



UNIVERSIDADE FEDERAL
DO RIO DE JANEIRO

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

FACULDADE DE LETRAS

**CONSTRUÇÕES HIPOTÁTICAS COM [VPP QUE] EM PORTUGUÊS E ITALIANO:
UMA ANÁLISE CONTRASTIVA BASEADA NO USO**

Juliana Barboza do Nascimento

Rio de Janeiro

2024

**CONSTRUÇÕES HIPOTÁTICAS COM [VPP QUE] EM PORTUGUÊS E ITALIANO:
UMA ANÁLISE CONTRASTIVA BASEADA NO USO**

JULIANA BARBOZA DO NASCIMENTO

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Linguística da Faculdade de Letras da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários à obtenção do Título de Mestre em Linguística.

Orientadora: Prof.^a. Dr.^a. Maria Maura Cezario

Coorientador: Prof. Dr. Dennis Castanheira

Rio de Janeiro

2024

CIP - Catalogação na Publicação

N244c Nascimento, Juliana Barboza do
Construções hipotáticas com [Vpp] que em
português e italiano: uma análise contrastiva
baseada no uso / Juliana Barboza do Nascimento. -
Rio de Janeiro, 2024.
98 f.

Orientadora: Maria Maura da Conceição Cezario .
Coorientador: Dennis da Silva Castanheira.
Dissertação (mestrado) - Universidade Federal do
Rio de Janeiro, Faculdade de Letras, Programa de Pós
Graduação em Linguística, 2024.

1. Linguística Funcional Centrada no uso. 2.
Língua portuguesa. 3. Língua italiana. 4. Orações
hipotáticas. 5. Conectivos. I. Cezario , Maria
Maura da Conceição , orient. II. Castanheira, Dennis
da Silva , coorient. III. Título.

Elaborado pelo Sistema de Geração Automática da UFRJ com os dados fornecidos pelo(a)
autor(a), sob a responsabilidade de Miguel Romeu Amorim Neto - CRB-7/6283.

**CONSTRUÇÕES HIPOTÁTICAS COM [VPP QUE] EM PORTUGUÊS E ITALIANO:
UMA ANÁLISE CONTRASTIVA BASEADA NO USO**

Juliana Barboza do Nascimento

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Linguística da Faculdade de Letras da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários à obtenção do Título de Mestre em Linguística.

BANCA EXAMINADORA

Maria Maura da C. Cezario

Presidente Prof.^a Dr.^a Maria Maura Cezario
Universidade Federal do Rio de Janeiro (UFRJ)

Coorientador Prof. Dr. Dennis Castanheira
Universidade Federal Fluminense (UFF)

Prof.^a Dr.^a Karen Sampaio Braga Alonso
Universidade Federal do Rio de Janeiro (UFRJ)

Prof.^a Dr.^a Carla Valeria de Souza Faria
Università Degli Studi Di Trieste - Itália

Prof.^a Dr.^a Violeta Virginia Rodrigues (Suplente)
Universidade Federal do Rio de Janeiro (UFRJ)

Prof.^a Dr.^a Deise Cristina de Moraes Pinto (Suplente)
Universidade Federal do Rio de Janeiro (UFRJ)

Rio de Janeiro

2024

AGRADECIMENTOS

Começo essa seção de agradecimentos com um sentimento latente em mente: alívio. Apenas as pessoas que estiveram ao meu lado nos últimos dois anos sabem o quão difícil foi trilhar o caminho do mestrado. Diversas dificuldades surgiram no caminho, mas apesar de tudo, eu consegui e isso só foi possível graças ao apoio de pessoas maravilhosas.

Roberto e Mônica, meus pais, vocês, mais do que ninguém, acompanharam todos os meus temores e todas as minhas crises relacionadas às minhas inquietações acadêmicas. Desde sempre, vocês me apoiaram naquilo que decidi fazer e nunca me deixaram desistir. Portanto, agradeço os conselhos, os momentos de colo e também os momentos de vibração a cada conquista minha. Todos esses sentimentos se estendem aos demais membros da **minha família**: minhas tias, meus tios, meu primo e meus avós, que nunca esconderam o orgulho e a felicidade sempre que eu avançava em mais uma etapa.

Maura, nossa parceria é antiga, mas meu carinho e admiração por você apenas crescem exponencialmente. Agradeço imensamente toda a paciência, todo o cuidado e tudo o que você fez por mim desde sempre. Sabemos o quanto foi difícil concluir essa etapa, mas te agradeço acima de tudo por não ter desistido de mim e não ter duvidado de meu potencial. Você continua sendo meu exemplo a se espelhar dentro e fora da academia. Obrigada por dividir comigo seus conhecimentos e se dedicar ao que construímos juntas ao longo desses anos.

Dennis, apesar da distância espacial que o tempo nos trouxe, você ainda é meu anjo da guarda. Sempre me sinto honrada ao lembrar que pude ter você como amigo e orientador durante essa trajetória. Você é um pesquisador admirável e um exemplo de professor que espero me tornar um dia. Obrigada por ser tão compreensivo, carinhoso e por confiar em mim. Só posso te agradecer infinitamente por tudo o que você fez e ainda faz por mim.

Estendo também o agradecimento a todos aqueles que cruzaram meu caminho e, de alguma maneira, contribuíram para minha formação durante o mestrado, sobretudo os professores do PPGLIN da UFRJ, os docentes e discentes do Discurso e Gramática, meu querido grupo de pesquisa, e também ao CNPq e a FAPERJ que financiaram e tornaram essa pesquisa possível.

Grazielle, obrigada por ser minha outra metade, por entender todos os meus sumiços e aparições repentinas, por me confortar sempre que eu precisava, por acreditar em mim, por me trazer alegria e por segurar minha mão em todos os momentos. Sou muito grata por existir no mesmo momento no mundo que você e por ter te encontrado.

Nicolas, você é uma das pessoas que mais acredita no meu potencial e é o melhor amigo que eu poderia pedir, disso eu tenho certeza. Obrigada por me ouvir e me incentivar em todos os momentos. Sem o seu apoio incondicional, eu não seria um terço da pessoa que eu sou.

Sara, como sempre você continua sendo minha guia e minha inspiração para me tornar uma profissional melhor e uma pesquisadora competente. Eu te agradeço imensamente por ter segurado minha mão nos momentos que eu achei que concluir essa pesquisa seria impossível. Obrigada por ler, revisar, me tirar do bloqueio que me tomou e por me ajudar a trilhar esse caminho. Acima de tudo, obrigada por ter me dado forças para não desistir.

Caroline, você sabe que existe um certo milagre nos encontros e, neste caso, eu sou muito grata por ter encontrado algo tão sagrado. Te agradeço imensamente por estar ao meu lado e por sempre me incentivar tanto. Suas palavras de apoio com certeza me impulsionaram

para que eu conseguisse concluir essa etapa e para que eu siga tentando o meu melhor. Obrigada por tudo o que você faz por mim e por me trazer tanta felicidade.

Brenda, eu nunca poderia dimensionar o carinho, a gratidão e o privilégio que eu tenho por ser sua amiga. Você provavelmente é a pessoa que mais acredita no meu potencial e, às vezes, ainda me pego surpresa com a sorte que foi conhecer você. Te agradeço por tudo o que você faz por mim. Seu apoio, seus conselhos e suas ajudas são inestimáveis. Obrigada por ler tudo o que eu escrevo, desde dos textos acadêmicos e ficcionais até as narrações mais básicas do meu dia. Obrigada por existir, ser minha amiga e escolher ficar ao meu lado todos os dias.

Paola, eu simplesmente não consigo mais me lembrar de como era minha vida antes de conhecer você. Todos os dias eu acordo e penso que gostaria de ser como você quando crescer e olha que eu já passei dessa fase tem tempo, mas não posso evitar quando você é uma das pessoas mais inspiradoras que eu conheço. Espero continuar ao seu lado para aplaudir todas as suas conquistas. Obrigada por ser minha razão, meu conforto e o inusitado que me salvou.

Amanda, eu não consigo compreender como em um ser humano tão pequeno existe tanta grandeza quanto dentro você. O carinho e a compreensão que compartilhamos uma com a outra é inestimável para mim. Você não tem ideia de como sua simples existência já faz meu dia mais feliz. Obrigada por entender minha loucura, por abrigar minhas paranoias, por sempre me ouvir e por ser uma pessoa tão incrível e que eu tenho orgulho de ter como minha amiga.

Mariana, nos últimos anos, nossas conversas semanais são o que mantém meu foco e minha própria sanidade, como você bem sabe. Finalmente esta etapa foi concluída e nós duas sabemos perfeitamente o quão difícil, mas te agradeço imensamente por me enxergar de uma forma tão positiva e por me incentivar a realizar esse exercício também. Obrigada todos os conselhos, todos os freios, todos os impulsionamentos e por todas as reflexões.

Por fim, agradeço a todos aqueles que de alguma maneira contribuíram para que eu chegasse até aqui. Fazendo empréstimos de minhas próprias palavras em minha monografia, espero que a busca por conhecimento seja o que continue me movendo e que eu nunca me esqueça da capacidade de transformação e dos sentimentos de liberdade e esperança que a educação pode proporcionar.

RESUMO

Esta dissertação tem como objetivo geral analisar comparativamente as construções oracionais em português e em italiano introduzidas por “dado que”, “posto que”, “dato che” e “posto che”, no que diz respeito às propriedades funcionais e discursivas, com o aparato teórico da Linguística Funcional Centrada no Uso, a qual une pressupostos do funcionalismo norte-americano (Hopper, 1987; Givón, 1984, 1995; Bybee, 2010) e o modelo de gramática proposto pela Gramática de Construções (Goldberg, 1995, 2006, 2019; Croft, 2001, Rosário; Oliveira, 2016; Pinheiro, 2020, Furtado da Cunha; Cezario, 2023). Adotamos o Princípio da Não-Sinonímia (Goldberg, 1995) para guiar a nossa hipótese geral de que, apesar de se inserirem em contextos semelhantes, essas construções apresentam especificidades estruturais e semântico-pragmáticas em cada língua e entre as línguas. Se há diferenças na forma, existe, em algum grau, uma função comunicativa diferente. Entendemos que as construções em análise fazem parte do esquema mais abstrato [[CONNECT] S V O]ORAÇÃO HIPOTÁTICA ADVERBIAL, em que CONNECT é o conectivo que introduz a oração; S é o sujeito; V o verbo e O o complemento ou adjunto. Em nossa pesquisa, CONNECT terá sempre a forma de [Xque], e assim temos o seguinte esquema [[(X) que] S V O]ORAÇÃO HIPOTÁTICA ADVERBIAL e desse esquema são instanciados outros como o adverbial condicional, causal e concessivo. Nesta tese, focaremos no esquema concessivo [[(V_{pp}) que]ORAÇÃO HIPOTÁTICA ADVERBIAL, no qual o *slot* V_{pp} é ocupado por verbos no particípio passado, neste caso, por “dado” e “posto” em português e “dato” e “posto” em italiano. Metodologicamente, foi realizada uma análise quali-quantitativa. Os dados foram retirados do Corpus do Português e do Corpus PAISÀ, na modalidade escrita, sendo textos de plataformas online, como blogs e enciclopédias, produzidos entre 2010 e 2014. Coletamos 360 dados, sendo 90 para cada construção estudada. A partir dos resultados encontrados, foi possível verificar que as construções estudadas, apesar de serem semelhantes na forma e no significado, possuem especificidades particulares de uso no contexto real de comunicação. Além disso, a pesquisa trouxe resultados relevantes para compreendermos como se definem *links* entre construções de um mesmo esquema e como falantes de diferentes línguas irmãs escolhem as construções.

Palavras-chave: Orações hipotáticas. Conectivos. Linguística Funcional Centrada no Uso.

Rio de Janeiro
Setembro de 2024

ABSTRACT

This dissertation has the general objective of comparatively analyzing the clause constructions in Portuguese and Italian introduced by “*dado que*”, “*posto que*”, “*dato che*” and “*posto che*”, with regard to functional and discursive properties, using the theoretical framework of Usage-Based Linguistics, which combines assumptions from North American functionalism (Hopper, 1987; Givón, 1984, 1995; Bybee, 2010) and the Construction Grammar model (Goldberg, 1995, 2006, 2019; Croft, 2001, Rosário; Oliveira, 2016; Pinheiro, 2020, Furtado da Cunha; Cezario, 2023). We adopted the Principle of Non-Synonymy (Goldberg, 1995) to guide our general hypothesis that, despite being inserted in similar contexts, these constructions have structural and semantic-pragmatic specificities. If there are differences in form, there is, to some degree, a different communicative function. We understand that the constructions under analysis are part of the more abstract scheme [[CONNECT] S V O] HYPOTACTIC ADVERBIAL CLAUSE, in which CONNECT is the connective that introduces the clause; S is the subject; V is the verb and O is the complement or adjunct. In our research, CONNECT always have the form of [Xque], and thus we have the following scheme [[(X) que] S V O] HYPOTACTIC ADVERBIAL CLAUSE and, from this scheme, others are instantiated such as the conditional, causal and concessive adverbial clauses. In this research, we will focus on the concessive scheme [[(V_{pp}) que] HYPOTACTIC ADVERBIAL CLAUSE, in which the V_{pp} slot is occupied by verbs in the past participle, in this case, by “*dado*” and “*posto*” in Portuguese and “*dato*” and “*posto*” in Italian. Methodologically, a qualitative and quantitative analysis was carried out. The data were taken from the Corpus do Português and the Corpus PAISÀ, in written form, being texts from online platforms, such as blogs and encyclopedias, produced between 2010 and 2014. We collected 360 pieces of data, 90 for each construction studied. From the results found, it was possible to verify that the constructions studied, despite being similar in form and meaning, have particular specificities of use in the real context of communication. Furthermore, the research brought relevant results to understand how links between constructions of the same scheme are defined and how speakers of different sister languages choose their constructions.

Keywords: Hypotactic clauses. Connectives. Usage-Based Functional Linguistics.

Rio de Janeiro
Setembro de 2024

RIASSUNTO

Questa tesi ha l'obiettivo generale di analizzare comparativamente le costruzioni proposizionali in portoghese e italiano introdotte da “dado que”, “posto que”, “dato che” e “posto che”, con riguardo alle proprietà funzionali e discorsive, con l'apparato teorico della Linguistica Funzionale Centrata sull'Uso, che combina i presupposti del funzionalismo nordamericano (Hopper, 1987; Givón, 1984, 1995; Bybee, 2010) e il modello della Grammatica delle Costruzioni (Goldberg, 1995, 2006, 2019; Croft, 2001, Rosario; Oliveira, 2016; Abbiamo adottato il Principio della Non Sinonimia (Goldberg, 1995) per guidare la nostra ipotesi generale in cui, pur essendo inserite in contesti simili, queste costruzioni hanno specificità strutturali e semantico-pragmatiche. Se ci sono differenze nella forma, c'è, in una certa misura, una diversa funzione comunicativa. Si comprende che le costruzioni in analisi fanno parte dello schema più astratto [[CONNECT] S V O] PROPOSIZIONE AVVERBIALE IPOTATTICA, in cui CONNECT è il connettivo che introduce la proposizione; S è il soggetto; V è il verbo e O è il complemento o l'aggiunto. Nella nostra ricerca CONNECT avrà sempre la forma di [Xque], e quindi abbiamo il seguente schema [[(X) que] S V O] PROPOSIZIONE AVVERBIALE IPOTATTICA e da questo schema se ne istanziano altri come l'avverbiale condizionale, causale e concessiva. In questa ricerca ci concentreremo sullo schema concessivo [(V_{pp}) que] PROPOSIZIONE AVVERBIALE IPOTATTICA, in cui lo *slot* V_{pp} è occupato dai verbi al participio passato, in questo caso da “dado” e “posto” in portoghese e “dato” e “posto” in italiano. Metodologicamente è stata effettuata un'analisi qualitativa e quantitativa. I dati sono stati presi dal Corpus do Português e dal Corpus PAISÀ, in forma scritta, ovvero testi provenienti da piattaforme online, come blog ed enciclopedie, prodotti tra il 2010 e il 2014. Abbiamo raccolto 360 dati, 90 per ogni costruzione studiata. Dai risultati incontrati è stato possibile verificare che le costruzioni studiate, pur essendo simili nella forma e nel significato, presentano particolari specificità di utilizzo nel contesto reale della comunicazione. Inoltre, la ricerca ha portato risultati rilevanti per comprendere come si definiscono i collegamenti tra costruzioni dello stesso schema e come i parlanti di lingue sorelle diverse scelgono le costruzioni.

