de texto do francês compõe 38,5% dos *frames* anotados, seguida pela anotação de texto do português, com 37%, e a anotação de vídeo, que representa 24,4%.

Para a anotação de texto corrido do português, foram anotadas 109 sentenças, que evocaram 105 *frames* distintos. A Tabela 2 apresenta os *frames* que foram evocados mais de cinco vezes pelo texto em português.

Tabela 2: *Frames* Evocados pela Anotação de Texto Corrido do Português

<i>Frame</i>	Ocorrência	%
Negação	16	6,8%
Frequência	8	3,4%
Desejar	9	3,8%
Evento_desejável	5	2,1%

Dizer	5	2,1%
Conhecimento	5	2,1%
demais <i>frames</i>	189	80,4%
Total	235	100%

Fonte: Elaborado pelas autoras, 2025.

Através das relações apresentadas pela Tabela 2, percebe-se que o *frame* Negação¹³ é o mais frequente no excerto de texto analisado, com 16 ocorrências. Ele foi evocado pelas ULs não.adv, nunca.adv e sem.prep. Após, Frequência, com 8 evocações, Desejar, com 9, Evento_desejável, Dizer e Conhecimento foram evocados 5 vezes cada um.

Nas anotações do francês, foram encontrados 109 *frames* distintos nas 100 sentenças anotadas, a diferença entre o número de sentenças anotadas do português e do francês é devida às reformulações textuais que acarretaram o agrupamento ou junção de algumas sentenças. Os *frames* com mais de cinco evocações estão representados na Tabela 3.

Tabela 3: *Frames* Evocados pela Anotação de Texto Corrido do Francês

<i>Frame</i>	Ocorrência	%
Negação	17	6,9%
Capacidade_de_ação	7	2,8%
Partes_do_corpo	7	2,8%
Desejar	6	2,4%
Agir_intencionalmente	6	2,4%
Estímulo_mental_focado_no_experenciador	6	2,4%
Afirmar_ou_negar	6	2,4%
Cenário_da_educação	5	2%
Certeza	5	2%
Prédios	5	2%
Chegar	5	2%
Declaração	5	2%
Animais	5	2%

¹³ Por convenção, o nome do *frame* é registrado na fonte *Courier New*.

Demais <i>frames</i>	160	65,3%
Total	235	100%

Fonte: Elaborado pelas autoras, 2025.

Assim como na anotação do português, o *frame* Negação lidera a lista, com 17 ocorrências, ele foi evocado pela expressão *ne + verbo + pas.adv*. A lista segue com os *frames* Capacidade_de_ação e Partes_do_corpo, com 7 ocorrências; Desejar, Agir_intencionalmente, Estímulo_mental_focado_no_experienciador e Afirmar_ou_negar todos com 6 evocações; Cenário_da_educação, Certeza, Prédios, Chegar, Declaração, Desejabilidade e Animais, que foram evocados 5 vezes.

Nas anotações de vídeo, foram anotados 155 *frames* nos seis primeiros minutos do vídeo, destes 55 *frames* distintos. Na metodologia adotada neste trabalho, assim como em Belcavelo *et al.* (2022) e Dornelas *et al.* (2024), os registros de *frame* da anotação de imagem dinâmica se dão a partir da relação entre imagem e discurso proferido pelas personagens, portanto, a identificação das ULs corresponde às imagens apresentadas na cena a partir do áudio original. Na Tabela 4, há a relação entre os *frames* com maior ocorrência.

Tabela 4: *Frames* Evocados na Anotação de Imagem

<i>Frame</i>	Ocorrência	%
Partes_do_corpo	13	8,3
Pessoas	8	5,1
Conhecer_pessoas	7	4,5%
Conhecimento	7	4,5%
Atividade	6	2,1%
Familiaridade	6	3,8%
Desejar	5	3,2%
Transportar	5	3,2%
Atividade_pausar	5	3,2%
Demais <i>frames</i>	93	60%
Total	235	99,6%

Fonte: Elaborado pelas autoras, 2025.

O *frame* *Partes_do_corpo* foi o mais evocado na anotação de vídeo, com 13 ocorrências, ele representa 8,3% dos *frames* evocados. *Pessoas* foi evocado 8 vezes. *Conhecer_pessoas* e *Conhecimento* foram evocados 7 vezes. *Atividade* e *Familiaridade* foram evocados 6 vezes. *Desejar*, *Transportar* e *Atividade_pausar* foram evocados 5 vezes.

