

DOI: 10.14393/DL29-v11n2a2017-8

Modelagem topológica do perfil sistêmico-funcional de manuais de instrução no espaço gramatical

Modeling topologically systemic-functional profiling of handbooks in the grammar space

Francieli Oliveira*

Thaís Campolina**

Giacomo Figueredo***

RESUMO: Embasado nos pressupostos da Linguística Sistêmico-Funcional e da Linguística de *Corpus*, o presente artigo investiga o modo pelo qual os sistemas gramaticais da oração contribuem para a produção de significados do texto. Mais especificamente, objetiva construir um modelo dinâmico da gramática que descreva a produção textual no manual de instrução em português brasileiro. Uma vez que a língua é constituída de probabilidades de ocorrência de funções que compõem subconjuntos, aos quais se denominam aqui, é possível descrevê-los à medida que as funções linguísticas são descritas. Para a modelagem, foi compilado um *corpus* de manuais de instrução e analisado segundo os sistemas principais da oração. Os resultados mostram como foi possível produzir um modelo, o qual aponta que as funções que realizam o manual prototipicamente são processo material, relacional e verbal, modo declarativo e jussivo e tema default e perspectiva. Além disso, configura-se como uma das formas de generalização das probabilidades maiores de escolhas dos sistemas gramaticais para esse gênero textual.

PALAVRAS-CHAVE: Modelagem sistêmico-funcional. Perfilação sistêmica. Linguística Computacional. Linguística de *Corpus*. Manual de instrução.

ABSTRACT: This paper adopts Systemic Functional Linguistics and Corpus Linguistics to investigate clause grammar systems and their contributions to text meaning-making. More specifically, it aims at building a dynamic model of grammar describing text production in handbooks in Brazilian Portuguese. Language is a probabilistic system of functions and these probabilities form subsets of textual genres. As the language functions are described, genres can be described and modelled in order to understand how functions build their meaning-making. A handbook corpus was compiled and profiled for the main clause grammar functions. The model shows that the prototypical handbook functions are material, relational and verbal process, declarative and jussive mood, and default and perspective theme. Accordingly, it generalizes the most frequent probabilities for choices in grammar systems for the handbook.

KEYWORDS: Systemic functional modeling. Systemic profiling. Computational Linguistics. Corpus Linguistics. Handbooks.

* Mestranda em Linguística Aplicada. Universidade Federal de Minas Gerais (UFMG).

** Graduada em Letras. Universidade Federal de Ouro Preto (UFOP).

*** Doutor em Linguística Aplicada. Departamento de Letras. Universidade Federal de Ouro Preto (UFOP).

1. Introdução

Campos disciplinares vinculados à Linguística e à Ciência da Computação, dentre eles a Linguística Computacional e a Inteligência Artificial, vêm crescentemente demandando contribuições mais efetivas dos estudos da língua para o Processamento de Linguagem Natural (PLN) (MATTHIESSEN; BATEMAN, 1991). Esse fato tem impulsionado a Linguística Aplicada a produzir conhecimento para oferecer modelos confiáveis da produção linguística, que não se baseiem apenas em regras formais de gramáticas livres de contexto; mas, de outra forma, tomem como parâmetro de processamento a Compreensão das Línguas Naturais (*Natural Languages Understanding*) (WINOGRAD, 1972; HALLIDAY, 2005).

De modo complementar, tem sido incrementada, no âmbito da Linguística Aplicada, a necessidade de implementação do processamento de línguas naturais na interação entre o humano e o computador (WINOGRAD, 2006), incorporando a máquina às suas práticas de pesquisa e aplicação. Dentre essas demandas, destaca-se a busca por modelos que extrapolem a ordem da oração, em particular voltando-se para a estrutura de textos e, por conseguinte, de gêneros textuais (MATTHIESSEN, 2015).

Situando-se nesse contexto, este artigo visa contribuir com soluções para as demandas do PLN relativamente à Linguística Aplicada. Dessa forma, tem por objetivo oferecer ao PLN um modelo linguístico dos sistemas gramaticais que realizam o gênero textual “manual de instrução”, pautado pelos pressupostos da Linguística de *Corpus* (BERBER SARDINHA, 2000; VIANA, 2011) e pelo arcabouço sistêmico-funcional (HALLIDAY, 2002; HALLIDAY; MATTHIESSEN, 2014).

Por conseguinte, este trabalho analisa as formas pelas quais a gramática da oração constrói a produção de significado do manual de instrução. Mais especificamente, produz um modelo capaz de descrever a maneira pela qual se dão as seleções sucessivas de opções nas redes dos sistemas (*system networks*) e, conseqüentemente, como essa dinâmica explica a organização do texto. A partir desse modelo, visa-se apresentar uma descrição gramatical do manual de instrução contribuindo para a descrição de gêneros textuais.

A Linguística de *Corpus* compreende que o estudo do funcionamento da língua deve ser realizado de forma empírica através de um *corpus* composto por textos falados e/ou escritos de forma espontânea (VIANA, 2011). Biber (2010) propõe que o estudo dos registros de uma língua deve ser realizado com base em um *corpus* composto por textos pertencentes ao gênero, para que seja possível investigar os padrões que o definem. Como a língua é um sistema de

probabilidades (HALLIDAY, 1991; BERBER SARDINHA, 2000), esses padrões apresentam um subconjunto de probabilidades de ocorrência de funções que fazem parte da construção de significados da língua. Isso significa que um gênero textual pode ser modelado exatamente desta maneira: um subconjunto de probabilidades de funções (HALLIDAY, 2002).

A Linguística Sistêmico-Funcional (LSF) (cf. HALLIDAY, 2002; HALLIDAY; MATTHIESSEN, 2014; MARTIN, 1992; MATTHIESSEN, 1995, dentre outros) compõe a base teórica e analítica deste artigo. Em primeiro lugar, essa escolha se deu pelo fato de essa teoria conceber a língua como um sistema de recursos que permite a representação, encenação e construção da experiência de mundo e das relações sociais dos falantes por meio de unidades de significado organizadas como configurações de funções, ou ‘textos’. Somando-se a isso, na qualidade de teoria abrangente, a LSF proporciona um arcabouço de descrição e análise do sistema linguístico como um todo, em seus diferentes estratos, de forma que a gramática pode ser examinada tanto em relação aos elementos que a expressam quanto em relação ao texto que organiza (HALLIDAY, 2010). Dessa maneira, a análise gramatical é capaz de revelar como se organizam os aspectos fundamentais dos gêneros textuais.

Sendo a produção de significado do manual de instrução o objeto de estudo deste artigo, cabe apontar que, do ponto de vista sistêmico-funcional, o significado é a propriedade da língua de relacionar funções organizadas sistemicamente, de forma que ao contraste entre elas se atribua um *valeur*¹ (SAUSSURE, 2006). Assim, o conceito de ‘função’ depende tanto dos recursos de produção de significado quanto da forma como a língua se organiza (HALLIDAY 2008). O modelo da língua ou o modelo de um gênero textual consiste em um conjunto de sistemas, a partir dos quais as relações entre os recursos geram as funções. Assim, toda e qualquer análise, descrição e aplicação sistêmico-funcionais são derivadas, em última instância, da dimensão sistêmica do sistema linguístico – denominada ‘dimensão do eixo’ (*dimension of axis*) (MARTIN, 2013).

A metodologia e a análise presentes neste artigo, portanto, procuram derivar seus protocolos de realização igualmente da dimensão do eixo. Para tanto, utilizam o *software* UAM CorpusTool (O’DONNELL, 2008) para análise semiautomática do *corpus* e, em seguida, a metodologia da perfilação sistêmica (FIGUEREDO, 2014), a qual estabelece distâncias topológicas entre as funções sistêmicas e seu movimento no espaço gramatical. A perfilação

¹ O termo *valeur* é adotado pela LSF e se conserva a palavra original em francês, mesmo quando empregado em outras línguas (cf. MARTIN, 2013).

gramatical – ou a aplicação da perfilação sistêmica à gramática – possibilita criar um modelo da produção de significado nos textos, de modo a descrever e prever seu funcionamento (OLIVEIRA, 2015).

Neste artigo, a definição de ‘modelagem’ parte da LSF e implica a criação de um modelo linguístico para futura aplicação, que pode variar desde áreas como o ensino e a análise do discurso até a Linguística Computacional, em estudos que envolvem o Processamento de Linguagem Natural (PLN). Dessa maneira, a modelagem proposta neste artigo se configura como uma resposta positiva à demanda da Linguística Aplicada por aplicações sistêmicas derivadas da dimensão do eixo, as quais envolvam o estudo de sistemas, escolhas e funções na constituição de textos/tipos de texto (MAHBOOB; KNIGHT, 2010) incluindo sua variação de frequências e probabilidades de conformação ao gênero textual (HALLIDAY, 1991; LEMKE, 1993).

Especificamente neste artigo, a modelagem sistêmico-funcional de perfilação é realizada em um *corpus* de manuais de instrução, tendo também como meta construir um modelo dinâmico da gramática que descreva a produção textual no manual de instrução em português brasileiro. Com isso, este trabalho apresenta, com base na perfilação gramatical, os recursos linguísticos utilizados para compreender os significados construídos no manual de instrução.