Parole chiave: Proposizioni ipotattiche. Connettivi. Linguistica Funzionale Centrata Sull'Uso.

Rio de Janeiro
Setembro de 2024

LISTA DE TABELAS

Tabela 1: Valor semântico das orações hipotáticas em português	48
Tabela 2: Valor semântico das orações hipotáticas em italiano	49
Tabela 3: Ordenação das orações hipotáticas em português	52
Tabela 4: Ordenação das orações hipotáticas em italiano	52
Tabela 5: <i>Status</i> informacional das orações hipotáticas em português	59
Tabela 6: <i>Status</i> informacional das orações hipotáticas em italiano	60
Tabela 7: Codificação formal dos sujeitos das orações hipotáticas em português.....	71
Tabela 8: Codificação formal dos sujeitos das orações hipotáticas em italiano.....	72
Tabela 9: Correferencialidade entre os sujeitos das orações matrizes e hipotáticas em português	74
Tabela 10: Correferencialidade entre os sujeitos das orações matrizes e hipotáticas em italiano	75
Tabela 11: Frequência de tipo e de ocorrência dos itens verbais da oração hipotática em português	76
Tabela 12: Frequência de tipo e de ocorrência dos itens verbais da oração hipotática em italiano	78
Tabela 13: Comparação entre os usos das orações hipotáticas em português e italiano	81

LISTA DE GRÁFICOS

Gráfico 1: Cruzamento entre valor semântico e ordenação com [dado que]or	54
Gráfico 2: Cruzamento entre valor semântico e ordenação com [posto que]or	55
Gráfico 3: Cruzamento entre valor semântico e ordenação com [dato che]or	56
Gráfico 4: Cruzamento entre valor semântico e ordenação com [posto che]or	56
Gráfico 5: Cruzamento entre <i>status</i> informacional e ordenação com [dado que]or.....	61
Gráfico 6: Cruzamento entre <i>status</i> informacional e ordenação com [posto que]or.....	62
Gráfico 7: Cruzamento entre <i>status</i> informacional e ordenação com [dato che]or.....	63
Gráfico 8: Cruzamento entre <i>status</i> informacional e ordenação com [posto che]or	63
Gráfico 9: Simultaneidade modo-temporal das orações hipotáticas em português.....	67
Gráfico 10: Simultaneidade modo-temporal das orações hipotáticas em italiano	68

LISTA DE FIGURAS

Figura 1: Rede esquemática da construção [(X) que] no português brasileiro	17
Figura 2: Estrutura simbólica de uma construção	27
Figura 3: Níveis de esquematicidade	30
Figura 4: Relação entre valor semântico e ordenação	36
Figura 5: O <i>continuum</i> entre as orações concessivas, condicionais e causais.....	38
Figura 6: <i>Corpus</i> do Português.....	41
Figura 7: <i>Corpus</i> Paisà – Pesquisa Avançada	42
Figura 8: Rede semântica dos conectivos em português.....	85
Figura 9: Rede semântica dos conectivos em italiano.....	86

LISTA DE QUADROS

Quadro 1: Processos de combinação de orações.....	33
Quadro 2: Correspondência de tempos e modos verbais entre Português e Italiano	65
Quadro 3: Fatores de integração entre orações com as construções em análise	92

SUMÁRIO

LISTA DE TABELAS.....	10
LISTA DE GRÁFICOS.....	11
LISTA DE FIGURAS	12
LISTA DE QUADROS	13
INTRODUÇÃO.....	15
1. PRESSUPOSTOS TEÓRICOS	21
1.1. Linguística Funcional Centrada no Uso.....	21
1.2 Processos cognitivos de domínio geral	26
1.3 O conceito de construção e as suas características	26
1.4 Esquematicidade, composicionalidade e produtividade	29
2. REVISÃO DA LITERATURA	32
2.1 Encaixamento de orações.....	32
2.2 Grau de integração entre as orações.....	34
2.3 Ordenação, valor semântico e status informacional	34
2.4 Os conectivos “dado que”, “posto que”, “dato che” e “posto che”	36
2.5 O valor semântico: o <i>continuum</i> das orações causais, condicionais e concessivas	38
2 METODOLOGIA.....	41
2.1 <i>Corpus</i> do Português e <i>Corpus PAISÀ</i>	41
2.2 Etapas e fatores de análise	42
3. ANÁLISE	44
3.1 Valor semântico, ordenação e status informacional da oração hipotática	44
3.1.1 Valor semântico.....	44
3.1.2 Ordenação da oração hipotática	50
3.1.3 Valor semântico x Ordenação	53
3.1.4 <i>Status</i> informacional.....	57
3.1.4 <i>Status</i> informacional x Ordenação	60
3.2 Elementos de integração	64
3.2.1 Simultaneidade Modo-Temporal.....	65
3.2.2 Codificação formal do sujeito da oração hipotática	68
3.2.3 Correferencialidade dos sujeitos da oração hipotática e da oração matriz.....	73
3.3 Frequência de tipo e de ocorrência dos itens verbais da oração hipotática	75
4 REFLEXÕES SOBRE A ANÁLISE COMPARATIVA: CONSIDERAÇÕES FINAIS	81
REFERÊNCIAS BIBLIOGRÁFICAS	95

INTRODUÇÃO

Esta dissertação se propõe a analisar as construções oracionais hipotáticas introduzidas por “dado que” e “posto que”, no português brasileiro, e “dato che” e “posto che”, no italiano, com o objetivo de mapear seus contextos de usos e, conseqüentemente, realizar um estudo contrastivo entre ambas as línguas. Baseando-se no Princípio da Não-Sinonímia (Goldberg, 1995), que postula que não existem formas diferentes que evidenciem significados iguais, postulamos que as construções conectoras do português devem ter suas particularidades de uso, assim como as construções conectoras em italiano.

Sendo assim, realizamos um estudo sincrônico sobre essas construções, descrevendo o modo como são usadas nos dois idiomas, em termos de frequência de uso, especificidades de forma-sentido e a maneira como tais construções se organizam em suas respectivas redes linguísticas, tendo como base a Linguística Funcional Centrada no Uso (cf. Barlow; Kemmer 2000; Bybee, 2010; Cezario; Furtado, 2013; Diessel, 2015; Goldberg, 2006; Hilpert, 2014; Traugott, 2008).

Por ser um estudo sobre construções, se torna relevante mencionar alguns princípios fundamentais que regem os estudos construcionistas (cf. Bybee, 2010; Goldberg, 1995; 2006; Hilpert 2014; Traugott; Trousdale, 2013). São eles: (I) a língua é um sistema de pareamentos de forma-sentido convencionalizados e relacionados entre si, abarcando desde itens lexicais a construções sintagmáticas; (II) as estruturas das línguas emergem a partir do uso linguístico; (III) a língua é formada por uma rede de nós e pelas ligações entre esses nós; (IV) o conhecimento da língua é um inventário de construções; (V) os processos cognitivos de domínio geral regem a cognição humana, sendo responsáveis pela linguagem e por outras áreas da cognição.

A motivação deste trabalho se dá a partir da monografia de Nascimento (2022), na qual foi realizado um estudo sobre os usos das construções iniciadas pelos conectivos “visto que”, “dado que” e “posto que” no português brasileiro, sob uma perspectiva sincrônica. Foram encontradas diferenças de usos dessas construções e, a partir disso, hipotetizamos que, como as línguas são parte da mesma família linguística, sendo oriundas do latim, deveriam haver características em comum entre seus usos.

Por outro lado, sendo línguas diferentes, haveria também possivelmente diferenças no uso dos conectores em estudo. É necessário levar em conta que, apesar das construções estudadas nessa pesquisa serem consideradas virtualmente equivalentes ao se comparar uma língua com a outra, podendo ser consideradas sinônimas em uma perspectiva hegemônica, é

importante destacar que é impossível transpor integralmente um sentido dado a uma construção de uma língua pra outra, uma vez que não há um significado estável e único. Cada sistema linguístico possui suas próprias especificidades, desta forma, da mesma maneira como não existe um sinônimo perfeito (Goldberg, 1985), não existiria também uma tradução que reflita perfeitamente a semântica que uma construção carrega.

Segundo Cezario, Santos Silva e Santos (2015), as orações introduzidas por esses conectivos poderiam ser esquematizadas como [[(X) que] (S) V (O)]. Nesse esquema, o *slot X* pode ser preenchido por diferentes unidades linguísticas, como “toda vez”, “se bem” ou “ainda”, por exemplo, instanciando as construções “toda vez que”, “se bem que” e “ainda que”. Ainda referente a esse esquema, o “S” representa o possível sujeito desta oração¹, “V” o verbo e “O” o possível objeto ou complemento verbal.

Defendemos que um esquema semelhante exista na rede linguística do italiano, porém, além de suas especificidades semântico-pragmáticas, haveria também uma sutil diferença em relação a estrutura, considerando que o conectivo “que” possui uma grafia diferente na língua italiana, sendo este grafado como “che”. Desta forma, o esquema em italiano seria codificado como [[(X) che] (S) V (O)].

Em ambos esses esquemas, haveria ainda um outro nível de abstração, codificado como [(V_{pp}) que] (S) V (O)], em português, e [(V_{pp}) che] (S) V (O)], em italiano, isto porque as construções estudadas nesta pesquisa têm como característica comum o *slot X* sendo preenchido por uma forma verbal originalmente no particípio passado².

Na rede taxonômica, num nível mais baixo, teríamos, por fim, os padrões construcionais, ou melhor, as construções que são investigadas nesta pesquisa: [(dado) que] (S) V (O)], [(posto) que] (S) V (O)], em língua portuguesa, e [(dato) que] (S) V (O)] e [(posto) che] (S) V (O)], em língua italiana. Para facilitar nossa exposição, cada construção oracional estudada será denominada de microconstrução da construção esquemática maior.

Também é necessário ressaltar que, apesar do enfoque no esquema que codifica abstratamente as orações hipotáticas, as orações principais ligadas a elas são igualmente relevantes para esta pesquisa, pois não existe hipotaxe sem oração principal. Desta forma, os

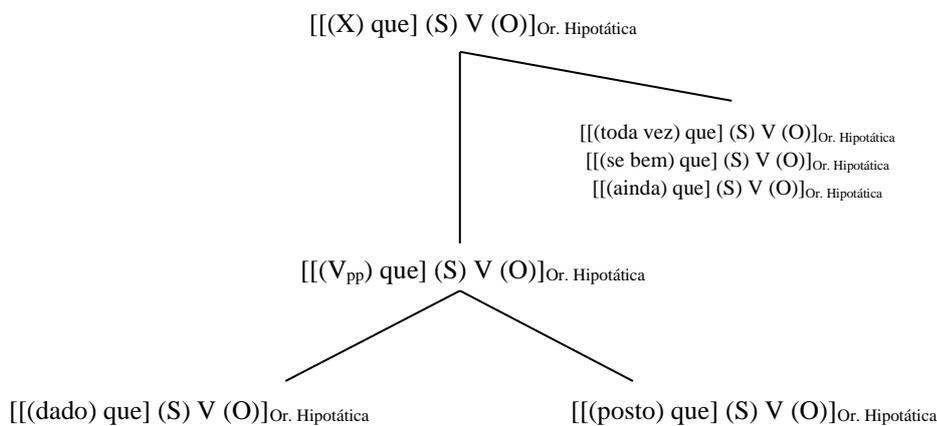
¹ No caso deste trabalho, o *slot* do sujeito não é de preenchimento opcional, visto que foram considerados apenas os dados em que o sujeito da oração hipotática e da principal existem, devido a sua relevância para alguns dos fatores investigados nesta pesquisa.

² É importante destacar que essa codificação de subesquema foi escolhida apenas como forma de agrupar esses conectivos, já que as sequências [V_{pp} que] e [V_{pp} che] teriam já sofrido um processo de *chunking*, e, a partir disso, o *slot* [V_{pp}] não é mais percebido como um espaço disponível para alocar verbos particípios passados, já que as construções instanciadas por esses esquemas já são entendidas como um único bloco linguístico pelos usuários de ambas as línguas.

esquemas investigados nesta pesquisa são ainda maiores, considerando que contemplamos em nossos dados tanto a oração hipotática instanciada por [X que] quanto a oração principal ligada a ela, que pode aparecer antes, depois ou perpassada por esta construção hipotática, como veremos no capítulo de Análise.

Como anteriormente mencionado, de acordo com essa perspectiva, a mente do falante funciona como um grande arcabouço de construções que são conectadas entre si. As relações instauradas pelos esquemas supramencionados podem ser melhor observadas na figura a seguir:

Figura 1: Rede esquemática da construção [(X) que] no português brasileiro



Fonte: Adaptado de Nascimento (2022)

Na figura, podemos observar as microconstruções estudadas nesta pesquisa, referentes à língua portuguesa. Sendo um padrão bastante produtivo, a construção [X que] abrange diferentes semânticas e aparece em plurais contextos (Cezario, Santos Silva e Santos, 2015). Além disso, devido aos processos cognitivos de domínio geral (Bybee, 2010), como a analogia e o *chunking*, esse padrão construcional tem seu *slot* preenchido por diferentes unidades linguísticas, como exemplificado alguns parágrafos acima. Sendo assim, não é de se surpreender que um padrão semelhante (praticamente idêntico a respeito da estrutura) seja encontrado em outras línguas românicas, como é o caso do italiano.

Ademais, é importante destacar que a rede esquemática supramencionada se relaciona apenas com o nível estrutural dessas construções, não levando em conta os aspectos semântico-pragmáticos, como valor semântico ou função, que também regulam e distribuem a forma como tais esquemas se organizam dentro da mente do usuário da língua.

Por fim, ainda buscando uma simplificação para a codificação das construções aqui investigadas, assumiremos a nomenclatura [dado que]or. como equivalente ao subesquema [[(dado) que] (S) V (O)]Or. Hipotática. O mesmo sendo válido para as demais microconstruções.

Tendo isso estabelecido, a seguir apresentamos alguns exemplos das microconstruções estudadas nesta pesquisa:

(1)“Uma vez já escrevi que não sei o que é avaliar arte, **dado que é muito pessoal e subjetivo**”. (*Corpus* do Português)

(2)“Gostaria de ter uma resposta firme à minha indagação, **posto que a legenda é dúbia**”. (*Corpus* do Português)

(3)“Imparare il Napoletano può risultare difficile, **dato che non esiste una grammatica convenzionata** e non c’è nessuna autorità che ne stabilisca il vocabolario e la relativa grammatica”. (*Corpus* PAISÀ)

[Aprender o napolitano pode ser difícil, **dado que não existe uma gramática convencionalizada** e não tem nenhuma autoridade que estabeleça o vocabulário e a relativa gramática.]