Ao analisar os *frames* frequentes em cada uma das anotações, foi feito um levantamento geral com os 635 *frames* nas anotações em conjunto, a fim de identificar aqueles mais frequentes em todo o material de análise. A Tabela 5 mostra a ocorrência e porcentagem dos seis *frames* mais frequentes nas anotações de texto corrido e de imagem em conjunto.

Tabela 5: *Frames* Evocados nas Anotações de Texto Corrido e de Imagem

<i>Frame</i>	Ocorrência	%
Negação	33	5,1%
<i>Partes_do_corpo</i>	24	3,7%
Desejar	20	3,1%
Conhecimento	15	2,3%
Prédios	12	1,8%
Opinião	12	1,8%
Demais <i>frames</i>	525	82,6%
Total	235	99,6%

Fonte: Elaborado pelas autoras, 2025.

Assim como nas anotações de texto corrido, o *frame* *Negação* inicia a lista com 33 ocorrências. *Partes_do_corpo*, que foi frequente na anotação das imagens do curta, obteve 24 evocações. *Desejar*, que se apresentou nas listas das três anotações, foi evocado 20 vezes. *Conhecimento* esteve na lista da anotação de texto corrido do francês e na anotação de imagem dinâmica, ao todo, foram encontradas 15 evocações. *Prédios* ocorreu 12 vezes e se esteve na lista da anotação de texto corrido do francês. Por fim, *Opinião* foi evocado 12 vezes na soma das anotações.

Através das análises, foram encontradas semelhanças e divergências entre os *frames* mais frequentes nas anotações. Em cada levantamento dos *frames* mais recorrentes, identificou-

se um ou mais coincidentes com o levantamento geral. Concluída essa etapa, passou-se à análise dos resultados obtidos com a aplicação da Métrica de Similaridade Semântica.

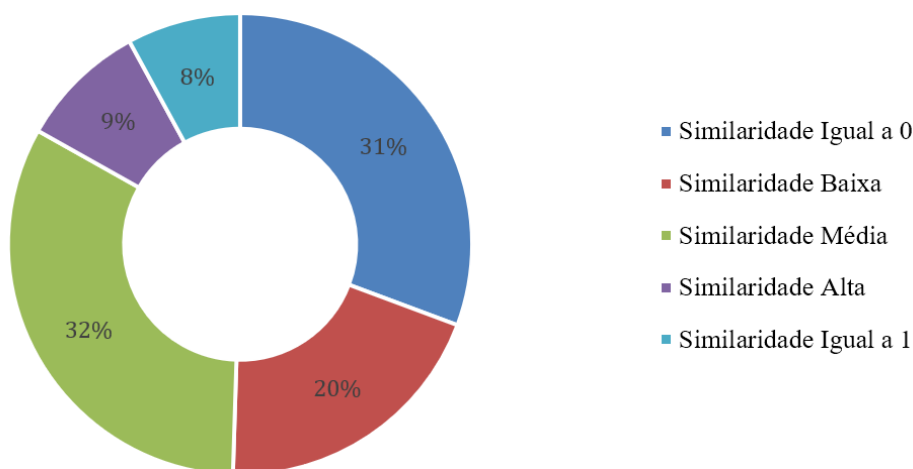
3.2 Resultados da Métrica de Similaridade Semântica

A Métrica foi realizada a partir do resultado das anotações de texto corrido do português e do francês. A anotação de imagem dinâmica não integrou a métrica devida à metodologia utilizada para realizar as anotações, que se difere da metodologia de anotação de texto corrido.

Foram anotados 117 pares de sentenças agrupados em cinco grupos: Similaridade Igual a 0, Similaridade Baixa (entre 0 e 0.41), Similaridade Média (entre 0.41 e 0.69), Similaridade Alta (entre 0.69 e 0.99) e Similaridade Igual a 1. No entanto, alguns grupos de sentenças receberam valor negativo (-1), casos em que não foi possível realizar o cálculo da métrica pela falta de sentença correspondente, estes casos agregaram o grupo de similaridade Nula. O Gráfico 1 mostra a distribuição dos grupos de sentenças conforme a pontuação de similaridade semântica.

O Gráfico 1 representa 100% dos grupos de sentenças analisados, ao todo, 101 grupos de sentenças. Os grupos de similaridade Média e Igual a 0 representam a maioria do material de análise, 32% e 31% do material, respectivamente. Em seguida, o grupo de Similaridade Baixa compõe 20% do material de análise. Os grupos com similaridade Alta e Igual a 1 representam 9% e 8%, respectivamente.