Ao final, este artigo demonstra como o arcabouço da LSF e Linguística de *Corpus* permite não somente analisar e descrever o gênero textual manual de instrução, como também, criar um modelo que aponte as probabilidades linguísticas que caracterizam esse texto. Uma vez que o modelo possui potencial de implementação computacional, ele pode contribuir para a interface entre o PLN e a Linguística Aplicada.

2. Pressupostos teóricos

Halliday (2002) aponta que, na tentativa de compreender melhor a língua, os falantes tentam o tempo todo estabelecer padrões, os quais são somente criados mediante as diferenças existentes na língua. Sendo assim, os significados são compreendidos por meio do contraste entre funções, ou *valeur* (GEASON, 1965; SAUSSURE, 2006). Pensando nisso, a descrição linguística consiste em conhecer a língua a fim de se estabelecer uma explicação para o seu funcionamento, de forma que essa explicação represente a modelagem linguística. Nesse sentido, Halliday (2002) afirma que um modelo da língua deve contemplar seus diferentes

contextos para diferentes propósitos, pois não é somente a base para a compreensão da língua, pelo modelo também é possível enxergar as informações relevantes sobre outros aspectos linguísticos.

Para modelar a língua, primeiramente, é necessária a sua descrição. Nesse sentido, a LSF oferece o arcabouço para a investigação de sistemas que trazem as possibilidades de escolhas de uma língua (HALLIDAY et al., 1964), assim como o seu significado e sua diferença com relação às outras escolhas (SAUSSURE, 2006). Por meio da organização sistêmica, na dimensão do eixo, é possível descrever, contribuir com a teoria linguística e modelar a língua.

Para a construção de um modelo da produção textual, toma-se a LSF como teoria para especificamente abordar três questões importantes (HALLIDAY, 2002; HALLIDAY, 2008): (i) a generalização da descrição de um *corpus* específico para o gênero textual; (ii) o grau de variação entre textos distintos e a acomodação de sua variabilidade no modelo; e (iii) a aplicabilidade do modelo para necessidades reais do uso da língua.

Como forma de atender a essas três questões, adota-se a modelagem topológica da gramática. O mapa topológico é a representação espacial das funções de um determinado sistema gramatical (MARTIN; MATTHIESSEN, 1991; FIGUEREDO, 2014). As funções são derivadas das opções (*features*) em um sistema. Por exemplo, o sistema de MODO possui em um primeiro nível duas opções, indicativo e imperativo. Destas opções se derivam as funções gramaticais de Indicativo e Imperativo.

Além disso, quando duas opções são derivadas de um mesmo sistema são denominadas formas agnatas. Neste caso as opções são, por um lado, semelhantes por serem derivadas de um mesmo sistema; mas, por outro lado, são opostas pelo fato da seleção de uma opção implicar na exclusão da outra opção. No sistema de MODO, por exemplo, indicativo e imperativo são agnatas. Com isto, são semelhantes por ambas serem derivadas do sistema de MODO e, concomitantemente, são opostas por que a seleção de indicativo implica na exclusão de imperativo. Esta relação de semelhança-e-oposição se denomina agnação (GLEASON, 1965, p. 199).

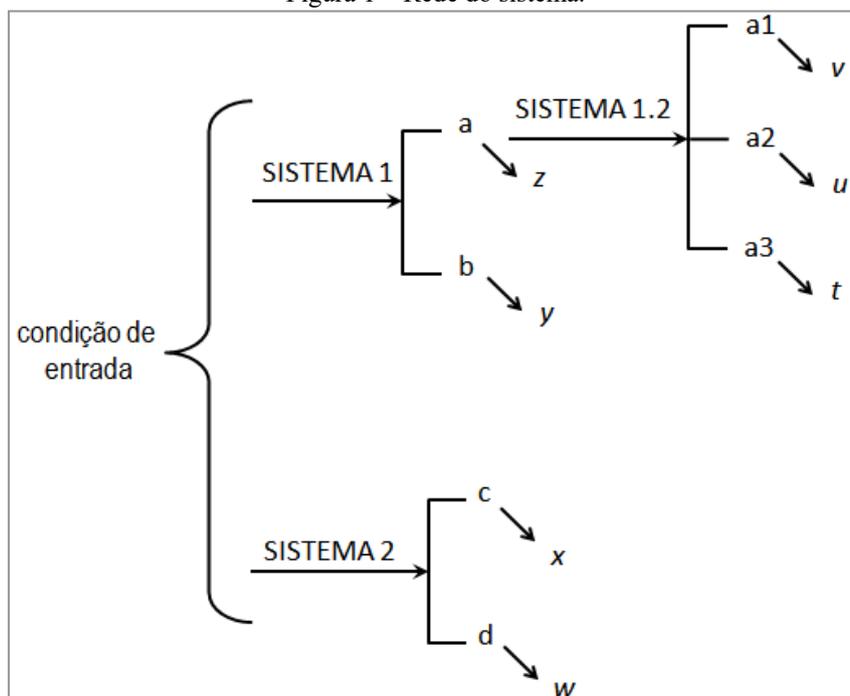
A partir deste pensamento sistêmico, é possível compreender que o significado não está exatamente em nenhuma opção sistêmica, mas na relação entre opções. Por conseguinte, nenhuma função linguística possui valor “em si”; o valor se encontra na relação de semelhança-e-oposição sistêmica. Este pensamento sistêmico, introduzido nos estudos linguísticos por Saussure (2006), foi caracterizado da seguinte forma.

Verifiquemos inicialmente que, mesmo fora da língua, todos os valores parecem estar regidos por esse princípio paradoxal. Eles são sempre constituídos: (1) por uma coisa dessemelhante, suscetível de ser trocada por outra cujo valor resta determinar; (2) por coisas semelhantes que se podem comparar com aquela cujo valor está em causa (p. 134).

O significado relacional das funções linguísticas se denomina valor, ou *valeur*. No mapa topológico, o *valeur*, que é estabelecido pela agnação, é representado pela variação em função da distância – *i.e.*, a distância gramatical, que é medida em *valeur* e suas reactâncias – assim, quanto maior o *valeur* entre duas funções, maior a distância entre elas. Assim, a distância entre as funções Indicativo e Imperativo é pequena, pois o *valeur* que as relaciona é pequeno. Por outro lado, a distância entre as funções Indicativo: Interrogativo: Polar e Imperativo: Jussivo: Ordem é maior, pois o *valeur* que as relaciona é relativamente grande.

O primeiro passo para a produção de modelos está no entendimento de como a LSF concebe a organização da língua como sistema. Uma vez que o objetivo da organização sistêmica é apresentar o funcionamento da língua, cada sistema corresponde a uma classe; assim, uma rede de sistemas corresponderá a uma parte da língua que se quer descrever para, posteriormente, ser modelada. Dessa forma, o sistema constitui-se como a base para a modelagem linguística. A Figura 1 apresenta a disposição do sistema em rede.

Figura 1 – Rede do sistema.



Fonte: os autores.

O sistema possui relações que são determinantes para a sua criação e interpretação. Como mostra a Figura 1, o primeiro elemento do sistema é a condição de entrada, o qual determina a unidade que será descrita. Em seguida apresentam-se os princípios de classificação, que são chamados de sistemas – escritos em caixa alta (MARTIN, 2013). Cada sistema possui opções de classes ('a', 'b', 'c', 'd', 'a1', 'a2', 'a3') e cada classe é apresentada com sua realização ('z', 'y', 'x', 'w', 'v', 'u', 't').

Existe uma relação de oposição no sistema, como por exemplo, entre a classe 'a' e a classe 'b'. Apesar de serem geradas pelo mesmo sistema, são opostas. Essa relação de semelhança-e-oposição, como visto anteriormente, é a agnação (GLEASON, 1965). Nesse sentido, a classe 'a' e a classe 'b' são agnatas. Além disso, quanto mais à direita estiver o sistema, maior o nível de distinção entre os itens, sendo esse nível de distinção chamado de delicadeza (MARTIN, 1992). Após a descrição de uma unidade, é possível visualizar as relações entre as categorias e o significado de cada uma dentro da língua em funcionamento.

Nesse sentido, devido à possibilidade de construção de sistemas de categorias linguísticas (HALLIDAY, 2002), o texto passa a ser compreendido como o produto de escolhas que geram significados nos diversos contextos de situação. Assim, a língua passa a ser um conjunto de textos diversos, organizados, primeiramente, pela dimensão do eixo (URE, 1989; MARTIN, 2013).

A organização da língua conforme a dimensão do eixo abre a possibilidade para a cosseleção de subsistemas para um mesmo fenômeno. Esta cosseleção, por sua vez, permite que os subsistemas se agrupem de forma consistente na categorização de determinado fenômeno. Assim, o estrato gramatical se organiza não só pelos subsistemas, mas, igualmente, pelo agrupamento desses na forma de cosseleções. A cada um desses agrupamentos dá-se o nome de metafunção (ou a função que um grupo de subsistemas desempenha no próprio sistema). Na gramática, as metafunções são três: ideacional, interpessoal e textual.