(4)“Dunque la retorica può essere concepita come un settore della pragmatica, **posto che la persuasione può essere intesa come una delle possibili azioni eseguibili con il linguaggio**”. (*Corpus* PAISÀ)

[Então a retórica pode ser concebida como um setor da pragmática, **já que a persuasão pode ser entendida como uma das ações possíveis que podem ser realizadas com a linguagem.**]

Neste trabalho, foram analisados 180 dados de língua portuguesa e 180 dados em língua italiana, totalizando 360 ocorrências, sendo 90 dados para cada uma das microconstruções. Os dados em português foram coletados do *Corpus* do Português enquanto os do italiano foram retirados do *Corpus* PAISÀ, ambas plataformas online que reúnem diversos textos escritos da primeira metade da década de 2010, como será mais bem exposto na Metodologia.

Sendo assim, assumimos a hipótese geral de que, apesar de o português brasileiro e o italiano serem duas línguas de origem latina, o funcionamento semântico-pragmático dessas microconstruções nessas duas línguas é diferente, porque, de acordo com a abordagem baseada no uso, construções são específicas de cada língua e variações de sentido e de forma estão sempre ocorrendo devido ao caráter eminentemente dinâmico que constitui qualquer sistema linguístico.

Nessa perspectiva, os usuários memorizam os contextos de uso de cada construção, levando em conta suas diferenças, ainda que mínimas, de forma e função (Bybee, 2010; Diessel, 2019, Traugott; Dasher, 2002). Entretanto, como línguas provindas da mesma língua mãe, há também muitos aspectos gramaticais e funcionais em comum.

Levando isso em conta, pontuamos, como objetivos gerais para esta pesquisa, descrever e analisar as propriedades formais e semântico-pragmáticas de cada uma das microconstruções e, conseqüentemente, realizar uma comparação dessas microconstruções, tanto

intralinguisticamente, visando evidenciar as principais semelhanças e diferenças das microconstruções dentro de um único sistema linguístico, assim como entre a língua portuguesa e a italiana. Para isso, definimos os seguintes objetivos específicos:

a) analisar e descrever as relações semânticas instauradas por cada uma das microconstruções conquanto às suas orações principais;

b) analisar a relação entre valor semântico, ordem e informatividade das orações hipotáticas, tendo como base os estudos de Diessel (2013) e Neves (1999);

b.1) observar as preferências colocacionais de cada microconstrução, considerando a ordem das orações hipotáticas em relação às orações matrizes;

b.2) observar a informatividade das orações hipotáticas, mapeando se tais microconstruções veiculam (ou não) informações pressupostas pragmaticamente;

c) medir o grau de integração entre as orações hipotáticas e suas principais, identificando suas preferências sintáticas, de acordo com os estudos de Givón (1975);

c.1) analisar o tempo e modo do verbo tanto da oração hipotática quanto da oração principal e verificar se são ou não simultâneos;

c.2) classificar como se codifica o sujeito da oração hipotática;

c.3) verificar se o sujeito da oração hipotática e da oração principal são correferentes;

d) analisar, a partir da proposta de Bybee (2010), a frequência de tipo e de ocorrência dos itens verbais da oração hipotática que aparecem com cada microconstrução.

A partir disso, elaboramos, por consequência, as seguintes hipóteses específicas, tomadas como norteadoras para esta pesquisa:

a) presume-se que, de acordo com as gramáticas consultadas para esta pesquisa, [dado que]or. apresente um valor condicional ou causal, enquanto [posto que]or. Tenha um valor concessivo. Em relação ao italiano, espera-se que [dato che]or. exprima um valor causal, enquanto [posto che]or. seja encontrada com o valor condicional;

b) a intenção comunicativa do falante, bem como o tipo de informação veiculada pela oração hipotática, influenciam nas posições dos conectores, de acordo com os estudos realizados por Diessel (2013) e Neves (1999);

b.1) em relação ao vínculo semântico e a ordenação, esperamos que orações condicionais sejam antepostas, visto que normalmente primeiro se enuncia a ocorrência de uma condição e depois sua consequência; conquanto às orações causais postulamos que elas sejam pospostas, pois normalmente se enuncia primeiro o efeito, expressado pela oração principal, e depois a causa, expressada pela oração hipotática; enquanto isso, a respeito das orações concessivas, esta

ordem depende do objetivo de quem a produz, pois tendem a aparecer antepostas quando possuem uma função de tópico, em que se retoma aquilo que já foi dito, e pospostas quando tem a função de adendo³;

b.2) conquanto a informatividade e a ordenação das orações hipotáticas, em seu uso canônico, as orações adverbiais antepostas têm a função de apresentar informações que são pragmaticamente pressupostas, fornecendo uma base temática para novas informações afirmadas em orações subsequentes. Sendo assim, presumimos que orações antepostas tenham tendência à pressuposição. Enquanto isso, as orações adverbiais pospostas são potencialmente mais independentes da oração principal e por isso tendem a fornecer novas informações. Desta forma, postulamos que orações pospostas tendem a ser não-pressupostas;

c) postula-se que uma maior integração semântico-pragmática implica também em uma maior integração sintática entre as orações, isto pode ser medido através da simultaneidade modo-temporal dos eventos tanto na oração hipotática quanto na principal, a correferencialidade entre os sujeitos de ambas as orações e também pela codificação formal do sujeito na oração hipotática em forma de anáfora. Sendo assim, sabendo que cada microconstrução tem suas próprias especificidades estruturais, esperamos que elas apresentem diferentes graus de integração entre suas orações hipotáticas e suas principais;

d) devido a pesquisa realizada por Nascimento (2022), que investigou alguns desses conectivos previamente em língua portuguesa, e considerando também a semelhança estrutural compartilhada por essas línguas, espera-se encontrar um número significativo de verbos de ligação preenchendo o *slot* V das orações hipotáticas, tanto em português quanto em italiano;

d.1) além disso, também acreditamos que a frequência de tipo desses verbos seja divergente e que uma determinada microconstrução apresente um número superior de verbos que pode preencher esse *slot* em comparação com a(s) demais, demonstrando diferentes níveis de produtividade.

Esta dissertação é estruturada da seguinte forma: o capítulo 1 apresenta os Pressupostos Teóricos relevantes para este estudo, trazendo também uma breve revisão da literatura acerca de trabalhos predecessores, que são essenciais para o desenvolvimento dessa pesquisa; no capítulo 2, explicitaremos a metodologia envolvida nesse trabalho e as características dos *corpora* utilizados; o capítulo 3 detalha a análise qualitativa e quantitativa efetuada; e, por fim, nos últimos capítulos, serão apresentadas as considerações finais e as referências bibliográficas.

³ Para que houvesse uma análise ainda mais completa acerca da hipótese sobre o comportamento deste conectivo, seria importante a presença de um terceiro fator: a função. Porém, como não foi incorporado na presente pesquisa, este fator ainda é um possível objeto de investigação para trabalhos futuros.

1. PRESSUPOSTOS TEÓRICOS

Neste capítulo serão apresentados os pressupostos teóricos que norteiam esta pesquisa, havendo um principal destaque para a Linguística Funcional Centrada no Uso (LFCU). Na primeira seção, trazemos alguns dos preceitos básicos pertencentes a essa abordagem funcional-cognitiva, depois seguiremos para a explicação dos mecanismos de domínio geral e sua influência no uso e, por fim, apresentaremos o conceito de construção e a relação da Gramática de Construções com o fenômeno da variação.

1.1. Linguística Funcional Centrada no Uso

A Linguística Funcional Centrada no Uso postula que existe uma estreita relação entre a estrutura da linguagem e o seu uso nos ambientes reais de comunicação (cf. Bybee, 2010; Cezario; Furtado da Cunha, 2013; Diessel, 2015). Sendo assim, esse modelo não se limita apenas aos aspectos formais da língua, sendo também consideradas em sua análise questões pragmáticas, semânticas e cognitivas. Dessa forma, essa abordagem combina os pressupostos da Linguística Funcional norte-americana e da Linguística Cognitiva.

Para o Funcionalismo norte-americano, existe uma conexão entre gramática e discurso. Nessa abordagem, os dois interagem e se influenciam mutuamente. Dessa forma, a gramática é compreendida como uma estrutura que muda e se adapta constantemente em função das transformações que ocorrem no discurso (cf. Givón, 1995). Sendo assim, a língua é concebida como um mecanismo usado para fins comunicativos, que, como tal, não pode ser analisado como um objeto independente, mas que funciona a partir de uma estrutura maleável. Portanto, de acordo com essa visão, não existe uma separação intransponível entre língua (gramática) e fala (discurso).

A Linguística Funcional Centrada no Uso se aproveita dos conhecimentos da Linguística Cognitiva e trata o conhecimento da linguagem como um reflexo das habilidades cognitivas de cada indivíduo, como a capacidade de categorizar, compreender e usar metáforas e metonímias, além de aspectos relacionados ao processamento linguístico e à experiência humana e, portanto, às atividades culturais pessoais e sociais. Nesse modelo, a gramática é entendida como uma rede constituída de construções que se organizam na mente do falante de maneira hierárquica, na qual integram-se propriedades fonológicas, morfossintáticas, semânticas e pragmáticas.

Essa rede é vista como um conjunto de nós (pareamentos de forma-função) que estão conectados entre si. Advindo da perspectiva cognitivista, esse conceito de Gramática de Construções (Goldberg, 1995; 2006) se trata de uma idealização de um arquimodelo genérico,

caracterizado pela hipótese de que o conhecimento linguístico tem a forma de uma rede de unidades simbólicas (Pinheiro, 2016).

Portanto, conforme a abordagem funcional-cognitiva, a cognição é materializada na interação e reflete o funcionamento de nossos pensamentos como indivíduos, concretizando formas únicas de expressão individual e refletindo o fato de estarmos inseridos em um ambiente social e cultural. Sendo assim, a partir da nossa experiência e interação, moldamos e somos moldados, o que faz com que soframos alterações desse ambiente sociocultural ao mesmo tempo que o modificamos.

Dessa forma, os eventos de uso orientam a formação e o funcionamento do sistema linguístico do falante, ao passo que o falante contribui para a manutenção e possíveis variações e mudanças do sistema linguístico pertencente a uma comunidade. Assim, o falante produz impacto (ao realizar reanálises, analogias e/ou quaisquer outros processos que implicam alterações e/ou extensões no emprego de expressões linguísticas) não apenas em seu próprio sistema linguístico, mas também nos sistemas de outros falantes.

Sendo assim, como esclarece Martelotta (2011), o sistema linguístico tem um caráter eminentemente dinâmico ou emergente, já que nasce da adaptação das habilidades cognitivas humanas a eventos de comunicação específicos e se desenvolve a partir da repetição ou ritualização desses eventos. Esses, por sua vez, acabam desempenhando um papel duplo no esquema comunicativo, pois, ao mesmo tempo em que constituem o resultado da atuação no sistema linguístico, fornecem também um input para os sistemas dos demais falantes, atuando também como um processo de feedback.

Hopper (1998) também utiliza o termo emergente para explicar como a estrutura linguística se remodela continuamente, fazendo com que seus usuários (re)criem padrões gramaticais a todo o momento, a depender da força informativa do discurso. Dessa forma, discurso e gramática se retroalimentam, visto que as construções linguísticas são o resultado da atuação do sistema linguístico ao mesmo tempo que o moldam, funcionando em um ciclo.

A partir disso, é possível lidar não apenas com a estabilidade de determinados padrões gramaticais, mas também com a instabilidade presente em um sistema pautado na pragmática e nos processos cognitivos que categorizam e também organizam o pensamento linguístico. Dessa forma, de acordo com essa teoria, o uso e a cognição constituem a base da capacidade humana para a linguagem.

Barlow e Kemmer (2000) propõem quais seriam as principais características que definem esta teoria da linguagem. A primeira delas é que existe uma relação íntima entre

estrutura linguística e instâncias de uso. A dinâmica intrínseca da linguagem decorre da constante atualização da rede de representações abstratas da gramática do falante por meio dos eventos de uso linguístico. Dessa maneira, as instâncias de uso mais específicas emergem a partir do uso da língua, para posteriormente dar origem a padrões mais gerais e abstratos.

Nesse processo, as instâncias de uso concreto e específico “alimentam” os padrões mais gerais, da mesma maneira que estes, por sua vez, licenciam essas concretizações. Em resumo, as instâncias operam de baixo para cima (*bottom-up*), mas eventualmente assumem também a função de instanciadoras, agindo de cima para baixo (*top-down*). Sendo assim, a partir do momento que uma construção é estabelecida, novas possibilidades são abertas, fazendo com que surjam possíveis espaços (*slots*), que podem ser preenchidos por unidades linguísticas semelhantes ou completamente inovadoras.

A exemplo, teria sido por meio do uso recorrente de conectivos como “visto que”, “ainda que” e “toda vez que” que o padrão construcional abstrato [(X) que S V O] se consolidou no sistema linguístico dos falantes de língua portuguesa, fazendo com que, eventualmente, esse padrão se tornasse produtivo para o surgimento de novas construções.

A segunda característica seria a importância da frequência. Ambos os autores chamam a atenção para a atuação desse fator na estruturação, na manutenção e no funcionamento do sistema linguístico. De um lado, os enunciados reais que são recorrentemente ativados na gramática e efetivados no discurso se tornam rotinas discursivas e essa repetição acaba moldando a língua e implicando em efeitos sobre suas unidades. Do outro, os enunciados também podem conter estruturas inovadoras, que surgem de intenções comunicativas de um falante individual ou de ressignificações de estruturas já existentes.

Desse modo, a rotinização de determinados padrões linguísticos é um importante fator que pode tanto conservar formas já existentes, facilitando seu acesso na mente do falante; torná-las obsoletas, mesmo que ainda continuem armazenadas na rede linguística do falante, elas seriam momentaneamente desativadas, devido a sua baixa ou inexistente ocorrência(s); e/ou surgir novas estruturas linguísticas.

A terceira diz respeito à compreensão e produção integradas ao sistema linguístico. Sob essa ótica, a maneira como o falante emprega o sistema linguístico também faz parte de seu entendimento acerca da própria língua. As diversas expressões presentes em um sistema linguístico são pragmaticamente marcadas, não sendo usadas de forma aleatória ou acidental, o que evidencia a sistematização e o conhecimento intrínseco, validando a premissa de que o sistema é influenciado pela experiência. Assim, à medida que os eventos de uso impactam a

estrutura e o funcionamento do sistema linguístico do falante, essa estrutura também está intrinsecamente relacionada com os processos mentais que ocorrem durante o uso da linguagem (Bybee, 2010).

A quarta característica está relacionada ao foco no papel da aprendizagem e da experiência na aquisição da linguagem. A estruturação do sistema do falante ocorre mediante a produção e compreensão da linguagem que surge na interação entre os participantes de uma cena comunicativa, juntamente com todos os elementos associados a essa interação. Dessa maneira, o sistema linguístico e a sua aquisição estão em constantes processos de formação, sendo influenciados pelas experiências que o falante vivencia em seus eventos de uso.

Outro ponto evidenciado pelos autores é que as representações linguísticas são emergentes. Essa característica destaca que a língua está sujeita a mudanças, não sendo composta por estruturas fixas e categorias fechadas. A linguagem humana é fundamentada em padrões linguísticos que podem se repetir ou não em um discurso.

Nesse contexto, um padrão específico pode ser potencialmente fortalecido à medida que os falantes o utilizam ou então enfraquecido (ou desativado) na rede linguística, caso haja uma baixa ou inexistente frequência em seu uso. Dessa forma, ocorre uma contínua renovação e reorganização dentro do sistema linguístico a partir das novas experiências vividas pelo seu usuário.

Os autores também destacam a importância de dados reais na descrição e construção da teoria. O modelo baseado no uso fundamenta-se na análise de contextos autênticos de produção linguística, partindo do princípio de que a linguagem se desenvolve na interação, reafirmando assim a simbiose existente entre discurso e gramática. Em virtude disso, a análise e descrição de dados são realizadas por meio da coleta e observação de ocorrências reais, que, por sua vez, são usadas para embasar a descrição dessa teoria.