Gráfico 1: Resultados da Métrica de Similaridade Semântica em Porcentagem



Fonte: Elaborado pelas autoras, 2025.

[pas^{Negação}], n'est-ce pas?

As sentenças possuem um *frame* semelhante evocado, que é Conhecimento, evocado pelas ULs *saber.v* e *savoir.v*. No entanto, as demais ULs evocaram *frames* distintos, e, apesar de terem uma ligação próxima na rede semântica, como *Desejar* herdeiro de *Emoção_com_foco_no_experienciador*, houve o salto entre os nós, o que reduziu a pontuação de similaridade semântica. O valor final da Métrica é a média da pontuação encontrada através do espelhamento das sentenças, isso possibilitou que os pares de sentenças que evocam *frames* de maneira desigual atingissem valores mais altos pela alta similaridade entre os *frames* correspondentes. Dessa forma, há gradação da pontuação de similaridade entre os grupos com similaridade Alta, Média e Baixa.

Na Similaridade Média, foram encontrados grupos de sentenças que se assemelham ao grupo de Similaridade Alta. No entanto, a correspondência entre os *frames* não foi suficiente para elevar a pontuação, uma vez que *frames* distintos e distantes ou que não se relacionam na rede semântica foram evocados pelos pares. Um exemplo está nas sentenças #692093 e #627782.

692093 É só tirar a calça.

627782 J'enlève mon pantalon et le problème sera fini.

692093 É só [tirar^{Remover}] a [calça^{Vestuário}].

627782 J'[enlève^{Causar_movimento}] mon [pantalon^{Vestuário}] et le [problème^{Dificuldade}] sera [fini^{Processo_fim}].

Apesar de os pares de sentenças evocarem o *frame* *Vestuário*, *tirar.v* e *enlever.v* evocam *Remover* e *Causar_movimento*, o que gerou um maior salto entre os nós. Além disso, as demais ULs, *problème.n* e *finir.n*, foram adicionadas por uma segunda parte da sentença que não corresponde ao original. Possivelmente, o tradutor optou por adicionar essa descrição para explicitar que se tratava de um problema. Essa adição gerou uma desigualdade entre os *frames* evocados, o que, por consequência, reduziu a pontuação da métrica.

O grupo Similaridade Baixa é composto, principalmente, por pares de sentenças que evocam *frames* distantes e evocação desigual de *frames*, como nas sentenças 692104 e 627718.

No exemplo, as sentenças evocam *frames* completamente distantes na rede semântica, o que aumentou exponencialmente o salto entre os nós, reduzindo a pontuação de similaridade.

692104	Parar de brincadeira.	627718	Ça suffit.
692104	[Parar _{Atividade_pausar}] de [brincadeira _{Atividade}].	627718	Ça [suffit _{Suficiência}].

Através das análises das sentenças, houve reformulações e omissões, seguindo os postulados de Cintas e Remael (2021d), como a omissão de certas ULs, o que influenciou as evocações, e, por consequência, o resultado da métrica. Também há casos em que o legendador optou por expandir a sentença e explanar sobre elementos previamente citados. Como no exemplo #692093/#627782 (É só tirar a calça./ *J'enleve mon pantalon et le problème sera fini.*), exposto anteriormente, em que o legendador adicionou uma oração à sentença e, por consequência, mais *frames* foram evocados do que a sentença original.

Do lado oposto, no par de sentença #628300/#627768 (E você, foi bem?/*Et toi?*), o tradutor optou por simplificar a sentença, o que omitiu a UL *bem.adv*. Essas reformulações no texto traduzido ajustam o texto à língua alvo, no entanto, geraram menores pontuações de similaridade semântica, uma vez que houve alteração da quantidade de ULs nas sentenças.

4 Discussão

A discussão se inicia com os dados relativos aos *frames* evocados pelas anotações e parte para os dados da Métrica de Similaridade Semântica. Os *frames* mais frequentes nas anotações dialogam entre si. Todavia, o *frame* Negação, o mais frequente nas anotações de texto corrido, não foi encontrado na anotação de imagem dinâmica. Isto ocorreu, pois as representações de negação foram anotadas no vídeo por meio de *Afirmar_ou_negar*, uma vez que as entidades que afirmam ou negam são pessoas. Na metodologia de anotação adotada, as entidades são o foco da demarcação das *bounding boxes*, desse modo, no *frame* mencionado, um falante afirma ou nega um tópico ou uma mensagem.