A metafunção ideacional é responsável pela representação da experiência do mundo; para isso, ela utiliza o sistema da TRANSITIVIDADE, que gera as funções dos cinco Tipos de Processo do português brasileiro² (FIGUEREDO, 2011): Material, Mental, Verbal, Relacional e Existencial. A metafunção interpessoal é responsável pela interação existente entre os

² Ao contrário da gramática do inglês, a gramática do PB não apresenta motivação sistêmica para o Tipo de Processo Comportamental. Em PB, o comportamento emprega ora a gramática material, ora a gramática mental/verbal.

falantes; para isso, ela utiliza o sistema de MODO (FIGUEREDO, 2011), que gera cinco funções de interesse para a presente pesquisa, a saber, Imperativo: Jussivo, Imperativo: Sugestivo, Indicativo: Declarativo, Indicativo: Interrogativo: Polar e Indicativo: Interrogativo: Elemental. A metafunção textual é responsável por semiotizar os significados da metafunção ideacional e interpessoal em texto. Ela gera, por meio do sistema de TEMA, que é o “ponto de partida”³ (MARTIN; ROSE, 2007, p. 107) da mensagem, as funções: Default, Tema Elemental, Ângulo Fonte, Ângulo Ponto de Vista, Perspectiva, Intensivo (FIGUEREDO, 2011).

A posição relativa de uma opção sistêmica reflete o seu valor funcional, e o mapa topológico opera como uma “descrição dinâmica” da língua (LEMKE, 1993). Como consequência, as relações entre as opções sistêmicas da produção textual podem ser quantificadas, o que abre a possibilidade para seu comportamento (*i.e.*, variação) ser modelado.

A partir da distância topológica, é possível projetar como determinado texto se desenvolve no tempo, relativamente às funções que emprega. Em outras palavras, é possível descrever as funções na primeira oração do texto (tempo = 1), em seguida na segunda oração (tempo = 2), na terceira (tempo = 3) e assim por diante, obtendo-se uma “descrição instantânea” da língua.

Assim, um texto que esteja ainda em produção pode ir sendo descrito à medida que vai sendo produzido a cada instante. A descrição instantânea se define como uma descrição gramatical do texto em qualquer momento de sua produção, e não apenas do produto final; é a descrição do texto como processo.

Para tanto, tomam-se a oração como unidade de análise, as opções dos sistemas gramaticais fundamentais da oração TRANSITIVIDADE (trs.), MODO (mod.) e TEMA (tem.) como eixos de coordenadas para a variação de distância e a dinâmica textual para o emprego de cada oração na variação de tempo, em que o tempo tem como quantidade uma oração. Isso leva conseqüentemente à produção de um “vetor-oração” que descreve o movimento do texto entre as orações. A “distância gramatical” e a “sequência temporal” mostram como os sistemas gramaticais “se movem” no espaço, o que se denomina ‘dinâmica de sistemas’.

Como forma de aplicar essa concepção de modelo ao *corpus* da pesquisa, apresenta-se a seguir a metodologia utilizada nesta pesquisa.

³ “Starting point” (MARTIN; ROSE, 2007, p. 107).

3. Metodologia

A metodologia para a perfilação sistêmica se divide em duas partes. Na primeira, apresenta-se o *corpus* investigado; na segunda, os procedimentos que levam à análise, os quais se dividem em (1) estudo do *corpus* e (2) anotação sistêmica.

O presente trabalho utilizou os dados de um *corpus* em português brasileiro (PB) composto por dez⁴ textos caracterizados comumente como *manual*, como aponta Biber (2010) que a descrição das funções gramaticais para um gênero textual não precisa, de fato, ser realizada em textos completos, mas pode ser realizada por amostras. Por conseguinte, os textos foram compilados a partir da tipologia da língua no contexto de cultura (URE, 1989), a qual classifica os textos em oito processos sociosemióticos (explicar, relatar, recriar, compartilhar, fazer, recomendar, capacitar e explorar) e quatro modos de produção (escrito/monólogo, escrito/diálogo, falado/monólogo e falado/diálogo) (HALLIDAY; MATTHIESSEN, 2014). Mais especificamente, para a análise desta pesquisa, foi compilado um *corpus* de textos classificados como capacitar, escrito/monólogo.

Por conseguinte, o critério da tipologia da língua no contexto de cultura (URE, 1989) foi utilizado na compilação do *corpus* de manuais de acordo com as variáveis do contexto de situação, definidas pela LSF (HALLIDAY, 1978; PAGANO, 2015):

Campo:

- Explicação de um procedimento (MATTHIESSEN et al., 2008);
- Exposição das orientações para o leitor (TEICH, 1999).

Sintonia⁵:

- Grau de especialização da linguagem: especializado para não especializado (HALLIDAY, 1978; MATTHIESSEN et al., 2010);
- Relação com o texto: não-familiar e distante (HALLIDAY, 1978; MATTHIESSEN et al., 2010).

⁴ A quantidade de textos empregados no estudo se baseou na capacidade de análise manual, uma vez que ainda não existem estudos indicando um número de textos que seja significativo para a modelagem da dinâmica de sistemas.

⁵ Neste artigo, traduz-se ‘tenor’ como ‘sintonia’ e ‘relations’ como ‘relações’.

Modo:

- Fonte de coleta (HALLIDAY et al., 1964): instituições renomadas do Brasil;
- Modo textual: institucional, pragmático, eletrônico e com objetivo de guiar os leitores para ações específicas (HALLIDAY et al., 1964; ROSE, 2001).

Consoante os procedimentos supracitados, foi possível obter textos pertencentes especificamente ao tipo de texto de interesse deste trabalho, que, de acordo com pesquisas anteriores, será tratado, agora tecnicamente, como ‘manual’ (cf. AZENHA, 1995; LIMA, 2013).

Conforme o estabelecimento desses critérios, a Tabela 1 apresenta os manuais com sua numeração, etiqueta, classificação segundo o processo sociossemiótico e o modo de produção (HALLIDAY; MATTHIESSEN, 2014), fonte de coleta e número de *tokens* referente à macro-onda de informação⁶ (MARTIN; ROSE, 2007) analisada de cada manual.

Tabela 1 – Perfil do *corpus* de manuais capacitar, escrito/monólogo.

Manual	Etiqueta	Classificação	Instituição	Número de <i>tokens</i>
Manual técnico de Gemas	GEM_01	Capacitar/escrito/monólogo	IBGM- DNPM	1915
Manual de Procedimentos de Enfermagem	ENF_02	Capacitar/escrito/monólogo	Governo do Distrito Federal	793
Manual do Diretor técnico	DIR_03	Capacitar/escrito/monólogo	CREMERJ	1963
Manual técnico de hemovigilância	HEM_04	Capacitar/escrito/monólogo	ANVISA	1105
Manual de Biossegurança em odontologia – UFPB	BIO_05	Capacitar/escrito/monólogo	UFPB	835
Manual de redação da presidência da república	RED_06	Capacitar/escrito/monólogo	Presidência da República	1301
Manual de noções básicas de primeiros socorros	SOS_07	Capacitar/escrito/monólogo	UNICAMP	493
Manual técnico - utilização de produtos fitofarmacêuticos	FIT_08	Capacitar/escrito/monólogo	Direção Geral da Agricultura e Desenvolvimento Ambiental	1302
Manual de normas técnicas - combate à Dengue	DEN_09	Capacitar/escrito/monólogo	FUNASA	789

⁶ O termo macro-onda de informação apresenta o fluxo de informação no texto em nível macro. A palavra “onda” é utilizada dentro da LSF como termo técnico para se referir à estrutura do fluxo de informação do texto (MARTIN; ROSE, 2007, p. 188).

Manual técnico - criação de pacu	PAC_10	Capacitar/escrito/monólogo	Itaipu Binacional e Unioeste/Campus de Toledo	498
			TOTAL	10994

fonte: os autores.

Após a compilação do *corpus*, foram adotados dois procedimentos de análise: (1) estudo do *corpus*; e (2) anotação sistêmica.

(1) Estudo do *corpus*: primeiramente, para a construção do perfil sistêmico do manual, analisaram-se os textos pelo viés do fluxo de informação. Martin e Rose (2007) apontam que o fluxo de informação é a forma como a metafunção textual organiza seus significados para construir o texto. Nesse sentido, o Quadro 1 apresenta o *corpus* em relação ao fluxo de informação presente em cada texto.

Quadro 1 – Ondas de informação *corpus* de manuais capacitar/escrito/monólogo.

MANUAL	ELEMENTOS PARATEXTUAIS	CONTEÚDO TEXTUAL				REFERÊNCIAS
		Apresentação	Introdução	Desenvolvimento		
				Explicação	Como fazer	
GEM_01	x	X		X	X	
ENF_02	x	X		X	X	x
DIR_03	x	X		X	X	x
HEM_04	x	X	X	X	X	x
BIO_05	x	X		X	X	x
RED_06	x	X		X	X	x
SOS_07				X	X	x
FIT_08	x	X		X	X	x
DEN_09	x	X		X	X	x
PAC_10	x	X	X	X	X	

Fonte: os autores.

O Quadro 1 mostra que o manual possui três ondas de informação: (1) elementos paratextuais, que se referem aos elementos que aparecem antes do texto, como capa, carta ao leitor e índice; (2) conteúdo textual, que é composto por três ondas: apresentação, introdução e desenvolvimento; e (3) referências bibliográficas.