Outra característica seria uma relação íntima entre uso, variação sincrônica e mudança diacrônica. O uso seria um dos colaboradores para a ocorrência desses fenômenos, já que, na interpretação de um discurso, inevitavelmente ocorrem adaptações mínimas que afetam tanto o sistema interno do próprio falante, quanto o de seus interlocutores, o que pode resultar nos processos de mudança e variação linguística.

Enquanto no processo de mudança uma estrutura ou padrão linguístico passa a ter um número de ocorrências maior até que outro, eventualmente, deixe de ser usado em determinados contextos linguísticos (havendo, então, uma inativação no sistema linguístico do falante), no de variação linguística, dois ou mais padrões linguísticos coexistem (e competem) em contextos

discursivos semelhantes, tendo sua ativação influenciada por fatores pragmáticos e também internos ao sistema linguístico, que fazem o falante optar por uma variante em detrimento de outra. Dessa forma, ambos os processos refletem o caráter indiscutivelmente dinâmico presente nas línguas.

Outro aspecto considerado pelos autores é a interconectividade entre o sistema linguístico e os sistemas cognitivos não-linguísticos. De acordo com essa perspectiva, diferentes áreas da cognição interagem para que seja possível compreender e produzir enunciados linguísticos. Sendo assim, diversas habilidades cognitivas humanas, como a de categorização, atenção, memorização e criatividade, se unem para que a capacidade linguística se desenvolva. Os conhecimentos linguísticos, adquiridos através da experiência e do uso, envolvem diferentes níveis de complexidade, que são interpretados por este sistema cognitivo igualmente complexo. Dessa maneira, de acordo com esta teoria, a mente funciona holisticamente, interconectando todas essas diferentes esferas da capacidade humana.

Além disso, os autores também destacam o papel crucial do contexto no funcionamento do sistema linguístico. Como destacado anteriormente, o uso molda a forma como a língua se estrutura. Desta maneira, cabe dizer que toda construção gramatical é dependente do seu contexto de uso.

Assim, é defendido que o sentido não é definido pela forma, ainda que esta forneça pistas a respeito de seu significado dentro de um contexto específico, mas não o determina por si só. Sendo assim, essa abordagem lida com o que se chama de “riqueza de significado”, pois permite que o falante faça inferências e use sua criatividade no ato comunicativo.

Para mais, outro pressuposto presente neste modelo teórico e que rege nossa pesquisa é o Princípio da Não-Sinonímia. Postulado por Goldberg (1995), esse princípio instancia que não existem formas diferentes que tenham significados iguais, portanto, se existem mudanças na forma, a intenção comunicativa é diferente.

Dessa forma, uma construção iniciada pelo conectivo “visto que”, por exemplo, atingiria propósitos comunicativos específicos, não alcançados da mesma maneira pelas construções iniciadas por “posto que” e “dado que”. Portanto, acreditamos que mesmo que essas construções sejam semelhantes, tanto em sua forma quanto pragmaticamente, elas devem apresentar especificidades que as tornam únicas. Nessa perspectiva, como as construções são específicas em cada língua, acreditamos que haja também diferenças entre os usos dessas construções em português e em italiano.

1.2 Processos cognitivos de domínio geral

Ainda de acordo com essa abordagem teórica, Bybee (2010) reitera que a linguagem é uma faculdade que exhibe estrutura aparente e regularidade de padrões e mostra, ao mesmo tempo, uma considerável variação em todos os níveis. Sendo assim, apesar de estar em constante mudança, a língua possui certa regularidade e o sistema linguístico do falante é estruturado a partir de processos cognitivos de domínio geral.

O primeiro deles é a categorização, que se define como a capacidade de associar um elemento a um determinado conjunto por semelhança. É considerado parte de um domínio geral pois vários tipos de categorias, não apenas linguísticas, são criadas a partir da experiência. Enquanto isso, o segundo processo é chamado de *chunking*, que trata de sequências linguísticas reiteradamente produzidas que são armazenadas e acessadas na cognição como uma única unidade complexa.

Já o terceiro é o processo de analogia, pelo qual novos enunciados são criados baseados em enunciados já armazenados no sistema linguístico. Devido a especificidade das construções linguísticas e a forma como elas são formadas a partir do uso, a probabilidade e a aceitação de um novo item são gradientes e, por isso, se baseiam na extensão de similaridade com usos que já existem.

O processo seguinte é chamado de memória rica e está relacionado ao armazenamento de detalhes da experiência linguística, como detalhes fonéticos, os contextos de uso de determinadas estruturas, significados e inferências associados a enunciados. E, por fim, o último mecanismo apontado é a capacidade de realizar associações transmodais que possibilitam o *link* entre a forma e o seu significado.

Esses processos cognitivos explicam o funcionamento das línguas, a aquisição da linguagem e a mudança linguística também. Aqui não abordaremos questões relativas à aquisição e à mudança, pois não são escopos do nosso trabalho.

1.3 O conceito de construção e as suas características

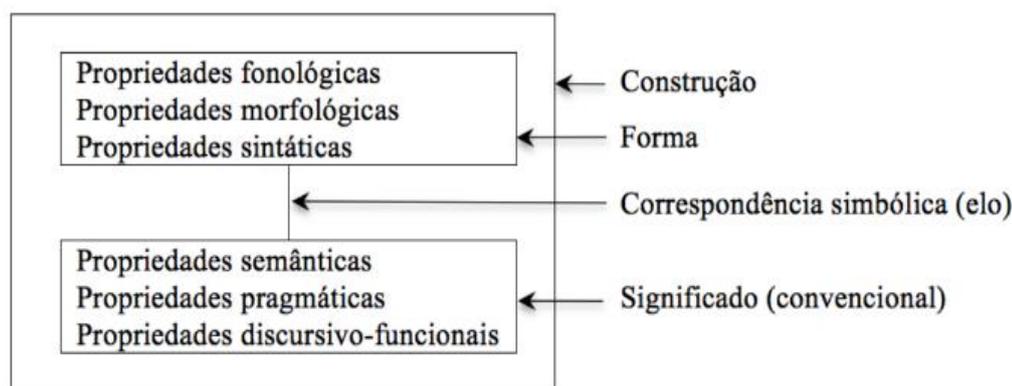
Outro conceito importante que emerge dessa abordagem baseada no uso é o de construção. De acordo com Traugott (1995, 2006), uma construção gramatical pode ser definida como o pareamento convencionalizado de forma e sentido, visto como um esquema simbólico a partir do qual se instanciam os componentes da gramática. A ideia crucial por trás da construção é que ela é um pareamento direto entre forma e significado que tem estrutura sequencial e pode incluir posições que são tanto fixas quanto abertas (cf. FILLMORE et al., 1988; GOLDBERG, 1995; CROFT, 2001).

Nos trabalhos de Goldberg (2006), o conceito de construção compreende desde um morfema simples (como o -s de plural) até um esquema mais complexo como os esquemas abstratos estudados nesta pesquisa, [(V_{pp}) que (S) V (O)] e [(V_{pp}) che (S) V (O)].

Este referencial teórico foi adotado, pois consideramos a língua como um grande arcabouço de construções que são definidas por um pareamento de forma e sentido, desde as mais simples até as mais complexas. Com base nos estudos de Croft (2001), a construção se define justamente pela junção das propriedades da forma (sintáticas, morfológicas e fonológicas) e da função ou significado (semânticas, pragmáticas e discursivo-pragmáticas).

Essa relação pode ser observada na imagem abaixo:

Figura 2: Estrutura simbólica de uma construção



Fonte: Croft (2001)

No que diz respeito à organização dessas construções, Croft (2001) destaca que as construções formam um elenco estruturado do conhecimento que o falante tem das convenções de sua língua que, normalmente, é representado por uma rede taxonômica de construções. Partindo desse princípio, qualquer construção com propriedades morfológicas, sintáticas, lexicais, semânticas, pragmáticas ou discursivo-funcionais únicas (idiossincráticas) deve ser representada como um nó independente na rede construcional, tendo o objetivo de capturar o conhecimento que o falante tem da sua língua.

Com base na perspectiva em que a língua funciona como rede de *links* entre as construções, as mudanças estão interconectadas e as redes podem se reorganizar e aumentar, tornando-se mais produtivas, ou então podem se contrair, fazendo com que essas construções se diluam gradualmente do sistema do usuário. (cf. Arena, 2015)

No caso das construções abordadas nesta pesquisa, não haveria um processo de mudança nem mesmo de apagamento de uma construção em detrimento da outra, mas sim uma situação

de coexistência dessas construções, dentro de seus próprios sistemas, que competem entre si para serem escolhidas pelo falante.

Hilpert (2021) postula que a competição na rede de construções pode ser entendida em termos de dois (ou mais) nós que estão conectados na rede. A competição surge se um falante deseja expressar um significado particular e esse significado está conectado a duas formas diferentes. A forma que o falante efetivamente escolhe depende da força do vínculo simbólico entre o significado e uma das respectivas formas.

A construção que possui uma conexão mais forte ganha (sendo esta definição bastante suscetível ao contexto comunicativo no qual o usuário está inserido no momento) enquanto a que tem uma conexão mais fraca perde. A conexão escolhida pode se tornar ainda mais aparente (e acessível) na rede linguística, ao mesmo tempo que a construção ignorada tenderia a se tornar ainda mais fraca do que antes.

Essa ideia de força de uma determinada construção está ligada a como o uso pode impactar as representações cognitivas do usuário da língua, sendo assim, pode estar relacionada a dois principais conceitos: o primeiro deles, já discutido alguns parágrafos acima, é a frequência de uso, pois acredita que, quanto mais um item é usado, mais facilmente será acessado em futuras interações comunicativas semelhantes.

Em segundo, se correlacionaria com o Modelo de Exemplos (Bybee, 2006; 2010; 2013; Pierrehumbert, 2001), que postula que o falante armazena todas as ocorrências de uma dada construção em termos de uma nuvem de exemplos, de forma que elementos mais fortemente relacionados – em termos de forma e/ou conteúdo – se encontrariam mais próximos na rede cognitiva. Da mesma maneira, elementos menos relacionados estariam mais distantes na rede. Sendo assim, pressupõe-se que o falante acessa o feixe de nós (pareamentos de forma-função) que estão correlacionados e escolhe aquele que é mais frequente e/ou adequado para seu propósito comunicativo.

Diessel (2019) acrescenta que quanto mais frequentemente um determinado link, ou um determinado padrão de links, é processado, mais forte é o(s) peso(s) das conexões e maior é a probabilidade de que essas conexões sejam reutilizadas no futuro. Isso explicaria também a maneira pela qual a estrutura de uma rede pode se modificar. Além disso, os modelos de rede podem mudar por meio de mecanismos específicos, que serão melhor evidenciados ao longo do texto, que criam ou excluem novos nós e novas conexões ou que servem para reconfigurar um conjunto já existente de nós e conexões.

Ademais, o autor destaca uma divisão em como categorizar as relações desses nós que existem na rede linguística do usuário. A primeira forma seria a relação simbólica, que conecta uma forma a um significado, enquanto a segunda é denominada como relação sequencial, que conecta os elementos linguísticos em sequências e, por fim, existiria também a relação taxonômica que conecta padrões linguísticos em diferentes níveis de abstração (como apresentado na Figura 1).

O modelo de Gramática de Construções adotou a ideia de que construções podem estar em competição se seus significados e funções são suficientemente similares. Construções gramaticais novas tendem a emergir em domínios que já são bem representados por outras construções. De acordo com Nascimento (2022), é isso que acontece com as construções [(X) que SVO] introduzidas por “visto que”, “dado que” e “posto que” no português do Brasil, já que estas possuem, na maioria das vezes, instâncias de uso praticamente idênticas, mas que, ainda assim, constituem formas diferentes que coexistem no sistema linguístico e que podem ser usadas para exprimir basicamente a mesma circunstância.

Sendo assim, Hilpert (2021) se questiona porque os usuários necessitariam de uma nova construção para exprimir determinado significado se já existe uma construção presente e bastante acessível em sua rede linguística.

De acordo com Croft (2000), quando ocorre uma competição entre construções, acontece a chamada “propagação”. Considerando que uma inovação linguística tenha sido criada, a propagação se preocupa em detalhar como ela se espalha por uma comunidade de falantes e não com aspectos relativos à forma ou semântica necessariamente. Dessa forma, a propagação não é funcionalmente motivada, pois não tem nada a ver com o quão bem uma forma é funcionalmente adaptada ou quão útil é o seu significado. Em vez disso, o autor argumenta que é socialmente motivado.

1.4 Esquematicidade, composicionalidade e produtividade

É importante também destacar que os três princípios mais importantes para a análise de construções, tanto do ponto de vista diacrônico quanto sincrônico, são a esquematicidade, a composicionalidade e a produtividade (Goldberg, 2002; Traugott; Trousdale, 2021).

A esquematicidade reflete o grau de abstração e generalização de uma construção e se fundamenta, especialmente, na capacidade de categorizar. Os esquemas podem ser compreendidos como padrões de experiência que ocorrem rotineiramente e que são cognitivamente fixados. Nessa perspectiva, esquemas linguísticos são grupos abstratos e gerais que podem ser analisados em razão das posições abertas ou fechadas que os constituem. Um

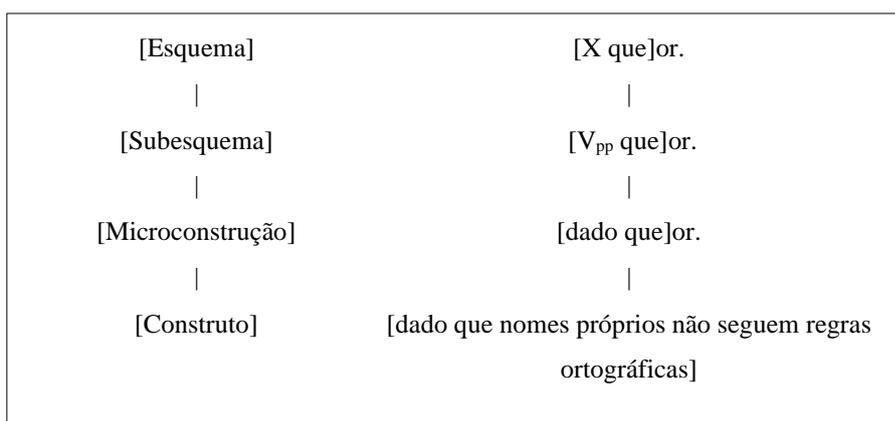
esquema altamente abstrato é pouco específico, apresentando *slots* abertos que, ao serem preenchidos, se tornam subesquemas e/ou microconstruções, que possuem um maior grau de especificação.

No caso das construções estudadas aqui, temos um maior nível de abstração codificado pelos esquemas [(X) que (S) V (O)] e [(X) che (S) V (O)], assim como temos também um nível menos abstrato, em que mais *slots* estão preenchidos, e que torna a construção mais específica, como é o caso dos esquemas investigados nesta pesquisa: [(dado) que S V (O)], [(posto) que S V (O)], [(dato) che S V (O)] e [(posto) che S V (O)].

Existem, dessa forma, diferentes níveis de esquematicidade (Traugott e Trousdale, 2021), em que, normalmente, há um esquema que representa o nível mais geral e abstrato, que vai se tornando cada vez menos esquemático conforme se aproxima dos níveis inferiores, sendo eles o subesquema e a microconstrução, respectivamente. Além desses, existem também os construtos, que representam as instâncias reais de uma construção, ou seja, é aquilo que realmente é produzido pelos usuários de um sistema linguístico.