Por outro lado, *Partes_do_corpo* foi representado treze vezes na anotação de vídeo e representa o *frame* mais presente na anotação. Isso ocorre, pois, quando um *frame* é evocado, os objetos que representam esse *frame* são anotados. Através destes exemplos, é possível perceber a diferença metodológica entre as anotações de texto e imagem.

Seguindo para a Métrica de Similaridade Semântica, apesar de os grupos de similaridade Baixa e Nula corresponderem a 51% do material de análise, este fator não interferiu na compreensão, uma vez que outros elementos, visuais e sonoros integraram a produção e auxiliaram a compreensão na garantia de manutenção do objetivo comunicativo do curta.

Foram encontrados dois tipos de sentenças longas, ou seja, sentenças com maiores quantidades de caracteres, aquelas originárias do português e aquelas produzidas pelo tradutor através da junção de sentenças curtas na língua fonte. As sentenças originárias do português, que foram traduzidas para o francês como uma ou mais sentenças curtas, não tiveram pontuação baixa de similaridade. No entanto, sentenças longas no francês, criadas a partir de sentenças curtas do português, obtiveram pontuações de similaridade mais baixas. Essas sentenças sofreram processos de reformulações e omissões (Cintas; Remael, 2021d), o que provocou a distância ou a falta de semelhança entre os *frames* evocados pelos pares correspondentes. Por outro lado, sentenças curtas em português que se mantiveram curtas no francês compõem, de modo geral, os grupos de similaridade Igual a 1 e Alta.

Por lidar com o par linguístico português-francês, houve casos em que o tradutor optou por reformular a sentença ou omitir informações a fim de seguir as regras gramaticais da língua alvo e as orientações para facilitação da legendagem. Essas mudanças no texto se mostraram necessárias para adequar a legenda ao contexto de cultura da língua para a qual foi traduzida.

Considerações Finais

A presente pesquisa demonstrou que a Semântica de *Frames*, associada ao Modelo de Primazia do *Frame* e à Métrica de Similaridade Semântica, pode ser um recurso eficaz para analisar relações semânticas em contextos de tradução interlinguística e multimodal, especialmente no par português-francês. Os resultados confirmaram que, na legendagem, a construção de sentido não ocorre apenas pelo texto, mas na interação entre elementos verbais, visuais e sonoros, reforçando a premissa de Bateman *et al.* (2017) de que os modos comunicativos devem ser interpretados de maneira integrada.

A análise revelou que as escolhas tradutórias são essenciais para adequar as legendas ao contexto sociocultural da língua alvo. Mesmo quando a similaridade semântica foi baixa ou nula, outros elementos multimodais garantiram a manutenção do sentido geral do curta-

metragem. Esse achado é relevante para os Estudos da Tradução, pois demonstra que a tradução audiovisual deve ser vista como um processo de diálogo contínuo entre texto e imagem.

O agrupamento das sentenças segundo a Métrica de Similaridade Semântica mostrou que sentenças curtas em português mantiveram uma alta similaridade ao serem traduzidas para o francês, enquanto sentenças longas ou fragmentadas apresentaram uma menor correspondência de *frames*, devido a processos de reformulação e omissão (Cintas; Remael, 2021). Assim, a análise baseada em *frames* evidenciou que as diferenças culturais e linguísticas influenciam diretamente as estratégias tradutórias e a similaridade semântica.

Além de atuar no enriquecimento dos bancos de dados da FrameNet Brasil, a pesquisa contribuiu significativamente ao aplicar a metodologia de anotação semântica multimodal em um estudo comparativo de tradução audiovisual. A integração de anotações textuais e visuais revelou como os *frames* “maximamente comparáveis” oferecem um recurso útil para avaliar traduções e sugerem novas possibilidades de análise em pesquisas futuras.

Por fim, espera-se que este trabalho incentive o desenvolvimento de estudos que combinem Semântica de *Frames*, Tradução Audiovisual e anotação semântica multimodal, contribuindo tanto para a criação de práticas tradutórias mais eficazes quanto para o aprimoramento e a criação de sistemas de inteligência artificial, como as previstas pela rede ReINVenTa.