A marcação ‘x’ indica as ondas de informação que estão presentes nos manuais. Ademais, a marcação ‘X’, em negrito e caixa alta, aponta a presença das ondas de informação ‘Explicação’ e ‘Como fazer’ em todos os manuais no ‘Desenvolvimento’. A seção ‘Explicação’ expõe uma explicação sobre o assunto tratado; e a seção ‘Como fazer’ explica os procedimentos de como realizar a atividade determinada no manual.

Como aponta o Quadro 1, a onda de informação ‘Desenvolvimento’ é uma hiperonda de informação (MARTIN; ROSE, 2007), composta por várias ondas menores que possuem a mesma forma padrão: explicação e como fazer. Essas ondas menores de informação são a parte do manual nas quais a língua é usada para facilitar o comportamento do falante em determinada atividade. Sendo assim, essa seção é composta por textos capacitar, os quais compõem o *corpus* de análise deste artigo.

(2) Anotação sistêmica: para essa etapa, o *corpus* foi etiquetado e armazenado segundo os critérios de etiquetamento e armazenamento do Catálogo da Língua Brasileira (CALIBRA) e, posteriormente, incorporado ao *software* UAM CorpusTool (O’DONNELL, 2008) para a análise sistêmico-funcional semiautomática⁷.

Após essas duas etapas, foi realizada a anotação das frequências relativas à TRASITIVIDADE, MODO e TEMA, por serem os três sistemas gramaticais da oração relativamente às metafunções ideacional, interpessoal e textual.

Após a análise do *corpus* segundo as categorias da LSF, as orações descritas são colocadas no espaço gramatical segundo o mapa topológico de cada metafunção, estabelecendo-se assim suas distâncias relativas (FIGUEREDO, 2014). Como ilustração, a Figura 2 apresenta o mapa topológico do TEMA em português brasileiro e o Quadro 2 mostra as funções relativas ao mapa.

⁷ O *software* UAM CorpusTool (O’DONNELL, 2008) permite incorporar os sistemas que serão utilizados na análise do *corpus*. Ao final da análise, *software* apresenta as frequências relativas a cada ocorrência categorizada pelo pesquisador.

Quadro 3: Perfil sistêmico parcial do Manual de Biossegurança em Odontologia.

ORAÇÃO	REALIZAÇÃO	TRANS (x)	MODO (y)	TEMA (z)
BIO_05_01	Assepsia: é o conjunto de medidas que utilizamos para impedir a penetração de microrganismos num ambiente que logicamente não os tem,	39 ⁸	20	10
BIO_05_02	logo um ambiente asséptico é aquele que está livre de infecção.	39	20	41
BIO_05_03	Antissepsia: é o conjunto de medidas propostas para inibir o crescimento de microrganismos ou removê-los de um determinado ambiente, podendo ou não destruí-los	39	20	10
BIO_05_04	e para tal fim utilizamos antissépticos ou desinfetantes.	14	20	41
BIO_05_05	É a destruição de micro-organismos existentes nas camadas superficiais ou profundas da pele, mediante a aplicação de um agente germicida de baixa causticidade, hipoalergênico e passível de ser aplicado em tecido vivo.	39	20	10
BIO_05_06	Degermação: Significa a diminuição do número de microrganismos patogênicos ou não, após a escovação da pele com água e sabão.	39	20	10
BIO_05_07	É a remoção de detritos e impurezas depositados sobre a pele.	39	20	10
BIO_05_08	Sabões e detergentes sintéticos, graças a sua propriedade de umidificação, penetração, emulsificação e dispersão, removem mecanicamente a maior parte da flora microbiana existente nas camadas superficiais da pele, também chamada flora transitória,	14	20	10
BIO_05_09	mas não conseguem remover aquela que coloniza as camadas mais profundas ou flora residente.	14	20	10
BIO_05_10	Desinfecção: é o processo pelo qual se destroem particularmente os germes patogênicos e/ou se inativa sua toxina ou se inibe o seu desenvolvimento.	39	20	10
BIO_05_11	Os esporos não são necessariamente destruídos.	16	20	10
BIO_05_12	Esterilização: é processo de destruição de todas as formas de vida microbiana (bactérias nas formas vegetativas e esporuladas, fungos e vírus) mediante a aplicação de agentes físicos e ou químicos,	39	20	10
BIO_05_13	Toda esterilização deve ser precedida de lavagem e enxaguadura do artigo para remoção de detritos.	14	20	10

Fonte: os autores.

⁸ Cada código descreve uma função gramatical no texto, a seguir: 39 (Processo: Relacional: Identificativo & Intensivo); 14 (Processo: Material: Transitivo & Transformativo); 16 (Processo: Material: Transitivo & Criativo); 20 (Modo: Declarativo); 10 (Tema: Default) e 41 (Tema: Perspectiva).

O Quadro 3 aponta o número de cada oração composto pelo código da etiqueta do texto. Por exemplo: BIO_05, que significa que este é o quinto texto compilado para o *corpus*, é o manual de biossegurança. Juntamente com essa etiqueta está o número referente a cada oração. Por exemplo: BIO_05_01, significa que esta é a primeira oração do texto BIO_05. Em seguida, a segunda coluna apresenta a realização linguística da oração.

A terceira coluna apresenta as funções do sistema de TRANSITIVIDADE do texto. Por exemplo, para a oração BIO_05_01, “Assepsia: é o conjunto de medidas...”, tem-se um Processo: Relacional Identificativo & Intensivo, que segundo o mapa topológico do sistema de TRANSITIVIDADE em português brasileiro é codificado com o número 39; ‘3’ para Relacional e ‘9’ para Identificativo & Intensivo. Já para a oração BIO_05_04, “e para tal fim utilizamos antissépticos”, tem-se um Processo: Material Transitivo & Transformativo, que, segundo o mapa topológico da TRANSITIVIDADE, recebe o número 14; ‘1’ para Material e ‘4’ para Transitivo & Transformativo.

A quarta coluna apresenta as funções do MODO. Por exemplo, para a oração BIO_05_01, tem-se o Modo Indicativo: Declarativo, classificado como 20; ‘2’ para Indicativo e ‘0’ para Declarativo.

A quinta coluna apresenta as funções do TEMA. Por exemplo, a oração BIO_05_01 possui Tema *Default* e é classificada como 10. No caso da oração BIO_05_04, a função do Tema é Intensivo: Perspectiva e por isso classificada com o número 41; ‘4’ para intensivo e ‘1’ para perspectiva.

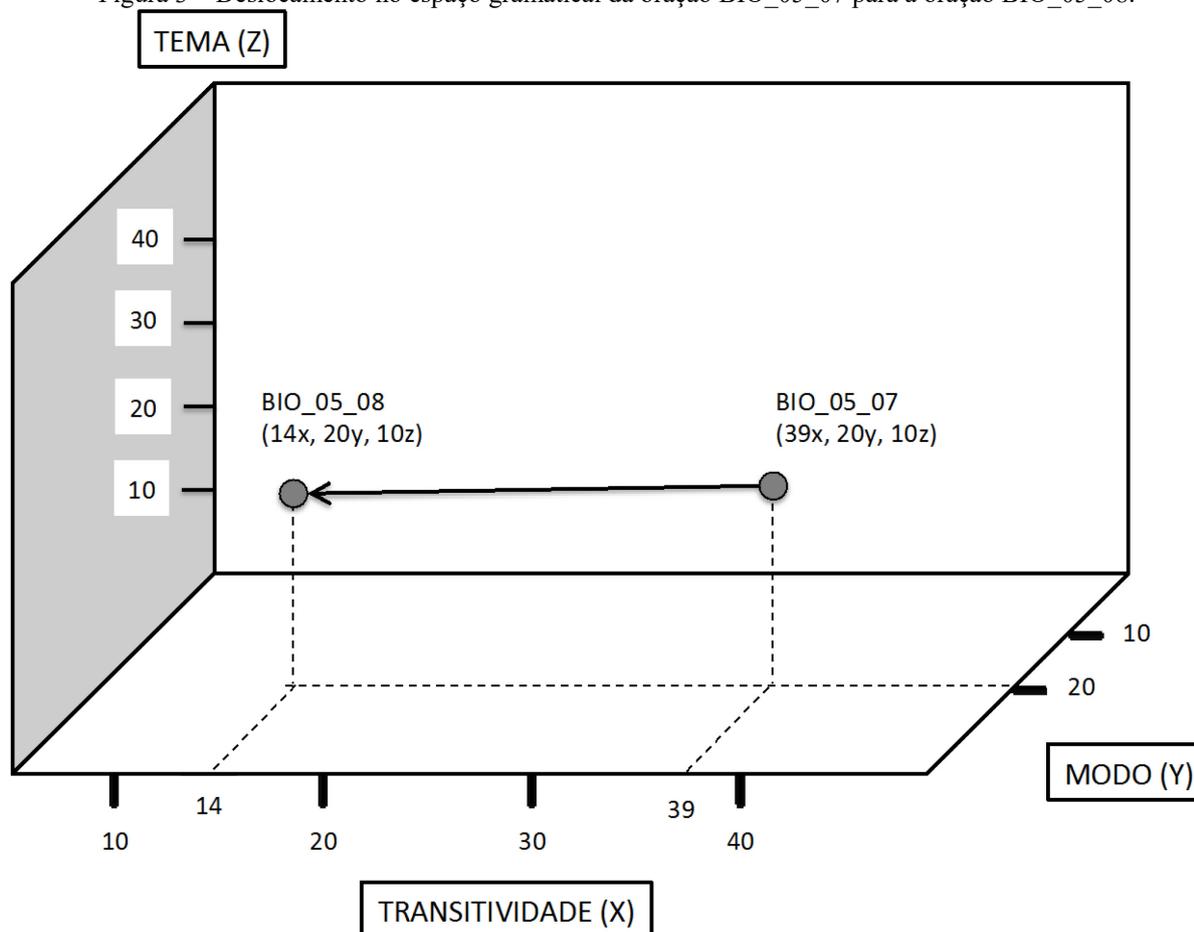
Esse tipo de quadro também foi abordado nos trabalhos de Lima (2013) e Braga (2013). Ele aponta todas as orações do texto com a disposição da TRANSITIVIDADE, do MODO e do TEMA, permitindo assim a visualização da mudança entre orações e, conseqüentemente, as mudanças de significado no manual.