Além disso, como evidenciado antes, o processo de generalização se dá numa ordem top-down, assumindo que a rede estoca os detalhes relevantes a partir da experiência dos usuários da língua, organizados em termos de princípios de categorização (Langacker, 2008), como pode ser observado na figura a seguir:

Figura 3: Níveis de esquematicidade



Fonte: Adaptado de Oliveira e Clemente (2022)

Outro princípio considerado é a composicionalidade, que mede o grau de transparência do elo entre forma e significado de uma construção. Ou seja, a análise dessa propriedade verifica se o significado de uma construção é, de algum modo, resultado da correlação de suas partes componenciais. Em outras palavras, a composicionalidade pode

ser pensada em termos de transparência entre o significado da parte e o significado do todo. Oliveira e Clemente (2022) citam como exemplos os diferentes usos de [a gente] no português brasileiro.

Em uma frase como “Normalmente, a gente daquele país tem cabelos claros”, o significado do determinante “a” e do substantivo “gente” contribuem para a construção do significado do sintagma nominal determinado singular e feminino, sendo entendido como “as pessoas”. Diferentemente, em “A gente viu o filme”, esses traços se perdem e há um significado altamente gramatical, que faz referência a primeira pessoa do plural. Esse significado não é derivado do significado das partes dessa construção sendo, portanto, menos composicional se comparado ao primeiro caso.

Tal princípio está intimamente ligado a um dos processos cognitivos de domínio geral: o *chunking*. Para que esse processo ocorra, é necessário que haja uma perda de composicionalidade, fazendo com que dois ou mais elementos linguísticos sejam lidos como um único bloco dotado de sentido.

É o que ocorreu com o esquema investigado nesta pesquisa, já que o *slot* [V_{pp}] não é mais percebido pelo falante como um espaço disponível para se alocar verbos participiais passados, em virtude de essas construções serem entendidas como um único bloco. Dessa forma, os verbos participiais perderam sua composicionalidade devido à alta frequência com que apareciam com o conectivo “que”. A partir disso, os construtos advindos desse subesquema são lidos como conjunções, dotadas de significado próprio.

Outro ponto a ser considerado é a produtividade, que está relacionada à extensibilidade e restrição de uma construção. Na análise da produtividade, avalia-se o grau em que um esquema sanciona construções menos esquemáticas. Um exemplo seria o esquema [Adj + -mente]ADV no português. Sendo extremamente produtivo, esse esquema sanciona diversos construtos como felizmente [feliz + -mente], agilmente [ágil + -mente] e tranquilamente [tranquilo + -mente]. Bem diferente do esquema [X que]CONNECT, que, por sua vez, possui um grau de produtividade consideravelmente mais baixo, já que não se tem novas conjunções se formando com tanta frequência no sistema linguístico.

Além disso, a produtividade também está ligada às restrições de um (sub)esquema. Na rede construcional dos conectores, no subesquema [V_{pp} que] são selecionadas para o preenchimento de seu *slot* apenas formas verbais de base participial, ocasionando as microconstruções [dado que], [posto que] e [visto que], por exemplo. Em contrapartida, não são aceitas microconstruções como [vai que] e [supondo que].

Dessa forma, a produtividade também pode ser medida, metodologicamente, através de fatores, sejam estes estruturais ou pragmáticos, como será observado nas próximas seções.

2. REVISÃO DA LITERATURA

Nesta seção, revisitaremos alguns trabalhos importantes sobre a questão da causalidade, o encaixamento de orações na perspectiva funcionalista e também traremos um breve resumo de alguns trabalhos predecessores que se mostraram relevantes para o destrinchamento da análise de conectivos de base participial e/ou do esquema abstrato [X que]conectivo.

2.1 Encaixamento de orações

Nas gramáticas tradicionais (Bechara, 2009; Cunha; Cintra, 2008; Rocha Lima, 1985), comumente as cláusulas existentes na língua são divididas em dois grupos: as de período simples, constituído por apenas uma oração, ou composto, quando existe a união de duas ou mais orações, que podem estabelecer entre si relações de dependência. No segundo caso, quando as orações são sintaticamente independentes entre si, ou seja, uma oração não exerce função sintática em relação à outra, o período ao qual pertencem é denominado composto por coordenação. Por outro lado, quando o período possui orações sintaticamente dependentes entre si, isto é, em que uma oração exerce uma função sintática em relação à outra, o período é denominado composto por subordinação.

Entretanto, segundo a visão funcionalista, as orações não poderiam ser divididas de forma estanque e dicotômica a partir apenas dos processos de coordenação e subordinação e, por isso, é defendido por este modelo teórico a análise dessas cláusulas por meio de um *continuum*. Dessa forma, ao estabelecerem graus de integração entre as orações na perspectiva da gramaticalização, Hopper e Traugott (1993) concebem que há três processos de combinação de orações: a parataxe, a hipotaxe e a subordinação.

Nessa nova perspectiva, a parataxe possui uma relação de independência relativa entre os núcleos que compõem o complexo oracional; sendo assim, o vínculo entre as orações depende apenas do sentido e da relevância da relação entre elas. Em outras palavras, a parataxe abrange tanto a justaposição quanto a coordenação.

Enquanto isso, a hipotaxe compreende dependência entre um núcleo e margens, mas não encaixamento da margem em um constituinte do núcleo, isto é, abarca orações em que há um núcleo e uma ou mais margens apresentam dependência relativa, não

estando totalmente incluídas em nenhum constituinte da oração núcleo, completando, assim, a estrutura argumental do verbo.

Nesse caso, a hipotaxe abrange aquelas orações que envolvem algum tipo de relação circunstancial, como condição, razão, propósito, tempo, espaço, maneira, meio, como as adverbiais e as adjetivas explicativas, que é o caso das orações investigadas nesta pesquisa.

Já a subordinação envolve dependência completa entre núcleo e margem(ns) e, portanto, encaixamento de toda a margem em um constituinte do núcleo (RODRIGUES, 2018). As orações denominadas subordinadas ou encaixadas, nesse caso, são aquelas que são partes constituintes de outras, tais como as substantivas e as adjetivas restritivas.

Podemos observar as principais diferenças entre esses processos na tabela a seguir:

Quadro 1: Processos de combinação de orações

parataxe	– dependência semântica	– encaixamento
hipotaxe	+ dependência semântica	– encaixamento
subordinação	+ dependência semântica	+ encaixamento

Fonte: Hopper e Traugott (1993)

Diante dessa descrição, depreende-se que a chamada “subordinação adverbial”, na verdade, constitui, para esses autores, um caso de hipotaxe circunstancial ou de realce, em razão de, nesse caso, não haver encaixamento de uma cláusula em outra, como na subordinação substantiva e adjetiva restritiva, mas uma combinação hipotática de realce, em que uma cláusula amplia outra circunstancialmente.

Sendo assim, ao contrário do que ocorre no encaixamento, em que uma oração está integrada estruturalmente e sintaticamente em outra, a articulação entre oração hipotática e sua principal não está sujeita à integração sintática e se relaciona com o aspecto organizacional do discurso. Portanto, as orações hipotáticas investigadas neste trabalho podem modificar ou expandir, de alguma forma, a informação contida na outra oração, estabelecendo, dessa forma, uma relação circunstancial.

2.2 Grau de integração entre as orações

Adicionalmente, também foi levado em consideração para esta pesquisa o Princípio da Integração. Com base nos estudos de Givón (1979), postula-se que uma maior integração semântica ou pragmática implica também em uma maior integração sintática entre as orações. Assim, ao avaliar o grau de integração entre as orações, o objetivo é verificar se as estratégias sintáticas manifestam diferentes níveis de gramaticalização. Estes podem ser inferidos por meio de fatores como a simultaneidade modo-temporal dos eventos nas orações, a correferencialidade dos sujeitos e a codificação formal do sujeito na oração hipotática.

Acredita-se que a realização do sujeito na oração hipotática por meio de anáfora pronominal ou por zero resulta em uma maior integração entre ela e a oração principal. Da mesma forma, se houver correferencialidade entre os sujeitos e se o tempo dos verbos nas orações hipotáticas coincidir com os das orações principais, o grau de integração entre oração matriz e oração hipotática será maior.

2.3 Ordenação, valor semântico e status informacional

Ademais, os estudos realizados por Diessel (2013) também se mostraram relevantes para este estudo, uma vez que o autor observa a variação estrutural das orações adverbiais levando em conta suas características semânticas, sintáticas e pragmáticas, considerando que as características morfossintáticas das orações adverbiais variam de acordo com sua posição e função, e que diferentes tipos semânticos de orações adverbiais podem ter diferentes propriedades estruturais.

Um elemento distintivo entre orações adverbiais e orações não subordinadas é sua disposição linear (conforme Haspelmath, 1995), que será referida aqui como “ordem”. Enquanto as orações coordenadas e as sentenças paratáticas frequentemente estão ligadas à sentença anterior, as orações adverbiais (ou hipotáticas) podem surgir tanto antes quanto depois da oração principal. Portanto, o arranjo linear que as orações principais e adverbiais assumem está estreitamente ligada à sua função pragmática.

Em seu uso canônico, as orações adverbiais antepostas têm a função de apresentar informações que são pragmaticamente pressupostas, fornecendo uma base temática para novas informações afirmadas em orações subsequentes. No entanto, se a oração adverbial seguir a oração principal, ela só poderá ser adicionada à estrutura anterior após a conclusão da oração principal. Sendo assim, frases complexas com orações adverbiais pospostas podem ser planejadas e processadas sucessivamente, isto é, uma oração de cada

vez, sugerindo que as orações adverbiais pospostas são potencialmente mais independentes da oração principal do que quando as orações adverbiais a precedem.

Sendo assim, outro aspecto a ser considerado é a estrutura informacional, que diz respeito ao conteúdo que o locutor e interlocutor compartilham (ou supõem compartilhar) no momento da interação verbal (Furtado da Cunha; Costa; Cezario, 2015). Esse conceito se relaciona ao modo como o falante organiza o próprio texto a partir dos seus objetivos comunicativos e também levando em conta as necessidades de seu interlocutor (Martelotta, 2003). Dessa forma, o discurso é moldado com base nas pressuposições que o locutor faz a respeito do conhecimento de seu interlocutor, supondo aquilo que ele previamente já conhece ou sabe e aquilo que será apresentado como uma novidade.

Com base em Prince (1981, 1992), os referentes podem ser classificados como: novos (subdivididos em novos-em-folha ou novos disponíveis), evocados (subdivididos em evocados textualmente ou situacionalmente) e inferíveis. Uma entidade é classificada como nova se for introduzida no discurso pela primeira vez, enquanto o conceito de evocado nos remete ao elemento situacionalmente acessível ou que já aparecera anteriormente no discurso, sendo, portanto, uma informação velha para o ouvinte ou leitor, que é tida como pressuposta. Já um referente inferível trata-se de um que não fora mencionado no discurso, mas é identificado pelo interlocutor por meio de um processo de inferência a partir de outras informações dadas.

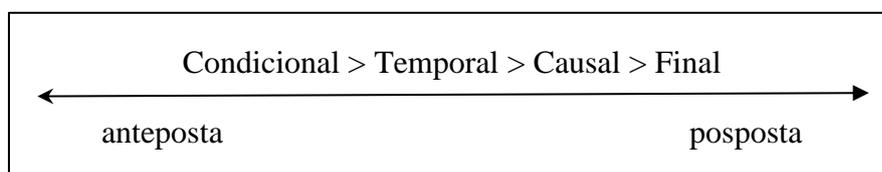
Assim sendo, a informatividade está intimamente ligada ao que chamaremos de pressuposição, composta por uma parte externalizada e explícita, cujo conteúdo semântico é aquilo que efetivamente se declara, e por uma outra parte interiorizada e implícita, cujo conteúdo semântico não é objeto direto de declaração e de comprometimento do falante, mas algo dado como uma espécie de conhecimento dado, partilhado pelos participantes do ato comunicativo.

Atrelado a esse conceito, de modo geral, acredita-se que orações hipotáticas que apresentam informações novas tendem a ocorrer após as orações matrizes, instanciando orações focais, enquanto aquelas que são anteriores às orações matrizes, normalmente, contêm uma informação já fornecida ou conhecida pelo interlocutor.

Diessel (2013) ainda une a posição linear assumida pelas orações hipotáticas a outro importante fator: o valor semântico. As orações adverbiais são frequentemente classificadas em diversos subtipos semânticos, os quais expressam relações temporais, condicionais, causais, finais e outras. De acordo com o autor, a relação semântica entre orações principais e adverbiais está correlacionada com sua ordem (cf. Diessel, 2001).

Sendo assim, de acordo com sua perspectiva, as tendências posicionais das orações condicionais, temporais, causais e intencionais podem ser delineadas da seguinte maneira:

Figura 4: Relação entre valor semântico e ordenação



Fonte: Adaptado de Diessel (2013)

Como pode ser percebido a partir do esquema supramencionado, nesta visão as orações condicionais tendem a preceder a oração principal enquanto as orações causais geralmente são colocadas após sua oração principal. Entretanto, os estudos realizados pelo autor não fornecem informações sobre as tendências das orações concessivas.

Devido a esse fato, recorreremos novamente ao estudo realizado por Neves (1999), que postula que a ordenação das orações concessivas é ambígua, pois depende do propósito comunicativo do interlocutor. Geralmente, aparecem antes de suas orações principais quando possuem a função de tópico, retomando, dessa maneira, informações que já haviam sido previamente fornecidas. Já se aparecem na posição posterior a oração principal à qual se liga, a oração concessiva assume a função comunicativa de adendo. Dessa forma, a posição da oração concessiva está muito mais relacionada com o grau de novidade que traz do que necessariamente com seu valor semântico.

Os estudos de Neves (1999, 2011) mostram-se ainda complementares aos de Diessel ao fomentar que as orações causais possuem essa tendência a posposição pois, de acordo com a autora, normalmente se enuncia primeiro o efeito, expressado pela oração principal, para depois ser enunciada a causa, expressada pela oração hipotática. Por sua vez, no caso das orações condicionais, primeiro se enuncia a ocorrência de uma condição, que pode ou não ser satisfeita e que é evidenciada pela oração hipotática, para depois ser enunciada o fator dependente para a concretização desta condição, expressa pela oração principal, evidenciando, dessa forma, a tendência à anteposição deste tipo de oração hipotática.

2.4 Os conectivos “dado que”, “posto que”, “dato che” e “posto che”

Partindo da ideia de que algumas conjunções poderiam apresentar valores semânticos não previstos pela tradição, foi realizada uma revisão da abordagem das

gramáticas em língua portuguesa e em língua italiana, buscando observar como elas apresentam as orações subordinadas adverbiais e as conjunções que as introduzem e que foram destacadas nesta pesquisa.

As gramáticas de Bechara (2009), Rocha Lima (1985) e de Cunha e Cintra (2008), utilizadas para esta revisão, apresentam visões que se assemelham, já que os três autores apontam que os conectivos em estudo são introdutores de orações adverbiais, classificando [posto que] como uma conjunção subordinativa concessiva e [dado que] como uma conjunção subordinativa condicional.

Além desses trabalhos, também foi consultada a gramática descritiva de Neves (1999), na qual são apresentadas as mesmas classificações supracitadas em relação à forma [posto que]; contudo, o conector [dado que] é classificado pela autora como uma conjunção adverbial causal, diferindo-se, portanto, das demais gramáticas consultadas.

Da mesma forma, no que se refere à língua italiana, foram consultadas algumas gramáticas italianas, buscando mapear quais sentidos normalmente são atribuídos a essas microconstruções. Entretanto, o cenário encontrado foi um pouco diferente. Em Dardano e Trifone (1999) e (2011), foi postulado que a conjunção subordinativa “dato che” tem um valor causal, enquanto “posto che” é entendida como uma conjunção subordinativa condicional. Enquanto isso, na gramática de Serianni e Castelvechi (1997), ambas as locuções conjuntivas são classificadas como causais. A Gramática Prática Michaelis, por sua vez, enumerou duas possibilidades de significado para [dato che]: condicional ou causal, ao mesmo tempo que classificou [posto che] com o valor condicional.