Agradecimentos

A pesquisa apresentada neste artigo atuou como forma de colaboração à ReINVenTA (Rede de Pesquisa e Inovação em Visão e Análise de Objetos Multimodais). A ReINVenTA é fomentada pela FAPEMIG, por meio do financiamento RED 00106/21, e pelo CNPq, por meio dos financiamentos 408269/2021-9 e 420945/2022-9. Informamos que Dra Maucha Gamonal é bolsista de pós-doutorado da CAPES (88887.015648/2024-00) do Programa PIPD e Dra. Adriana Pagano é bolsista de Produtividade em Pesquisa do CNPq (313103/2021-6) e bolsista do Programa FAPEMIG de fomento à internacionalização das ICTMGs. Registramos e agradecemos o auxílio do Dr. Ely Matos, pesquisador da FrameNet Brasil, que gerou a medida de similaridade semântica entre áudio e legenda, bem como o levantamento dos *frames* anotados em cada modalidade estudada.

Referências

- BATEMAN, Jhon; WILDFEUER, Janina; HIIPPALA, Tuomo. Introduction: the challenge of multimodality. In. BATEMAN, Jhon, WILDFEUER, Janina., HIIPPALA, Tuomo. *Multimodality: Foundations, Research and Analysis a Problem-Oriented Introduction*. Berlin/Boston: Walter de Gruyter GmbH, 2017. p. 7-21.
- BELCAVELLO, Frederico; VIRIDIANO, Marcelo; COSTA, Alexandre Diniz da; MATOS, Ely Edison da Silva; TORRENT, Tiago Timponi. Frame-Based Annotation of Multimodal

Corpora: Tracking (A)Synchronies in Meaning Construction. *Proceedings of the International FrameNet Workshop 2020: Towards a Global, Multilingual FrameNet*, Marseille, p. 23-30, mai., 2020. Disponível em: <https://aclanthology.org/2020.framenet-1.4/>. Acesso em: 07 jan. 2025.

BELCAVELLO, Frederico; VIRIDIANO, Marcelo; MATOS, Ely; TORRENT, Tiago Timponi. Charon: A FrameNet Annotation Tool for Multimodal Corpora. *Proceedings of the 16th Linguistic Annotation Workshop (LAW-XVI) within LREC2022*. Marseille, p. 91-96, jun. 2022, p. 91-96. Disponível em: <https://aclanthology.org/2022.law-1.11/>. Acesso em: 08 jan. 2025.

CINTAS, Jorge Díaz; REMAEL, Aline. Reconceptualizing subtitling. In: CINTAS, Jorge Díaz; REMAEL, Aline. *Subtitling: Concepts and Practices*. Oxon/New York: Routledge, 2021a. p. 1-31.

CINTAS, Jorge Díaz; REMAEL, Aline. The semiotics of subtitling. In: CINTAS, Jorge Díaz; REMAEL, Aline. *Subtitling: Concepts and Practices*. Oxon/New York: Routledge, 2021b. p. 91-117.

CINTAS, Jorge Díaz; REMAEL, Aline. Spatial and temporal features. In: CINTAS, Jorge Díaz; REMAEL, Aline. *Subtitling: Concepts and Practices*. Oxon/New York: Routledge, 2021c. p. 91-117.

CINTAS, Jorge Díaz; REMAEL, Aline. The Linguistic of Subtitling. In: CINTAS, Jorge Díaz; REMAEL, Aline. *Subtitling: Concepts and Practices*. Oxon/New York: Routledge, 2021d. p. 146-177.

CRESTANI, Fabio. Application of spreading activation techniques in information retrieval. *Artificial Intelligence Review*, [S. l.], v. 11, p. 453-482, 1997. DOI: <https://doi.org/10.1023/A:1006569829653>. Disponível em: <https://link.springer.com/article/10.1023/A:1006569829653#article-info>. Acesso em: 08 jan. 2025.

CZULO, Oliver. Aspects of a primacy of frame model of translation. In: CZULO, Oliver; HANSEN-SCHIRRA, Silva; HOFMANN, Sascha. *Empirical modelling of translation and interpreting*, Berlin: Language Science Press, 2017, p. 465-490.

DORNELAS, Leticia Grossi; GAMONAL, Maucha Andrade; PAGANO, Adriana Silvina. Análise semântica de audiodescrição em curta-metragem: uma abordagem multimodal a partir da Semântica de Frames. *Domínios de Lingu@gem*, Uberlândia, v. 18, p. 1-30, dez. 2024. DOI: <https://doi.org/10.14393/DLv18a2024-66>. Disponível em: <https://seer.ufu.br/index.php/dominiosdelinguagem/article/view/75272>. Acesso em: 08 jan. 2025.