Como complementação, acrescentam-se à perfilação gramatical a identificação e descrição dinâmica da mudança que acontece entre orações e o “movimento” da gramática pelo texto. Cada ponto do espaço representa uma oração composta pelas categorias de TRANSITIVIDADE (eixo x), MODO (eixo y) e TEMA (eixo z). As semirretas que ligam os pontos representam vetores (i.e., magnitude e direção) que apresentam o movimento de uma oração para outra.

Esse movimento pode ser observado, por exemplo, entre as orações BIO_05_07 “É a remoção de detritos e impurezas depositados sobre a pele” e BIO_05_08 “Sabões e detergentes

sintéticos, graças a sua propriedade de umidificação, penetração, emulsificação e dispersão, removem mecanicamente a maior parte da flora microbiana existente nas camadas superficiais da pele, também chamada flora transitória”. O vetor indica uma mudança no espaço gramatical no eixo da TRANSITIVIDADE, de uma oração relacional “[Degermação] é” para uma oração material “sabões... removem”. O Modo para ambas as orações, bem como o Tema se mantêm. Dessa forma, o movimento é zero. Com isso tem-se: BIO_05_07 (39x, 20y, 10z) → BIO_05_08 (14x, 20y, 10z). No espaço, tem-se:

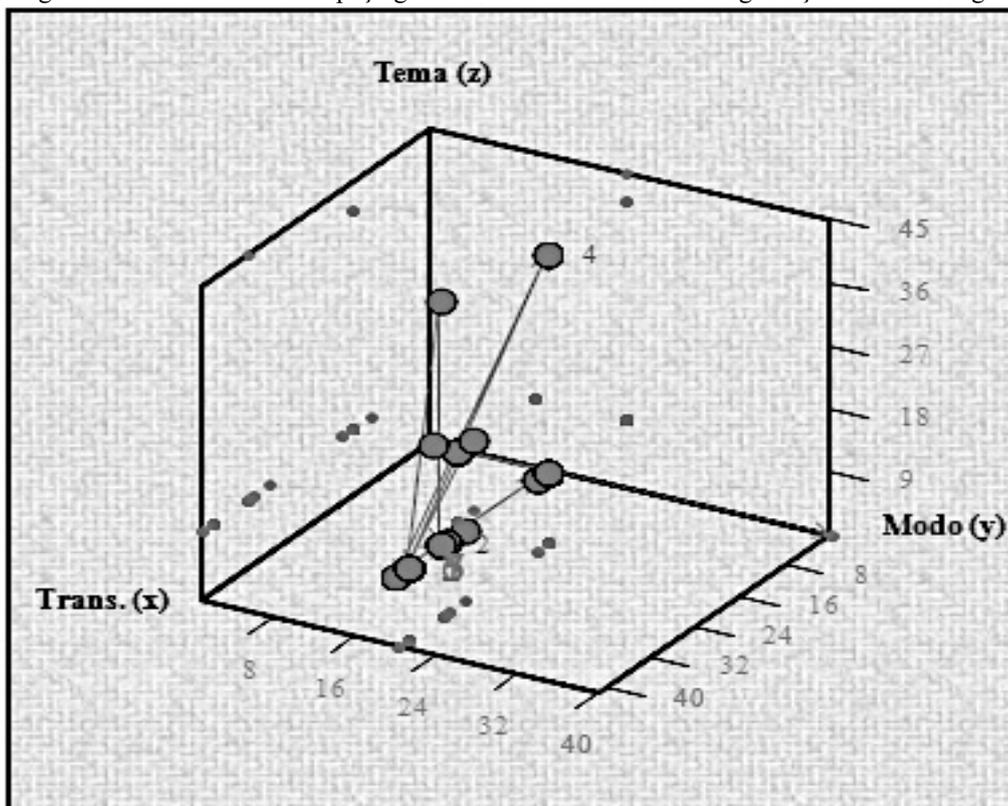
Figura 3 – Deslocamento no espaço gramatical da oração BIO_05_07 para a oração BIO_05_08.



Fonte: os autores.

A Figura 4 apresenta o deslocamento no espaço gramatical do Manual de Biossegurança em Odontologia.

Figura 4: Deslocamento do espaço gramatical do Manual de Biossegurança em Odontologia.



Fonte: os autores.

Esse procedimento foi realizado em todos os textos do *corpus* através de uma macro desenvolvida no Microsoft Office Excel para transferir a perfilação gramatical transcrita em códigos para um gráfico em 3D⁹. Esse tratamento computacional da perfilação permite a visualização e análise das mudanças que a língua faz, do ponto de vista gramatical, para criar o significado do manual (FIGUEREDO, 2014; OLIVEIRA, 2015). Por esse motivo, esse passo metodológico torna-se uma relevante contribuição para a Linguística Computacional e, como consequência, a Linguística de *Corpus*.

4. Resultados

No *corpus* de manuais, após a análise de cada manual no *software* UAM CorpusTool, os resultados de cada categoria foram obtidos em frequências, pela necessidade de se observar a contribuição das diferentes funções e, assim, a probabilidade de escolhas na produção de significado em cada texto. A Tabela 2 exhibe os resultados.

⁹ Obtida em: <http://www.doka.ch/Excel3Dscatterplot.htm>. Acesso em: dezembro de 2016.

Tabela 2: Resultados da análise do *corpus* de manuais em relação à TRANSITIVIDADE, MODO e TEMA retirados do software *UAM CorpusTool*.

TRANSITIVIDADE	
Material	62,17%
Relacional	29,84%
Verbal	5,68%
Existencial	2,31%
Mental	0,00%
TOTAL	100,00%
MODO	
Declarativo	72,47%
Jussivo	27,53%
Sugestivo	0,00%
Polar	0,00%
Elemental	0,00%
TOTAL	100,00%
TEMA	
<i>Default</i>	86,50%
Perspectiva	8,70%
Intensivo	4,26%
Ângulo ponto de vista	0,36%
Ângulo fonte	0,18%
Elemental	0,00%
TOTAL	100,00%

Fonte: os autores.

A Tabela 2 mostra as frequências relativas às delicadezas da TRANSITIVIDADE, do MODO e do TEMA. Na TRANSITIVIDADE, o Processo Material possui 62,17% das ocorrências, o Processo Relacional possui 29,84%, o Verbal 5,68%, o Existencial 2,31% e o Mental 0,00%. No MODO, o Declarativo obteve 72,47%, o Jussivo 27,53% e o Sugestivo, Polar e Elemental 0,00%. No TEMA, o Tema Default obteve 86,50%, o Perspectiva 8,70%, o Intensivo 4,26%, o Ângulo Ponto de Vista 0,36%, o Ângulo Fonte 0,18% e o Elemental 0,00%.

Os dados mostram que o *corpus* de manual é composto principalmente por Processo: Material e Processo: Relacional, Modo: Declarativo e Modo: Jussivo, e Tema: Default, Tema: Perspectiva e Tema: Intensivo. Por conseguinte, esses resultados apresentam o que é o manual: um texto que procura explicar procedimentos, identificar seu objeto de abordagem e apresentar algumas ordens. Halliday e Matthiessen (1999, p. 98) apontam que “nos gêneros textuais que

apresentam procedimentos, as orações materiais são normalmente imperativas e as relacionais indicativas”.¹⁰¹¹ A afirmação dos autores é corroborada pelo que foi encontrado na análise.

Mediante os resultados obtidos através do *corpus*, foi realizada a análise gramatical e semântica das ondas de informação de cada manual. Calcularam-se as frequências relativas a cada categoria supracitada em relação às ondas de informação e, por fim, o modelo, contendo as ocorrências mais frequentes deste gênero textual foi desenvolvido.

O modelo não é, na verdade, nenhum texto específico, criado por usuários do português brasileiro. É, sim, uma generalização das probabilidades maiores de escolhas dos sistemas gramaticais para esse tipo de texto. Assim, a análise que se segue não se refere especificamente a qualquer texto em particular, mas prevê comportamento da gramática oracional do português brasileiro para quaisquer textos caracterizados como manual.

O Quadro 4 a seguir expõe as duas primeiras ondas de informação do modelo. Devido ao fato de o modelo ser uma generalização do *corpus*, ou seja, “formado por partes gramaticais de todos os textos”, ele foi nomeado como Frankenstein:

Quadro 4 — As duas primeiras ondas de informação do Frankenstein.