Entretanto, a maior diferença entre as gramáticas de língua portuguesa e as de língua italiana, foi a ausência de [posto che] em seus quadros classificativos. Consultando também a Grande Grammatica Italiana di Consultazione (2001), La Grammatica Italiana Treccani (2012) e Le regole e le scelte: Manuale di Linguistica e di Grammatica Italiana (2019), todas classificaram [dato che] como um conectivo causal, porém, não foi encontrada nenhuma classificação do conectivo [posto che], o que difere fortemente do cenário em português, no qual, em todas as gramáticas consultadas, a presença de [posto que] foi observada.

Isso nos permite supor que o nível de produtividade e frequência desse conectivo nas duas línguas pode ser diferente, partindo do critério que muitos manuais e gramáticas linguísticos, ao menos os que foram consultados para essa pesquisa, não se dedicam a classificação e descrição do conectivo [posto che] em língua italiana.

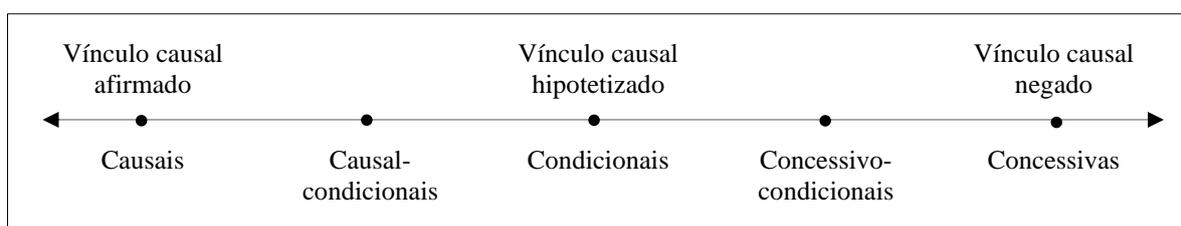
2.5 O valor semântico: o *continuum* das orações causais, condicionais e concessivas

A partir disso, torna-se relevante destacar a relação existente entre causa, condição e concessão. De acordo com Neves (1999), tanto as construções concessivas quanto as causais e condicionais expressam, de alguma forma, uma conexão “causal” em sentido mais amplo, assim como apresentam uma conexão condicional, devido a serem explicáveis em dependência de satisfação de necessidade (ou de suficiência) de determinadas condições.

Podemos dizer que, semanticamente, as relações concessivas, causais e condicionais se situam da seguinte forma: em um polo, caracterizado pela causalidade, há a relação de causa entre a hipotética e a matriz sendo afirmada; em um espaço intermediário, tido como das condicionais, há a relação de causa entre as duas orações sem ser afirmada ou negada; no outro polo, das concessivas, há o vínculo causal entre as orações.

Essas relações podem ser observadas na seguinte figura:

Figura 5: O *continuum* entre as orações concessivas, condicionais e causais



Fonte: Adaptado de Neves (1999)

No trabalho realizado por Nascimento e Castanheira (2020) e Nascimento (2022), foi verificado que as orações introduzidas por “dado que” e “posto que” apresentam, de forma geral, um maior número de ocorrências com o valor causal. É importante destacar que os resultados obtidos nessa pesquisa divergiram com aquilo que foi proposto pelas gramáticas tradicionais da língua portuguesa, isso porque as orações introduzidas por “posto que” e “dado que” registraram a ocorrência de multifuncionalidade, já que foram encontradas ocorrências causais, causal-condicionais e concessivas com a primeira conjunção e os valores causal e causal-condicional com a segunda.

Sob essa perspectiva, os valores semânticos não seriam tão rígidos quanto a tradição gramatical sugere. Questões internas à língua, bem como os contextos em que ela é empregada, estão intimamente ligados ao significado que as expressões apresentam.

Isso ilustra a natureza mutável das línguas humanas e sua propensão à variação, evidenciando uma abordagem mais dinâmica das questões gramaticais. Nesse sentido, os conectores existentes em um sistema linguístico não são categorizados de maneira inflexível, mas sim sujeitos a modificações no valor semântico de uma construção linguística específica, dependendo de fatores estruturais e discursivos.

Outro trabalho que se mostra relevante para esta pesquisa é a dissertação defendida por Batista (2013), cujo estudo permitiu que tivéssemos um panorama do que poderíamos encontrar com os conectivos escolhidos para essa pesquisa. Também se concentrando nas orações introduzidas por “dado que” e “posto que” na língua portuguesa, a autora optou por um recorte em que foram analisadas apenas as orações causais que apareciam com esses conectivos. A partir de seus resultados, foi percebido que essas orações apareciam, preferencialmente, postostas à oração principal; o tempo e modo verbais escolhidos, em sua maioria, era o presente do indicativo; e a maioria dos sujeitos eram não correferenciais, o que demonstra uma menor integração sintática entre as orações hipotáticas e suas matrizes.

Além disso, apesar de se tratar de um trabalho de natureza sincrônica, cabe também destacar o trabalho realizado por Oliveira e Clemente (2022), que investigou como o subesquema [V_{não-finito} que]_{Condicional} se estabilizou na rede de conectivos da língua portuguesa.

As autoras postulam que, a partir da emergência do conectivo [supondo que], outras três microconstruções emergiram na língua, sendo estas [considerando que], [dado que] e [posto que]. Através de processos de domínio cognitivo, supracitados nesta mesma seção, houve uma consolidação de [supondo que] na lista de conectivos condicionais disponíveis na língua, enquanto [posto que] e [dado que] sofreram um intenso processo de heterossemia, evidenciando um valor contrastivo, historicamente ligado à base causal presente na condicionalidade. Isso causou, ao longo do tempo, um deslocamento semântico dessas construções.

Diacronicamente, tanto “dado que” quanto “posto que” surgiram como conectores causais. Com o tempo, e devido a mudanças na construção da língua, esses conectores desenvolveram diversos significados que historicamente se relacionam com a base da causalidade. Entre esses significados, destacam-se o valor condicional, registrado nos séculos XIX e XX, e o valor concessivo, que surgiu no século XX (cf. Oliveira; Clemente, 2022).

De acordo com Amorim (2021), [dado que]_{causal} e [posto que]_{causal} teriam tido sua origem no ablativo absoluto do latim, tendo preservado dele traços que contribuem para sua atuação como exemplares mais próximos da zona da prototípica da causalidade. Assim sendo, são formas mais consolidadas e de fácil acesso para os falantes da língua, assim como o outro conectivo de base participial [visto que]. Além disso, como é amplamente aceito nos estudos sobre conectores adverbiais, o significado de causa é o mais fundamental e serve como base para o surgimento de significados mais abstratos, como condicionalidade e concessividade, que emergem de convencionalizações de projeções metafóricas da base causal, como explicitado a alguns parágrafos acima.

Dessa forma, neste capítulo foram apresentados os pressupostos teóricos que embasam nossa pesquisa acerca desses conectivos. Também foram trazidos alguns dos principais trabalhos da linha sobre conexão de orações e sobre os conectivos aqui investigados. Isso possibilita uma visão panorâmica e ampla da fundamentação funcionalista usada, bem como de algumas investigações já realizadas e que dialogam cientificamente com a nossa pesquisa.

2 METODOLOGIA

Neste capítulo, será apresentada a metodologia utilizada em nossa análise, assim como as etapas e fatores escolhidos para manipulação de nossos dados. Na seção 3.1, serão apresentados os *corpora* de onde nossos dados foram retirados e suas justificativas, enquanto na seção 3.2 serão expostos as etapas e os fatores escolhidos para nossa análise, baseados em nossas hipóteses e objetivos explicitados em nossa seção introdutória.

2.1 *Corpus* do Português e *Corpus* PAISÀ

Os dados foram coletados do *Corpus* do Português e do *Corpus* PAISÀ, plataformas online que reúnem diversos textos em língua portuguesa e língua italiana, respectivamente. Esses *corpora* foram escolhidos devido a sua semelhança, já que ambos são constituídos de dados oriundos de textos escritos, elaborados na primeira metade da década de 2010 e extraídos de plataformas online.

As ocorrências do português brasileiro foram retiradas especificamente da aba Web-Dialetos do *Corpus* do Português, que reúne cerca de 1 bilhão de palavras de páginas da internet de quatro países lusófonos: Brasil, Portugal, Angola e Moçambique. Foram considerados para esta pesquisa apenas textos pertencentes à variante brasileira, em modalidade escrita e que foram compilados entre 2013 e 2014.

Figura 6: *Corpus* do Português



The screenshot shows the website 'o corpus do português' with a navigation bar and a table of corpora. The table lists three corpora: 'Gênero / Histórico' (45 million words, created in 2006), 'Web / Dialetos *' (1 billion words, created in 2016), and 'NOW (2012 - 2019)' (1.1 billion words, created in 2018). The website also features a map of Portuguese-speaking countries and a 'TOUR' button.

		Corpus	Tamanho	Criado
1	Info	Gênero / Histórico	45 milhões de palavras	2006
2	Info	Web / Dialetos *	1 mil milhão de palavras	2016
3	Info	NOW (2012 - 2019)	1,1 mil milhão de palavras	2018

Clicar no link [Info] acima para mais detalhes.

Fonte: Davies; Ferreira (2006)

Disponível em: <https://www.corpusdoportugues.org>.

Acesso em: 30 de janeiro de 2024

Enquanto isso, para a coleta de dados em italiano, foi utilizado o *Corpus PAISÀ* (*Piattaforma per l'Apprendimento dell'Italiano Su corpora Annotati*) que conta com cerca de 250 milhões de palavras de textos em modalidade escrita, retirados da web entre setembro e outubro de 2010.

Figura 7: *Corpus Paisà*

The screenshot shows the PAISÀ web interface. At the top, there is a navigation bar with 'Benvenuti', 'Corpus dell'Italiano', 'Progetto PAISÀ', and 'Materiale di consultazione'. Below this, there are tabs for 'Ricerca Semplice', 'Ricerca Avanzata', 'Ricerca CQP', and 'Filtri'. The 'Ricerca Avanzata' tab is active. The main search area is divided into two sections, 'Parola 1' and 'Parola 2'. Each section has a 'Forma' dropdown (set to 'posto' for Parola 1 and 'che' for Parola 2), a 'Lemma' dropdown, and a 'POS' dropdown. There are also input fields for 'da' and 'a' (both set to '1') and a 'occorrenze' field. Below these are checkboxes for 'ignora maiuscole/minuscole' and 'ignora diacritici (come è, è)'. To the right of each section is a '# parole' field with a dropdown set to '0'. At the bottom, there is a pagination bar showing '1-15/987' and 'Vai alla pagina 1 / 66 Vai'. Below the search area, there is a table of search results with columns for the search terms and the corresponding text snippets.

Parola 1	Parola 2	Snippet
posto	che	organista titolare della Thomaskirche , era stato occupato anche da
posto	che	calcio del Napoli prendendo il fu di Ascarelli . Durante
posto	che	Non si conosce un altro attirati tanti giovani di tutti
posto	che	" ma non conservò questo due anni , avendo acconsentito
posto	che	sia più semplice trovare un entrar- vi . Prima di

Fonte: Lyding, V.; Stemle, E.; Borghetti, C.; Brunello, M.; Castagnoli, S.; Dell'Orletta, F.; Dittmann, H.; Lenci, A.; Pirrelli, V. (2014)

Disponível em: <https://www.corpusitaliano.it/it/index.html>

Acesso em: 30 de janeiro de 2024

No caso da língua italiana, foi utilizada a aba de *Ricerca Avanzata* (Pesquisa Avançada), em que se pode filtrar apenas dados que tenham as duas unidades linguísticas investigadas (“dato” + “che” e “posto” + “che”) em sequência, constituindo a estrutura dos conectivos que estamos investigando e, por sua vez, tornando a coleta mais rápida.

2.2 Etapas e fatores de análise

Após a definição de quais *corpora* seriam utilizados, a análise foi dividida em duas etapas: a primeira consistiu na coleta dos 90 primeiros dados de cada microconstrução, totalizando uma amostra de 360 dados⁴. A escolha desse número de

⁴ É importante destacar que os dados em língua portuguesa já haviam sido previamente coletados e analisados durante a fase de iniciação científica da autora, ocorrida entre os anos de 2018 a 2020. Sendo assim, esta primeira fase se concentrou na coleta apenas de dados em língua italiana.

dados foi arbitrária, isto porque, apesar de ambos os *corpora* apresentarem o número total de dados referentes à construção requisitada, grande parte desses resultados não se mostraram compatíveis com nosso objeto de estudo.

Isso se deve, em grande parte, a pluralidade que os termos “dado” e “posto” em português, assim como “dato” e “posto” em italiano possuem, isso porque “dado” e “dato” foram encontrados com o sentido de “dado estatístico”, enquanto “posto”, em ambas as línguas, pode ser encontrado com sentido de “cargo de trabalho” ou “local”. No caso, quando estes termos, dotados desses significados, eram encontrados, seguidos por um conectivo ou um pronome relativo “que”, eram desconsiderados de nossa mostra, tendo em conta que nosso objetivo era analisar microconstruções que são encabeçadas pelos conectivos [dado que], [posto que], [dato che] e [posto che].

Além disso, havia também um grande número de dados repetidos. Sendo assim, o quantitativo apresentado pelos *corpora* não correspondem diretamente ao nível de produtividade que essas microconstruções realmente apresentam em seus respectivos sistemas linguísticos.

Tendo isso em consideração, a fim de descrever o comportamento das orações hipotáticas em português e em italiano, mostrando suas características semelhantes e divergentes, adotamos os seguintes fatores de análise:

- (a) valor semântico das construções hipotáticas;
- (b) arranjo linear das construções hipotáticas em relação às matrizes;
- (c) estrutura informacional das construções hipotáticas;
- (d) simultaneidade modo-temporal dos verbos presentes nas orações hipotáticas e em suas orações principais;
- (e) codificação formal do sujeito da oração hipotática;
- (f) correferencialidade entre os sujeitos das orações hipotáticas e suas principais;
- (g) frequência *type* e *token* dos itens verbais pertencentes às orações hipotáticas.

A partir desses fatores, foi realizada uma análise qualitativa e quantitativa dos dados, além de alguns cruzamentos, embasados em hipóteses prévias que havíamos, realizados com o auxílio do programa *Excel*.

Cada um desses fatores será detalhado e exemplificado no próximo capítulo. Da mesma forma, as hipóteses específicas, apresentadas na Introdução deste trabalho, também serão retomadas junto de cada fator e expostas novamente na seção de Análise.

3. ANÁLISE

Neste capítulo, serão apresentados e discutidos os resultados obtidos na análise dos dados, com o propósito de examinar o comportamento das orações hipotáticas introduzidas por “dado que” e “posto que” na língua portuguesa, e “dato che” e “posto che” em língua italiana. Nosso objetivo é verificar a hipótese de que essas construções possuem propriedades funcionais e discursivas distintas em cada língua.

Entendemos, assim, que diferentes padrões de uso são adotados por cada construção, os quais estão vinculados ao contexto, bem como a questões sintáticas e semânticas, pragmáticas e discursivas. Essas tendências serão elucidadas a partir dos critérios de análise que foram empregados.

Primeiro, serão anunciados os resultados pertinentes à língua portuguesa e, em seguida, referentes à língua italiana. As comparações dentro do próprio sistema linguístico serão feitas o tempo todo, enquanto apresentamos os resultados. As considerações acerca desses resultados, sobretudo a comparação entre as duas línguas, assim como a corroboração ou não com nossas prévias hipóteses, serão apresentadas com mais detalhes em nosso último capítulo.