FILLMORE, Charles J.; Semântica de Frames. *Cadernos de Tradução*, Porto Alegre, n. 25, p. 25-54, jul-dez., 2009.

HALVERSON, Sandra L. Implications of cognitive Linguistics for Translation Studies. In: ROJO, Ana; IBARRETXE-ANTUÑANO, Iraide. *Cognitive Linguistics and Translation: advances in some theoretical models and applications*. Berlin/Boston: Walter de Gruyter GmbH, 2013, p. 33-73.

HOLMES, James S. The Name and Nature of Translation Studies. In: VENUTI, Lawrence. *The translation studies reader*, London/New York: Routledge, 2000. p.172-185.

- JAKOBSEN, Arnt; MAIA. Usando o Translog para registrar os acionamentos de teclas e mouse e rastrear os movimentos oculares de tradutores. *Revista Graphos*, João Pessoa, Edição Especial, p. 99-121, 2018. Disponível em: <https://periodicos.ufpb.br/index.php/graphos/article/view/40075>. Acesso 08 jan. 2025.
- KRESS, Gunther. Mode. In: KRESS, Gunther. *Multimodality: a social semiotic approach to contemporary communication*. Oxon/New York: Routledge, 2010. p. 79-102.
- ROJO, Ana; IBARRETXE-ANTUÑANO, Iraide. Cognitive Linguistics and Translation Studies: Past, present and future. In: ROJO, Ana; IBARRETXE-ANTUÑANO, Iraide. *Cognitive Linguistics and Translation: advances in some theoretical models and applications*. Berlin/Boston: Walter de Gruyter GmbH, 2013, p. 3-30.
- RUPPENHOFER, Josef; ELLSWORTH, Michael; PETRUCK, Miriam; JOHNSON, Christopher; SCHEFFCZYK, Jan. *FrameNet II: Extended Theory and Practice*. 2016.
- SALOMÃO, Maria Margarida Martins. *FrameNet Brasil: um trabalho em progresso*. *Calidoscópio*, São Leopoldo, v. 7, n. 3, p. 171-182, 2009. Disponível em: <https://www.redalyc.org/articulo.oa?id=571561888002>. Acesso em: 02 jan. 2025.
- SAMAGAI, Mairon Morelli. *A Permanência Semântica na Tradução Audiovisual: uma análise da legendagem pela ótica da Semântica de Frames*. 2023. 184f. Dissertação (Mestrado em Letras) – Universidade Federal de Juiz de Fora, Juiz de Fora, 2023. Disponível em: <https://repositorio.ufjf.br/jspui/handle/ufjf/16136>. Acesso em: 08 jan. 2025.
- VIRIDIANO, Marcelo; TORRENT, Tiago Timponi; CZULO, Oliver; MATOS, ALMEIDA, Arthur Lorenzi; MATOS, Ely Edison da Silva; BELCAVELLO, Frederico. The Case for Perspective in Multimodal Datasets. Proceedings of the 1st Workshop on Perspectivist Approaches to NLP. *NLPerspectives*. Marseille, p. 108-116, jun., 2022. DOI: <https://doi.org/10.48550/arXiv.2205.10902>.
- TORRENT, Tiago Timponi; ELLSWORTH, Michael J. Behind the Labels: Criteria for Defining Analytical Categories in FrameNet Brasil. *Veredas – Revista de Estudos Linguísticos*, Juiz de Fora, v. 17, p. 44-65, 2013. Disponível em: <https://periodicos.ufjf.br/index.php/veredas/article/view/25403>. Acesso em: 08 jan. 2025.
- TORRENT, Tiago; MATOS, Ely Edison da Silva Matos; BELCAVELLO, Frederico; VIRIDIANO, Marcelo; GAMONAL, Maucha, Andrade; COSTA, Alexandre Diniz da Costa; MARIM, Mateus Coutinho. Representing context in FrameNet: a multidimensional, multimodal approach. *Front Psychol.*, [S. l.], v. 13, e838441, 2022. DOI: <https://doi.org/10.3389/fpsyg.2022.838441>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/35444591/>. Acesso em: 08 jan. 2025
- WEBTOOL 4,02 [FNBr/server4/webtool402]. FrameNet Brasil. Disponível em: <https://webtool.framenetbr.ufjf.br/index.php/webtool/annotation/main/annotation/628286/0/c>. Acesso em: 08 jan. 2024.

Recebido em 11 de janeiro de 2025
Aceito em 30 de maio de 2025