ONDAS	ORAÇÕES	TRANS. (x)	MODO (y)	TEMA (z)
1	FRANK_01	39 ¹²	20	10
	FRANK_02	33	20	10
	FRANK_03	39	20	10
	FRANK_04	32	20	10
	FRANK_05	39	20	10
	FRANK_06	39	20	10
	FRANK_07	33	20	10
	FRANK_08	14	20	41
	FRANK_09	14	20	10
	FRANK_10	32	20	10
	FRANK_11	32	20	10
	FRANK_12	14	20	10
	FRANK_13	33	20	10

¹⁰ ...in procedural registers, material clauses are typically imperative, relational ones indicative.

¹¹ Todas as traduções de citações apresentadas neste artigo, salvo quando for indicado, foram feitas pelo autor.

¹² Cada código descreve uma função gramatical do Frankenstein, a seguir: 39 (Processo: Relacional: Identificativo & Intensivo); 33 (Processo: Relacional: Atributivo & Intensivo); 32 (Processo: Relacional: Atributivo & Intensivo); 14 (Processo: Material: Transitivo & Transformativo); 29 (Processo: Existencial: Permanência); 20 (Modo: Declarativo); 11 (Modo: Imperativo: Jussivo); 10 (Tema: Default); 41 (Tema: Perspectiva); 42 (Tema: Intensivo).

2	FRANK_14	14	20	42
	FRANK_15	29	20	10
	FRANK_16	14	20	10
	FRANK_17	14	20	10
	FRANK_18	14	20	10
	FRANK_19	14	11	10
	FRANK_20	14	11	10
	FRANK_21	14	11	10
	FRANK_22	14	11	10

Fonte: os autores.

O Quadro 4 apresenta o funcionamento do modelo em relação às duas primeiras ondas de informação. A primeira onda tem como objetivo identificar e atribuir características ao objeto de trabalho do manual. Para isso ela utiliza Processo: Relacional (39 – Identificativo & Intensivo, 33 – Atributivo & Intensivo, 32 – Atributivo & Possessivo) (HALLIDAY; MATTHIESSEN, 1999), o Modo: Declarativo (20) e o Tema: Default (10).

Nessa mesma onda também existem orações com o Processo: Material (14 – Transitivo & Transformativo) que ajudam na explicação do objeto apresentado (HALLIDAY; MATTHIESSEN, 1999) e quando há a troca do tipo de Processo (de uma oração 33 passando para o 14 – 33→14) existe uma mudança no Tema que está presente nessa troca.

A segunda onda é responsável pela explicação do objeto tratado. Para isso, utiliza-se basicamente de Processo: Material (14), do Modo: Declarativo (20), e de alguns comandos para o leitor, utilizando o Modo: Jussivo (11) e Tema: Default (10). No entanto, quando ocorre a troca de onda de informação (1→2) o Tema muda de Default para Intensivo (10/42).

Além do Processo: Material, existe no início da onda um Processo: Existencial (29 - Permanência), que faz a apresentação de um novo participante na explicação, o Existente (HALLIDAY; MATTHIESSEN, 1999). O Quadro 5 mostra as últimas três ondas de informação do Frankenstein:

Quadro 5: As três últimas ondas de informação do Frankenstein.

ONDAS	ORAÇÕES	TRANS. (x)	MODO (y)	TEMA (z)
3	FRANK_23	33 ¹³	20	41
	FRANK_24	14	20	10
	FRANK_25	16	20	10
	FRANK_26	14	20	10
	FRANK_27	14	20	10
	FRANK_28	11	20	10
	FRANK_29	14	20	10
	FRANK_30	14	20	10
	FRANK_31	41	20	41
4	FRANK_32	14	20	10
	FRANK_33	16	11	10
	FRANK_34	14	11	10
	FRANK_35	14	11	10
	FRANK_36	14	11	10
	FRANK_37	14	11	10
	FRANK_38	14	11	10
	FRANK_39	14	11	41
	FRANK_40	14	11	10
	FRANK_41	14	11	10
	FRANK_42	14	11	10
5	FRANK_43	14	20	42
	FRANK_44	32	20	10
	FRANK_45	41	20	10
	FRANK_46	14	20	10

Fonte: os autores.

A terceira onda, como a segunda, é também iniciada com uma mudança de Tema. O que era Default se torna um Tema: Perspectiva (10 → 41) que apresenta no início da oração uma circunstância de intensificação. Na mesma oração, há um Processo: Relacional (31 – Atributivo & Circunstancial) com o Modo: Declarativo (20) que caracteriza o objeto tratado no manual (HALLIDAY; MATTHIESSEN, 1999).

Em seguida, há uma sequência de orações que apresentam o Processo: Material (14 – Transitivo & Transformativo, 16 – Transitivo & Criativo, 11 – Intransitivo & Transformativo),

¹³ Cada código descreve uma função gramatical do Frankenstein, a seguir: 32 (Processo: Relacional: Atributivo & Intensivo); 33 (Processo: Relacional: Atributivo & Intensivo); 14 (Processo: Material: Transitivo & Transformativo); 16 (Processo: Material: Transitivo & Criativo); 41 (Processo: Verbal: Semiose & Não-recepção); 20 (Modo: Declarativo); 29 (Processo: Existencial: Permanência); 20 (Modo: Declarativo); 11 (Modo: Imperativo: Jussivo); 10 (Tema: Default); 41 (Tema: Perspectiva); 42 (Tema: Intensivo).

Modo: Declarativo (20) e Tema: Default (10), pois esta onda trata de uma explicação que termina com um Processo: Verbal (41 – Semiose & Não Recepção) e um Tema: Perspectiva (41). Este Processo: Verbal vem para nomear um novo participante, o que ganha suporte na afirmação de Martin e Rose (2007), “Os procedimentos de textos escritos dão ainda um passo além na mediação entre a expertise do autor e o leitor, direcionando cada etapa em relação a objetos e lugares cuidadosamente nomeados”.¹⁴

A quarta onda é responsável pela segunda parte de comandos do manual. Ela começa com explicação (14/20/10) e continua com um fluxo de dez orações de comando realizadas pelo Modo: Jussivo (11) (HALLIDAY; MATTHIESSEN, 1999). Existe uma mudança do tipo de Processo: Material (14 → 16), que acontece devido ao significado do Processo, quando o primeiro transforma um Participante no texto, e o segundo cria ou elimina um outro Participante (HALLIDAY; MATTHIESSEN, 1999).

Além do Processo, o Tema também sofre uma mudança na oração 39, o que era Default (10) passa a ser Tema: Perspectiva (41). Nesse sentido, como ele apresenta uma circunstância de intensificação que é colocada em posição temática como o primeiro elemento da oração. Ela pode ser causal, temporal ou condicional.

E por fim, a última onda trata de finalizar o texto (MARTIN; ROSE, 2007). Para isso ela explica, identifica e nomeia seu objeto do manual usando Processo: Material (14), Processo: Relacional (32) e Processo: Verbal (41) (HALLIDAY; MATTHIESSEN, 1999). Ademais, como ocorreu no início da segunda e terceira ondas, esta também é marcada por um Tema diferente do Default (10). Aqui o Tema: Intensivo (42) aparece na troca de onda de informação. Ele “implica em introdução de, ou contraste com, um novo assunto e subsequente mudança de fase” (FIGUEREDO, 2011, p. 154).

As ondas de informação do modelo, juntamente com a dinâmica existente entre elas e dentro delas, apresentam as probabilidades mais comuns de ocorrência no manual. Sendo assim, este Frankenstein não é o manual e sim um modelo criado para compreender como são realizadas, pelo falante, as escolhas gramaticais (HALLIDAY et al., 1964), a fim de construir o significado do manual. Para exemplificar o modelo, o Quadro 6 mostra a sua primeira onda de informação com exemplos do *corpus*. Os exemplos seguem a descrição de cada oração exibida no Quadro 4:

¹⁴ “Written procedures go a step further than this to mediate the author’s expertise, directing the learner what to do at each step, in relation to explicitly named objects and locations.” (MARTIN; ROSE, 2007, p. 180).

Quadro 6 — O Frankenstein com exemplos do *corpus*.