3.1 Valor semântico, ordenação e status informacional da oração hipotática

Como já abordado anteriormente, as pesquisas realizadas por Neves (1999) e Diessel (2013) revelam que existe uma estreita correlação entre o valor semântico de uma oração hipotática, sua posição em relação à oração matriz e o grau de novidade da informação que ela transmite. Com base nessas considerações, os resultados de cada um desses fatores, bem como a interseção entre eles, serão apresentados nas próximas subseções.

3.1.1 Valor semântico

Os estudos realizados por Neves (1999) evidenciam que existe um vínculo entre orações causais, concessivas e condicionais e que essas relações circunstanciais podem se organizar em um *continuum*, como evidenciado pela Figura 4, presente no capítulo de pressupostos teóricos.

A autora propõe que a relação causal diz respeito à conexão causa-consequência, ou causa-efeito, entre dois eventos. Enquanto a oração principal manifesta a consequência ou efeito, a oração hipotática (evidenciada em **negrito**) exprime a causa, como pode ser observado nos exemplos a seguir:

a) Oração hipotática causal

(5) “Inúmeras metodologias são consultadas pelos professores e todas elas têm seus aspectos positivos e negativos, **posto que ninguém desenhou uma solução única.**” (*Corpus* do Português)

(6) “Tecnicamente, si potrebbe notare che il Congresso di Vienna non si svolse mai realmente, **posto che il Congresso non si riunì affatto in sessione plenaria**, e la maggior parte delle discussioni avvenne in sessioni informali tra le grandi potenze.” (*Corpus* PAISÀ)

[Tecnicamente, se poderia notar que o Congresso de Viena nunca se desenvolveu realmente, **já que o Congresso não se reuniu sequer em sessão plenária**, e a maior parte das discussões aconteceram em sessões informais entre as grandes potências.]

No exemplo (5), é evidenciado que todas as metodologias possuem aspectos positivos e negativos e isso se deve ao fato de que não existe uma única solução. Sendo assim, é evidenciado o efeito (a existência de uma pluralidade de metodologias) de não se haver uma única solução. Da mesma forma, o exemplo (6) evidencia que o motivo pelo qual o Congresso Viena jamais se desenvolveu é porque este mesmo congresso nunca sequer se reuniu em uma sessão plenária. Desta forma, primeiro se evidencia o efeito (o não desenvolvimento deste Congresso), para depois ser atribuída uma causa para isso (o fato de que nunca houveram sessões plenárias).

Por outro lado, as orações condicionais, ainda segundo a autora, são observadas tradicionalmente pelas relações lógico-semânticas que elas exprimem. Uma construção condicional é estabelecida pela relação entre uma oração que exprime condição (oração hipotática) e uma que exprime o que é condicionado (oração principal). Essa relação entre elas se apoia em uma hipótese, razão pela qual o termo “período hipotético” é utilizado para designar esses tipos de construções na língua italiana e em alguns estudos clássicos da língua portuguesa.

Em relação aos nossos resultados, não foi encontrada nenhuma relação dita puramente condicional nas microconstruções em língua portuguesa. Ainda que seja muito difícil qualificar em qual parte do espectro uma oração hipotática está localizada, as construções encontradas em português se aproximam mais do polo das orações causais do que exclusivamente do valor condicional, sendo classificadas nesta pesquisa como orações causal-condicionais, que serão melhor explicadas e exemplificadas ainda neste capítulo. Entretanto, o mesmo não pode ser dito em relação às orações em italiano, já que podemos observar esse tipo de valor semântico no exemplo abaixo:

b) Oração hipotática condicional

(7) “Un altro esempio che viene fatto, non per il licenziamento ma per il fattore di rischio, è Kylie Minogue con il suo nuovo album “X”. Dicono che se non venderà abbastanza ci sarà un vero tracollo finanziario dell’azienda. Ma credete davvero che una come Kylie, posto che faccia flop, venga lasciata a piedi?” (Corpus PAISÀ)

[Outro exemplo que vem dado, não pela demissão, mas pelo fator de risco, é Kylie Minogue com seu novo álbum “X”. Dizem que se não vender o suficiente haverá um verdadeiro colapso financeiro da empresa. Mas vocês realmente acreditam que alguém como Kylie, caso faça flop, será deixada de lado?]

No exemplo (7), que fala sobre a carreira de Kylie Minogue, é evidente uma relação hipotética, em que se ilustra uma possível consequência caso determinada condição seja (ou não) satisfeita. Neste caso, a oração hipotática hipotetiza uma situação em que caso o novo trabalho da cantora seja um fracasso, pode haver como possível consequência o seu esquecimento ou negligência, por parte do público ou da própria empresa. Dessa forma, existe então uma relação entre a oração que exprime condição (neste caso, “posto che faccia flop”) e a que exprime o que é condicionado (“una come Kylie... venga lasciata a piedi?”).

Como evidenciado alguns parágrafos acima e apresentado no capítulo de pressupostos teóricos, existe uma relação intrínseca entre os valores causal, condicional e concessivo. Levando em conta os estudos relacionados por Neves (1999), esses sentidos podem ser combinados, ocasionando orações ditas híbridas, que apresentam mais de uma circunstância ao mesmo tempo.

Considerando que esses valores se organizam em um continuum, como demonstrado na Figura 4, os valores causal-condicional e condicional-concessivo também são possíveis no estudo de orações hipotáticas. No caso desta pesquisa, o último valor não foi encontrado, enquanto foi encontrado um percentual pequeno com as orações ditas causal-condicionais, tanto em língua portuguesa, quanto em língua italiana, como podemos observar nos dois exemplos abaixo:

c) Oração causal-condicional

(8) “**Dado que os gurus e os Budas não existem independentemente dos discípulos ou estudantes**, segue-se que nem os professores nem os discípulos existem como entidades totalmente independentes, como dois postos sólidos e

concretos, cada um deles existindo por si próprio mesmo se o outro nunca tivesse existido.” (*Corpus* do Português)

- (9) “Se i soggetti dello studio non fossero stati allenati, avrebbero tutti ottenuto un considerevole miglioramento della forza senza variazioni di massa muscolare entro le prime settimane di allenamento. In quel caso, questi rapidi adattamenti neurali si sarebbero verificati in entrambi i gruppi, ma non spiegherebbero l'eventualità che l'arginina apparentemente avesse triplicato il miglioramento nell'attivazione del sistema nervoso. Tuttavia, **dato che i soggetti erano allenati, la situazione è ancora più sconcertante.**” (*Corpus* PAISÀ)

[Se os sujeitos do estudo não tivessem sido treinados, teriam todos alcançado uma melhora considerável na força, sem variações de massa muscular nas primeiras semanas de treinamento. Nesse caso, estas rápidas adaptações neurais teriam sido verificadas em ambos os grupos, mas não explicariam por que a arginina aparentemente teria triplicado a melhora na ativação do sistema nervoso. Porém, **se / já que os sujeitos foram treinados, a situação é ainda mais intrigante.**]

Tanto o exemplo (8) quanto o (9) evidenciam uma relação mista de sentidos. No primeiro deles, além da situação de condição traçada entre a existência de budas e gurus e professores e seus discípulos, em que uma existe independente da outra, existe também uma relação de causalidade, pois a não relação de independência entre budas e gurus (causa) tem como consequência uma relação de não existência dessa dependência também entre professores e alunos, havendo, portanto, um paralelo comparativo entre essas entidades.

Conquanto ao exemplo (9), acontece algo semelhante. Ao mesmo tempo em que se evidencia uma condição (a possibilidade de os sujeitos participantes desse estudo terem sido treinados), evidencia também uma consequência em que, devido a esse fato, a situação é intrigante. Da mesma forma, existe uma relação de causalidade entre essas orações, devido ao treinamento desses sujeitos (a causa) teria como efeito tornar a situação (de variação de massa muscular) intrigante. Sendo assim, pode-se apresentar dois possíveis sentidos, que não se anulam: um em que se os sujeitos foram treinados, a situação é intrigante (valor condicional), assim como, porque os sujeitos foram treinados, a situação é intrigante (valor causal).

Em relação às orações concessivas, o sentido dessas construções pode ser definido como aquilo que contraria as expectativas. Neste caso, a oração hipotática concessiva tende a expressar um fato (ou alguma noção) que poderia se caracterizar como um

impedimento para a realização do evento presente na oração principal, entretanto, apesar desse empecilho, o evento é mantido.

Não houve ocorrências com esse sentido em língua italiana, mas alguns poucos casos foram encontrados com a oração introduzida por “posto que” em português, como podemos observar no exemplo abaixo:

d) Oração hipotática concessiva

(10) “É, pois, impossível, como justamente observa o mesmo autor, julgar com o eu intelectual a maneira de agir do eu afetivo. **Posto que a vida afetiva e a vida intelectual sejam demasiado heterogêneas para que uma se reduza à outra, procedemos sempre sem levar em conta a diferença** que separa os sentimentos da inteligência.” (*Corpus do Português*)

No exemplo (10), é evidenciado pela oração hipotática como, apesar de a vida afetiva e a intelectual serem bastante diferentes, isso não é considerado pelos seres humanos, pois estes continuam a agir sem levar em conta essa diferença. Sendo assim, este exemplo demonstra claramente a relação entre algo que deveria ser um empecilho (a diferença entre afeto e intelecto), mas que, no fim, não é levado em consideração.

Tendo isso em conta, apresentamos, na tabela abaixo, os resultados acerca dos valores semânticos das microconstruções em língua portuguesa:

Tabela 1: Valor semântico das orações hipotáticas em português

	Causal	Causal- Condicional	Condicional	Concessiva	Total
[dado que]or.	85 (94,4%)	5 (5,6%)	0	0	90 (100%)
[posto que]or.	84 (93,3%)	4 (4,4%)	0	2 (2,3%)	90 (100%)
Total	169 (93,9%)	9 (5%)	0	2 (1,1%)	180 (100%)

Fonte: Elaboração própria

A Tabela 1 apresenta os índices de presença dos sentidos identificados na coleta de dados das microconstruções em análise. As duas microconstruções tendem fortemente a apresentar sentido causal (ambas com quase 95% dos dados com esse sentido). Além disso, [dato que]or. se apresenta em algumas poucas ocorrências veiculando o sentido condicional (5,6% dos casos), assim como [posto que]or (4,4% dos casos). No caso das orações introduzidas por [posto que], existe uma terceira via em que, em alguns raros usos, essas orações podem apresentar o valor semântico concessivo (2,3% dos casos).

Como anteriormente mencionado, não foram encontrados casos de uso exclusivamente condicional com nenhuma das duas microconstruções em português. Também vale apontar que, por mais que existam os sentidos condicional e concessivo no que tange a essas construções, a disparidade entre a presença da semântica causal e dos outros sentidos encontrados é grande, visto que a proporção percentual bruta indica uma clara preferência pela causalidade (93,9% dos casos).

Passemos agora aos resultados desse fator com as microconstruções em língua italiana:

Tabela 2: Valor semântico das orações hipotáticas em italiano

	Causal	Causal- Condicional	Condicional	Concessiva	Total
[dato che]or.	87 (96,7%)	3 (3,3%)	0	0	90 (100%)
[posto che]or.	61 (67,8%)	0	29 (32,2%)	0	90 (100%)
Total	148 (82,2%)	3 (1,7%)	29 (16,1%)	0	180 (100%)

Fonte: Elaboração própria

Na tabela 2, é possível verificar que, novamente, o sentido de causalidade é o preferido dos conectivos aqui analisados. Assim como nas microconstruções equivalentes no português, [dato che]or. tende mais ao sentido de causalidade (96,7%) do que quando esse valor se combina à condicionalidade (3,3%). Já [posto che]or. se apresenta em um número significativo de dados com sentido de condicionalidade (32,2%), ainda que tenha

uma tendência maior (67,8%) à causalidade. O alto percentual de usos do sentido de condicionalidade de [posto che]or. em relação à falta de uso neste sentido com [dato che]or. revela uma importante diferença do emprego dos dois conectivos.

Em suma, o português e o italiano se distinguem principalmente no que tange ao sentido concessivo, presente no português, mas não no italiano. Também se distinguem no uso do sentido condicional, que só se apresenta nas microconstruções do italiano, tendo em vista as microconstruções que estudamos.

3.1.2 Ordenação da oração hipotática

Outro fator analisado foi a ordenação das orações hipotáticas em relação às suas orações principais. De forma geral, a literatura demonstra que existem duas posições canonicamente esperadas: a anteposta e a posposta. Entretanto, além dessas, também foram encontradas ocorrências de orações intercaladas, apenas em língua italiana.

Estas orações se caracterizam por se inserirem dentro de suas principais e aparecerem, normalmente, isoladas por alguma pontuação, como vírgulas, travessão, parênteses ou colchetes. Pragmaticamente, pode-se dizer que essas orações tendem a emitir uma opinião, uma observação, uma ressalva ou até mesmo uma advertência para o interlocutor. Entretanto, a função específica exercida por esse tipo de ordenação não foi investigada nesta pesquisa.

Tendo isto estabelecido, a seguir apresentamos essas três possíveis colocações das orações hipotáticas (evidenciadas em negrito) em relação às suas orações principais (sublinhadas):

a) Anteposta à oração matriz

(11) “**Dado que o valor comercial das galinhas “gastas” é baixíssimo, há produtores que sequer se dão ao trabalho de as vender para abatedouros.**” (*Corpus do Português*)

(12) “Ma **dato che l'associazione tra colore e numero è arbitraria** si possono pensare ad infinite associazioni che darebbero infiniti risultati diversi.” (*Corpus PAISÀ*)

[“Mas **dado que a associação entre cor e número é arbitrária, se pode pensar em infinitas associações** que dariam infinitos resultados diferentes”.]

b) Intercalada com a oração matriz

(13) “Intendo dire che il mercato del carbonio, **posto che funzioni** (e questo è ancora da dimostrare), è un mercato e come tale ha creato un'altra opportunità per fare

soldi a chi si occupa di finanza (non credo siano le persone che ne necessitano maggiormente)”. (*Corpus PAISÀ*)

[“Quero dizer que o mercado de carbono, **desde que funcione** (e isso ainda precisa ser demonstrado), é um mercado e, como tal, criou outra oportunidade para aqueles envolvidos no setor financeiro ganharem dinheiro (não creio que sejam eles as pessoas que majoritariamente necessitam disso)”].

c) Posposta à oração matriz

(14) “Envolvidos pelo delírio não percebem que a sordidez golpista já faz parte do conhecimento do mundo e nada vai mudar, **posto que o golpe é real**.” (*Corpus do Português*)

(15) “Se si installa da Live CD, i pacchetti di traduzione nella lingua prescelta al momento dell'avvio vengono installati automaticamente durante l'installazione, **posto che sia rilevata una connessione internet funzionante**.” (*Corpus PAISÀ*)

[“Se se instala a partir do Live CD, os pacotes de tradução na língua pré-escolhida no momento da inicialização serão instalados automaticamente durante a instalação, **desde que seja detectada uma conexão de Internet que funcione**.”]

Os exemplos (11) e (12) ilustram como as orações antepostas se comportam, se posicionando antes de suas orações principais, enquanto o exemplo (13) demonstra como as orações hipotáticas intercaladas se apresentam, se localizando entre o sujeito e o predicado da oração principal, sendo evidenciada pelas vírgulas. Já os exemplos (14) e (15) elucidam a forma como as orações pospostas aparecem, se colocando logo após suas orações principais.