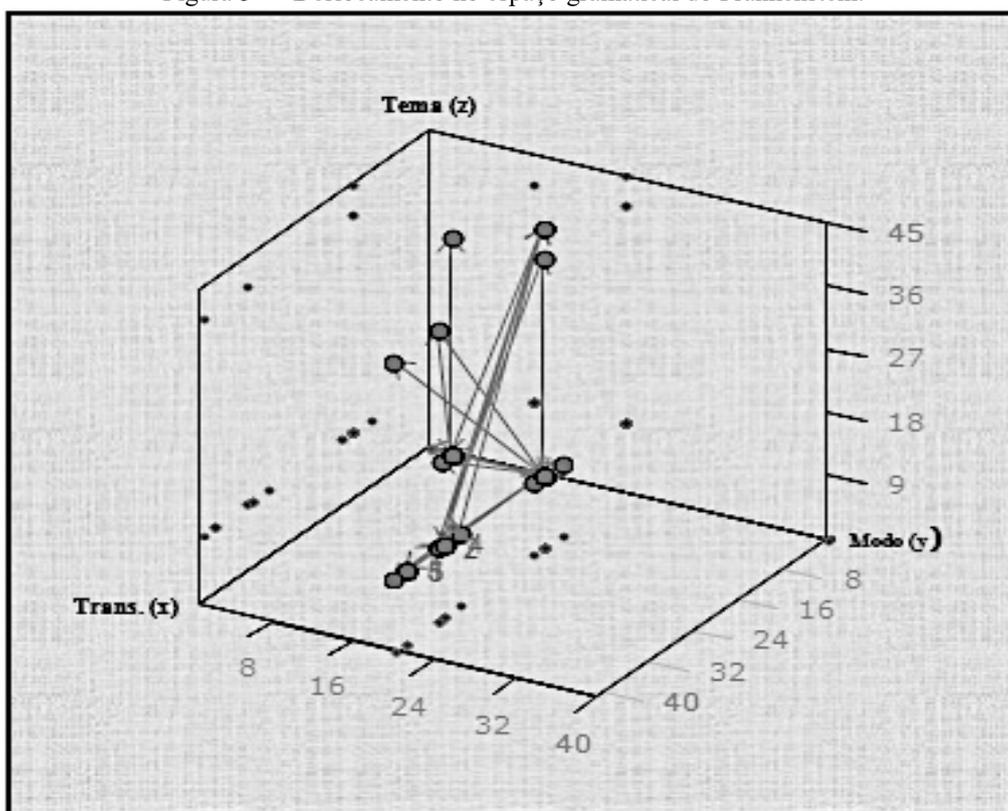
ORAÇÃO	EXEMPLOS
FRANK_01	Assepsia: é o conjunto de medidas que utilizamos para impedir a penetração de microrganismos num ambiente que logicamente não os tem
FRANK_02	Essas complicações na maioria das vezes são inevitáveis.
FRANK_03	Os utensílios utilizados (baldes, bacias, sacos plásticos, entre outros) devem ser limpos sempre que seja observado resíduo (Figura 18).
FRANK_04	Além do itinerário que lhe compete, com os quarteirões numerados, o pessoal da operação deve dispor das tarefas de cada dia e de croquis com o desenho e a disposição de todos os quarteirões (quadras) de cada área.
FRANK_05	Trata-se de uma exigência constitucional (Constituição, art.39,III e 83),
FRANK_06	Esta situação nada mais é do que uma obrigação contratual.
FRANK_07	neste caso a participação dos envolvidos no processo é fundamental.
FRANK_08	Quando necessário, deve-se discriminar a massa (peso) individual das gemas que compõem a joia.
FRANK_09	A marcação se fará com lápis-cera azul ou preto no cateto esquerdo de cada ângulo do quarteirão.
FRANK_10	Ela deve conter o plano de governo, exposição da situação do País e solicitação de providências que julgar necessárias (Constituição, art. 84, XI)
FRANK_11	Os números e os sinais devem ter cinco centímetros de altura.
FRANK_12	Os quarteirões receberão numeração crescente, do número um ao infinito.
FRANK_13	A lavagem de mãos é obrigatória para todos os componentes da equipe de saúde bucal;

Fonte: os autores.

O Quadro 6 mostra o Frankenstein que foi criado com partes distintas dos manuais do *corpus*. Apesar de ele não apresentar os mesmos participantes nas orações e de não haver conectividade entre um termo e outro do ponto de vista lexical, este modelo está gramaticalmente correto no sentido de que ele representa ideacionalmente, encena interpessoalmente e constrói textualmente a primeira onda de informação do *corpus* que trata principalmente da identificação do objeto abordado pelo manual. Ademais, o modelo faz isso a partir das probabilidades maiores de escolha de funções dos sistemas envolvidos para cada oração.

O significado desse tipo de texto, o manual, além de ser compreendido pelas suas funções gramaticais, também é compreendido pela disposição das mudanças que ocorrem de uma oração para outra. A Figura 5 exhibe o deslocamento no espaço gramatical do Frankenstein desta pesquisa:

Figura 5 — Deslocamento no espaço gramatical do Frankenstein.



Fonte: os autores.

Como na Figura 4, a Figura 5 descreve o movimento gramatical realizado pelo Frankenstein. Primeiramente ocorrem as escolhas gramaticais em 'x', 'y' e 'z'. Através dessas escolhas são formadas as orações (representadas pelos pontos). Na sequência, observa-se o deslocamento (representado pelas setas) de uma oração para outra. Este conjunto de escolhas e movimentos forma o modelo de descrição dinâmica, ou a dinâmica textual (segundo a dinâmica discursiva apresentada em Lemke no artigo de 1993).

No deslocamento, as ligações entre os pontos podem ser convertidas em vetores, por descreverem movimento. Com isso, apresentam as mudanças gramaticais em relação ao tempo (FIGUEREDO, 2014). Sendo assim, os vetores mostram a quantidade de mudança gramatical de uma oração para outra. O Quadro 7 exhibe as equações vetoriais da primeira onda de informação do Frankenstein:

Quadro 7 — Equações vetoriais do desenvolvimento da primeira onda de informação do Frankenstein.

ORAÇÃO	TRANS./MODO/TEMA	EQUAÇÕES VETORIAIS	QUANTIDADE DE MUDANÇA
FRANK_01	39/20/10	$r = \langle 39, 20, 10 \rangle + t \langle -6, 0, 0 \rangle$	$r = \langle 39 - 6t, 20, 10 \rangle$
FRANK_02	33/20/10	$r = \langle 33, 20, 10 \rangle + t \langle 6, 0, 0 \rangle$	$r = \langle 33 + 6t, 20, 10 \rangle$
FRANK_03	39/20/10	$r = \langle 39, 20, 10 \rangle + t \langle -7, 0, 0 \rangle$	$r = \langle 39 + 7t, 20, 10 \rangle$
FRANK_04	32/20/10	$r = \langle 32, 20, 10 \rangle + t \langle 7, 0, 0 \rangle$	$r = \langle 32 + 7t, 20, 10 \rangle$
FRANK_05	39/20/10	$r = \langle 39, 20, 10 \rangle + t \langle 0, 0, 0 \rangle$	$r = \langle 39, 20, 10 \rangle$
FRANK_06	39/20/10	$r = \langle 39, 20, 10 \rangle + t \langle -6, 0, 0 \rangle$	$r = \langle 39 - 6t, 20, 19 \rangle$
FRANK_07	33/20/10	$r = \langle 33, 20, 10 \rangle + t \langle -19, 0, 31 \rangle$	$r = \langle 33 - 19t, 20, 10 + 31t \rangle$
FRANK_08	14/20/41	$r = \langle 14, 20, 41 \rangle + t \langle 0, 0, -31 \rangle$	$r = \langle 14, 20, 41 - 31t \rangle$
FRANK_09	14/20/10	$r = \langle 14, 20, 10 \rangle + t \langle 18, 0, 0 \rangle$	$r = \langle 14 + 18t, 20, 10 \rangle$
FRANK_10	32/20/10	$r = \langle 32, 20, 10 \rangle + t \langle 0, 0, 0 \rangle$	$r = \langle 32, 20, 10 \rangle$
FRANK_11	32/20/10	$r = \langle 32, 20, 10 \rangle + t \langle -18, 0, 0 \rangle$	$r = \langle 32 - 18t, 20, 10 \rangle$
FRANK_12	14/20/10	$r = \langle 14, 20, 10 \rangle + t \langle 19, 0, 0 \rangle$	$r = \langle 14 + 19t, 20, 10 \rangle$
FRANK_13	33/20/10		

Fonte: os autores.

O mapeamento das funções gramaticais e o estabelecimento do descolamento gramatical apresentado na metodologia por meio da perfilação gramatical constituem, assim, o modelo topológico do manual.

Por fim, cabe ressaltar que as frequências relativas apresentadas para o modelo do manual podem coincidir – e, com efeito, coincidem – com frequências relativas de outros gêneros textuais. A alta incidência de Processos Materiais encontrada no manual é compartilhada, por exemplo, com gêneros que envolvem a representação de mudança material no mundo, tais como a descrição de processos físicos, ou de atividades de trabalho, entre outras. Já a frequência do Modo encontrada no manual pode ser comparada à frequência de outros gêneros tais como os manuais de guia e de como fazer, receitas, procedimentos, entre outros. Esta constatação revela que o manual é, por um lado, um gênero separado de outros; porém, por outro lado, existe uma contiguidade entre este e outros gêneros, o que se verifica pela

frequência comparável entre as funções gramaticais. Assim, a construção de modelos de gêneros e sua relação com outros modelos de gêneros semelhantes abre espaço para a constituição de famílias de gêneros e de processos sociossemióticos: explicar, relatar, recriar, compartilhar, fazer, recomendar, capacitar e explorar (URE, 1989; HALLIDAY; MATTHIESSEN, 2014).

5. Considerações finais

Tendo como foco principal apresentar uma forma de contribuição de aplicações da Linguística de *Corpus* e LSF à análise e descrição da produção de textos, este artigo apresentou a modelagem do gênero textual ‘manual’ de forma que esta possa subsidiar trabalhos futuros em Linguística Aplicada, mais especificamente, trabalhos envolvendo o Processamento de Linguagem Natural.

Através da metodologia de perfilação gramatical, a qual possibilita compilar e analisar corpora gramaticais (FIGUEREDO, 2014), foi possível construir um modelo de produção textual pautado pelas escolhas em sistemas gramaticais. A justificativa para uma metodologia nova para os estudos de produção textual a partir da gramática se deve ao fato de que, em geral, as descrições dos gêneros textuais carecem de descrições relativas à dinâmica textual. Assim, este artigo oferece uma contribuição por conseguir localizar as escolhas nos sistemas e explicar, por meio das probabilidades mais frequentes, o comportamento sistêmico, neste caso específico, do manual.

O modelo, que não representa qualquer texto especificamente, configura-se como uma das formas de generalização das probabilidades maiores de escolhas dos sistemas gramaticais para esse gênero textual. Assim, a análise que se segue tem o potencial para ser aplicada, de uma forma geral, a textos caracterizados como manual por ser capaz de prever, em grande medida, qual é o comportamento futuro do texto. Assim, em qualquer momento do texto, existe o potencial para se saber aonde a gramática “se movimenta”.

Como ilustração, é possível saber que a partir da oração-média 32, ou quarta onda de informação, a explicação continua com um fluxo de dez orações de comandos, com mudança de Modo Declarativo para Imperativo: Jussivo, com uma quantidade de mudança gramatical alta do ponto de vista interpessoal, mas baixa do ponto de vista ideacional e textual, ou 14/20/10 → 16/11/10, ou $r^{\rightarrow} = \langle 2+34t, 9, 0 \rangle$.

Desta maneira, o artigo também contribui, em outra frente, para os estudos em LSF, pois mostra o potencial robusto dessa teoria para lidar com novas formas de representação da produção de textos. Assim, focando-se na função com número maior de ocorrências, o artigo mostrou um conjunto de possibilidades de análise advindas da perfilação gramatical, tais como a frequência da função, a sua distribuição dentro dos textos, a dispersão ao longo dos textos, bem como os agrupamentos que forma, nos padrões oracionais, com outras funções. Para as pesquisas em Linguística Aplicada – de base sistêmico-funcional – a perfilação gramatical mostrou como é possível compilar e extrair dados de corpora, inclusive para a forma como são empregados na dinâmica textual.

Referências bibliográficas

BERBER SARDINHA, T. Linguística de corpus: histórico e problemática. **D.E.L.T.A.**, v. 16, n. 2, p. 323-367, 2000.

BIBER, D. What can a corpus tell us about registers and genres? In: O'KEEFFE, A.; McCARTHY, M. (Ed.). **The Routledge Handbook of corpus Linguistics**. London, New York: Routledge, 2010. p. 241-254. <https://doi.org/10.4324/9780203856949.ch18>

BRAGA, A. **A popularização da popularização da ciência: gramaticalização da interação, avaliação, representação da experiência, construção e organização discursiva**. 2013. 97 f. Monografia (Bacharelado em Tradução) – Departamento de Letras, Universidade Federal de Ouro Preto, Mariana, 2013.

BUNGE, M. **La investigación científica: Su estrategia y su filosofía**. Barcelona: Ediciones Ariel, 1969.

FIGUEREDO, G. **Introdução ao perfil metafuncional do português brasileiro: contribuições para os estudos multilíngues**. 2011. 385 f. Tese (Doutorado em Linguística Aplicada) – Programa de Pós-Graduação em Estudos Linguísticos, Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2011.

FIGUEREDO, G. Uma metodologia de perfilação gramatical sistêmica baseada em corpus. **Letras & Letras**, v. 30, n. 2, jul./dec., 2014. p. 17-45. <https://doi.org/10.4324/9780203856949.ch18>

GLEASON, H. **Linguistics and English grammar**. New York: Holt, Rinehart & Winston, 1965.

HALLIDAY, M. A. K. **Language as social semiotic: the social interpretation of language and meaning**. London & Baltimore: Edward Arnold & University Park Press, 1978.

_____. Towards Probabilistic Interpretations. In: VENTOLA, E. (Ed.). **Functional and Systemic Linguistics: Approaches and Uses**. Berlin/ New York: Mouton de Gruyter, 1991. p.39-61. <https://doi.org/10.1515/9783110883527.39>

_____. **On grammar**. London: Continuum, 2002.

_____. **Computational and quantitative studies**. London: Continuum, 2005.

_____. Working with meaning: towards an applicable linguistics. In: WEBSTER, J. (Ed.). **Meaning in Context: implementing intelligent applications of language studies**. London and New York: Continuum, 2008.

_____. Pinpointing the choice: meaning and the search for equivalents in a translated text. In: MAHBOOB, A.; KNIGHT, N. (Ed.). **Applicable Linguistics**. London e New York: Continuum, 2010.

HALLIDAY, M. A. K.; MATTHIESSEN, C. **Construing experience as meaning: a language based approach to cognition**. London: Cassell, 1999.

_____. **Halliday's introduction to functional grammar**. 4ed. Oxford e New York: Routledge, 2014. A

HALLIDAY, M. A. K.; McINTOSH, A.; STREVESEN, P. **The linguist sciences and language teaching**. London: Longmans, 1964.

LEMKE, J. Text Production and Dynamic Text Semantics. In: VENTOLA, E. (Ed.). **Functional and systemic linguistics: approaches and uses**. Berlin: Mouton de Gruyter, 1991. <https://doi.org/10.1515/9783110883527.23>

_____. **Discourse, Dynamics, and Social Change**. Cultural Dynamics. v.6, n. 1, p. 243-275, 1993. <https://doi.org/10.1177/092137409300600107>

LIMA, K.C.S.; **Caracterização de registros orientada para a produção textual no ambiente multilíngue: um estudo baseado em corpora comparáveis**. Tese (Doutorado em Linguística Aplicada) – Belo Horizonte: Faculdade de Letras, UFMG/PosLin, 2013.

MAHBOOB e KNIGHT, N. Applicable linguistics: an introduction. In: _____. (Ed.). **Applicable Linguistics**. London e New York: Continuum, 2010.

MARTIN, J. R. **English text: system and structure**. Philadelphia and Amsterdam: John Benjamins Publishing Company, 1992. <https://doi.org/10.1075/z.59>

_____. **Systemic functional grammar: a next step into the theory – axial relations**. Beijing: Higher Education Press, 2013.

MARTIN, J.; MATTHIESSEN, C. Systemic typology and topology. In: CHRISTIE, F. (Ed.). **Literacy in social processes**. Darwin: Centre for Studies in Language in Education, Northern Territory University, 1991.

MARTIN, J.; ROSE, D. **Working with discourse: meaning beyond the clause**. 2. ed. London: Continuum, 2007.

MATTHIESSEN, C. **Register in the round: registerial cartography**. *Funct Linguist* 2015.

MATTHIESSEN, C.; BATEMAN, J. **Text generation and systemic functional linguistics: experiences from English and Japanese**. London: Pinter Publishers, 1991.

MATTHIESSEN, C.; TERUYA, K.; WU, C. Multilingual studies as a multi-dimensional space of interconnected language studies. In: WEBSTER, J. (Ed.). **Meaning in Context: implementing intelligent applications of language studies**. London and New York: Continuum, 2008.

MATTHIESSEN, C.; TERUYA, K.; LAM, M. **Key terms in systemic functional linguistics**. London and New York: Continuum, 2010.

O'DONNELL, M. The UAM CorpusTool: software for corpus annotation and exploration. In: BRETONES CALLEJAS, C. *et al.* (Ed.). In: **Applied linguistics now: understanding language and mind**. Almería: Universidad de Almería, 2008.

OLIVEIRA, F. S. Comparação linguística e perfilação gramatical sistêmica em um corpus combinado. **Revista Estudos da Linguagem**, v. 23, Edição Especial, p. 727-768, 2015.

PAGANO, A. S. et al. Boas Práticas para a Pesquisa nas Ciências da Saúde. In: TORRES, H. D. C.; REIS, I. A.; PAGANO, A. S. **Empoderamento do Pesquisador nas Ciências da Saúde**. Belo Horizonte: FALE/UFMG, 2015. Cap. 9, p. 139-142.

ROSE, D. **The Western Desert Code: an Australian cryptogrammar**. Canberra, The Australian National University: Pacific Linguistics, 2001.

SAUSSURE, F. **Curso de Linguística Geral**. São Paulo: Cultrix, 2006.

TEICH, E. Contrastive linguistics and translation studies revisited. In: GIL, A. (Org.). **Modelle der Translation: Grundlagen für Methodik, Bewertung, Computer Modellierung**. Frankfurt am Main e Berlin: Lang, 1999.

URE, J. **Text types classified by situational factors**. Manuscrito, 1989.

VIANA, V. Linguística de corpus: conceitos, técnicas e análises. In: VIANA, V.; TAGNIN, S. (Org.). **Corpora no ensino de línguas estrangeiras**. São Paulo: HUB Editorial, 2011, p. 25-95.

WINOGRAD, T. **Understanding natural language**. New York: Academic Press, 1972.

WINOGRAD, T. **Shifting viewpoints: Artificial intelligence and human-computer interaction**. USA: Elsevier B. V., 2006, p. 1256-1258.

Bibliografia

CÂNDIDO-JR, A. **Análise bidirecional da língua na simplificação sintática em textos do português voltada à acessibilidade digital**. 2013. 225f. Tese (Doutorado em Ciências de Computação e Matemática Computacional). Instituto de Ciências Matemáticas e da Computação, Universidade de São Paulo, São Carlos, 2013.

Artigo recebido em: 15.12.2016

Artigo aprovado em: 15.04.2017