As hipóteses acerca dessas preferências colocacionais serão discutidas nas próximas subseções, entretanto, é importante enfatizar que os estudos de Diessel (2013) relacionam diretamente este fator ao de informatividade, realçando que a ordem das orações hipotáticas está relacionada ao grau de novidade das informações que elas trazem, assim como, ainda de acordo com o autor, mas também levando em conta os estudos de Neves (1999), ao sentido que elas possuem.

Com base nessas considerações, reunimos na tabela abaixo os resultados pertinentes a esse fator em relação à língua portuguesa:

Tabela 3: Ordenação das orações hipotáticas em português

	Anteposta	Intercalada	Posposta	Total
[dado que]or.	44 (48,9%)	0	46 (51,1%)	90 (100%)
[posto que]or.	10 (11,1%)	0	80 (88,9%)	90 (100%)
Total	54 (30%)	0	126 (70%)	180 (100%)

Fonte: Elaboração própria

A Tabela 3 indica que [dado que]or. apresenta uma quase homogeneidade quanto à ordenação, podendo aparecer quase igualmente anteposta (48,9% dos casos) e posposta (51,1% dos casos). Por outro lado, [posto que]or. apresenta uma preferência pela ordenação posposta bem mais significativa (88,9% dos casos), havendo a posição anteposta numa frequência bem menor (11,1% dos casos).

Se considerarmos as porcentagens brutas totais, poderíamos afirmar que nas microconstruções parece haver uma preferência pela ordenação posposta. No entanto, a distribuição de ordenação da estrutura [dado que]or. não estaria plenamente de acordo com essa constatação. Sendo assim, é possível que a ordenação seja um diferenciador de [dado que]or em relação [posto que]or.

As orações em italiano parecem ser mais categóricas quanto à ordenação, como demonstra a Tabela 4:

Tabela 4: Ordenação das orações hipotáticas em italiano

	Anteposta	Intercalada	Posposta	Total
[dato che]or.	7 (7,8%)	0	83 (92,2%)	90 (100%)
[posto che]or.	26 (28,9%)	3 (3,3%)	61 (67,8%)	90 (100%)

Total	33 (18,3%)	3 (1,7%)	144 (80%)	180 (100%)
--------------	---------------	-------------	--------------	---------------

Fonte: Elaboração própria

A Tabela 4 indica que [dato che]or é predominantemente usado em posição posposta (92,2%) à oração principal, permitindo-se pouquíssimo à posição anteposta (7,8%) e em nenhum caso à posição intercalada. Já [posto che]or. apresenta maior pluralidade de usos quanto a ordenação, sendo preferencialmente posposta (67,8%), porém, há uma significativa quantidade de orações antepostas (28,9%) e alguns poucos casos de orações intercaladas (3,3%).

3.1.3 Valor semântico x Ordenação

Como foi destacado no capítulo teórico, na visão de Diessel (2013), as orações condicionais tendem a preceder a oração principal enquanto as orações causais geralmente são colocadas após sua oração principal. O autor não faz nenhuma previsão acerca de qual seria a predileção das orações concessivas, sendo assim, de acordo com um estudo realizado por Neves (1999), postula-se que a ordenação das orações concessivas é ambígua, pois depende do propósito comunicativo do interlocutor.

Ainda de acordo com este trabalho realizado por Neves (1999), é defendido pela autora que a tendência de as orações condicionais serem antepostas se deve ao fato de que primeiro se enuncia a ocorrência de uma condição, neste caso a oração hipotática, que pode ou não ser satisfeita, para depois ser enunciada a oração que depende da concretização desta condição, expressa pela oração principal.

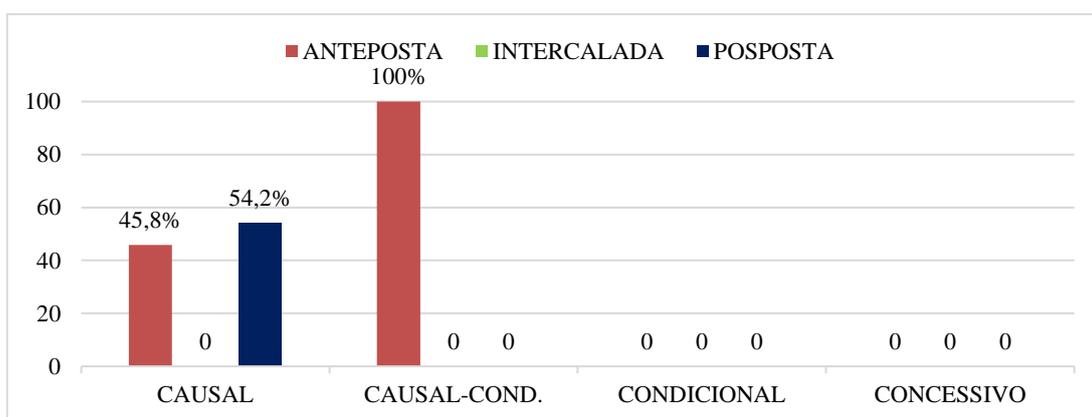
No caso das orações causais, é dito que normalmente se enuncia primeiro o efeito, expressado pela oração principal, e depois a causa, expressada pela oração hipotática. Já em relação às orações concessivas, esta ordem depende do objetivo de quem a produz, pois tendem a aparecer antepostas quando possuem uma função de tópico, em que se retoma aquilo que já foi dito, e pospostas quando tem a função de adendo.

De forma geral, essa função de adendo parece igualmente se aplicar às orações intercaladas, que aparecem bem no meio entre o sujeito e predicado de sua oração principal, já que, como mencionado na subseção de Ordenação, essas orações tendem a

funcionar como uma observação ou ressalva pertinente ao interlocutor. Entretanto, os estudos supramencionados não fazem menção a esta possível ordenação.

Tendo essas considerações em vista, realizamos o cruzamento entre Valor Semântico e Ordem, buscando verificar se, primeiro, os resultados serão compatíveis com essas hipóteses, e, em segundo, se as orações apresentam divergências de comportamento em relação a este fator. O gráfico a seguir combina os valores semânticos evocados por [dado que]or e suas preferências colocacionais:

Gráfico 1: Cruzamento entre valor semântico e ordenação com [dado que]or

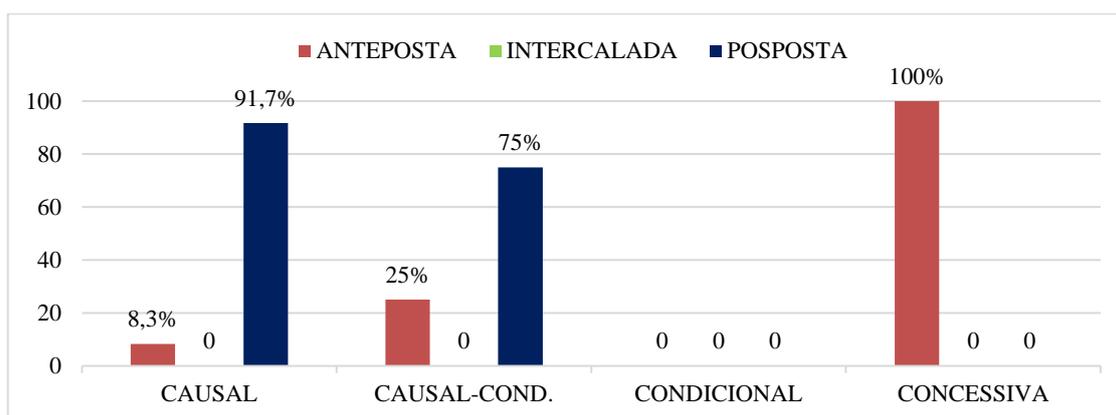


Fonte: Elaboração própria

Em relação aos resultados de [dado que]or, como anteriormente mencionado, não foi encontrado nenhum caso de oração intercalada na língua portuguesa. Entretanto, no que tange às orações causais, não parece haver uma predileção clara sobre a ordenação, já que 45,8% dos casos são antepostos e 54,2% dos casos são pospostos, evidenciando uma diferença mínima entre as duas posições. Enquanto isso, as orações que evocam o sentido de causa combinado a condicionalidade, ditas orações causal-condicionais, apresentaram todas as suas ocorrências na posição anteposta. Os demais valores semânticos não apresentaram nenhuma ocorrência com esta microconstrução.

No gráfico a seguir, verificamos o cruzamento entre valores semânticos e ordenação presentes nos dados coletados da estrutura [posto que]or:

Gráfico 2: Cruzamento entre valor semântico e ordenação com [posto que]or



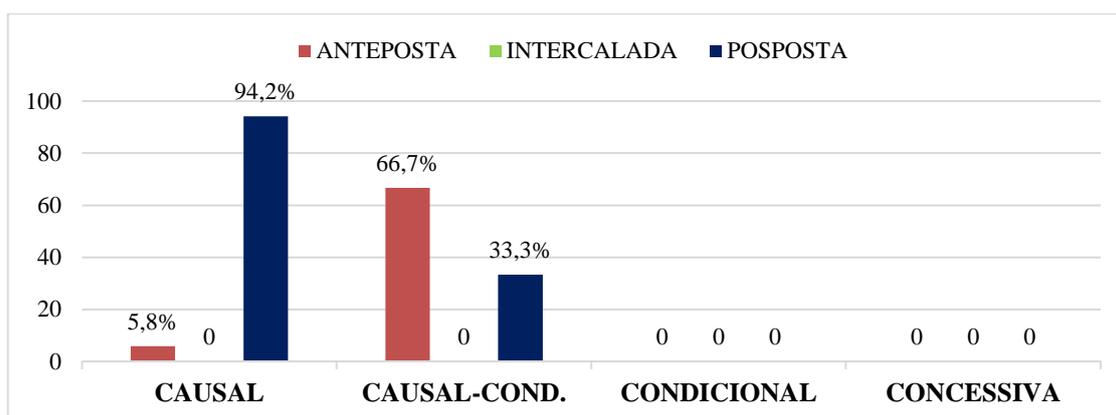
Fonte: Elaboração própria

No gráfico 2, [posto que]or parece seguir a tendência da outra microconstrução do português em análise. Assim como ocorreu com os resultados de [dado que]or, o sentido de causalidade apresenta uma preferência pela posposição, representando 91,7% dos casos, enquanto a anteposição totaliza apenas 8,3% dessas ocorrências.

Porém, diferentemente da microconstrução analisada anteriormente, cujos resultados correlacionam as orações causal-condicionais à anteposição, neste caso essas orações apresentam uma preferência pela posposição, havendo 75% de suas ocorrências ocorrendo após a oração principal, sendo a anteposição apenas 25% dos casos. Além disso, essa microconstrução apresenta ainda um terceiro valor semântico: a concessão. Os resultados mostram que 100% dos casos em que este sentido aparece, a posição predileta é a anteposição. E, por fim, da mesma forma que a oração [dado que]or, analisada anteriormente, não houve casos de orações plenamente condicionais com esta microconstrução.

Na página a seguir, apresentaremos os resultados dos cruzamentos entre valores semânticos e ordenação das microconstruções do italiano, iniciando por [dato che]or:

Gráfico 3: Cruzamento entre valor semântico e ordenação com [dato che]or

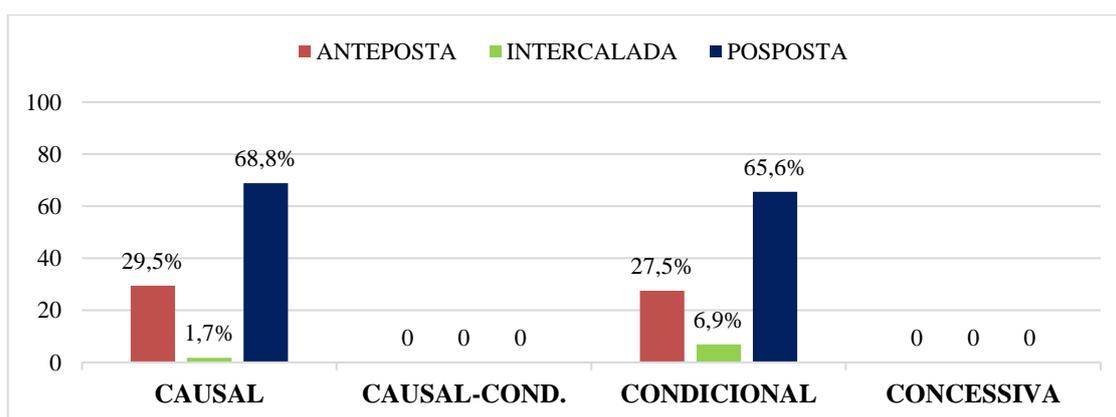


Fonte: Elaboração própria

No tocante à [dato che]or, as orações causais possuem uma clara predileção sobre a ordenação, apresentando 94,2% dos seus casos na posição posposta, contra apenas 5,8% na posição anteposta. Em relação às orações causal-condicionais, o resultado foi diferente, já que estas tendem a aparecer antes de suas orações principais, fazendo com que a posição anteposta seja equivalente à 66,7% dos casos, enquanto a posição posposta totaliza 33,3% dos casos. Em relação aos demais valores, não foi encontrada nenhuma ocorrência, assim como não houve casos de orações intercaladas.

No gráfico abaixo, apresentamos o cruzamento em relação à [posto che]or:

Gráfico 4: Cruzamento entre valor semântico e ordenação com [posto che]or



Fonte: Elaboração própria

Esse gráfico demonstra que em relação à causalidade, as orações introduzidas por “posto que” possuem uma tendência à posposição, totalizando 68,8% dos casos. Entretanto, apesar disso o número de orações antepostas se mostra significativo,

constituindo 29,5% dos casos. Concomitantemente, a posição intercalada aparece em terceiro lugar e em uma escala muito menor, totalizando apenas 1,7% dos casos.

O outro valor encontrado com esta microconstrução foi o valor condicional, visto que não foram encontradas ocorrências com os outros valores semânticos além dos supracitados. Da mesma forma que as orações causais, este sentido possui uma predileção pela posição posposta, apresentando 65,6% de seus casos com essa colocação, porém, também apresenta um número ligeiramente expressivo de orações antepostas, que constituem 27,5% dos casos. Conquanto às intercaladas, o número de orações que apareceram nesta posição é pequeno, sendo apenas 6,9% dos casos, porém, constitui uma porcentagem ligeiramente maior se comparado às ocorrências com as orações causais supramencionadas.

Em relação às nossas hipóteses, todas as quatro microconstruções corroboraram que a causalidade está associada à posposição, uma vez que a maioria das orações com esse valor semântico apareceram após suas orações principais. Entretanto, apesar de a tradição linguística sugerir que a posição preferida das orações condicionais é a anteposta, nossa pesquisa não corroborou com essa hipótese, pois as orações condicionais encontradas se apresentaram preferencialmente pospostas às suas matrizes.

Enquanto a causalidade tende a posposição, a condicionalidade tende à anteposição, sendo assim, não havia uma hipótese definida sobre qual seria a colocação preferida das orações causal-condicionais, presentes em nossos resultados. Porém, enquanto [dado que]or. e [dato che]or. apresentaram a maior parte de suas ocorrências na posição anteposta, os resultados de [posto que]or. se diferiram, visto que 75% das ocorrências apareceram pospostas às suas matrizes.

Da mesma forma, não existia uma hipótese específica para as orações concessivas, uma vez que seu posicionamento pode variar de acordo com o propósito comunicativo, sendo ambíguo. Todavia, as orações com este valor semântico, encontradas apenas nas ocorrências da microconstrução [posto que]or., apresentaram uma preferência categórica pela posição anteposta.

3.1.4 Status informacional

Em relação ao status informacional, a tradição postula que orações hipotáticas podem trazer informações pressupostas, que já foram dadas no discurso antecedente ou que o leitor/ouvinte possa deduzir através de pistas deixadas pelo escritor/locutor, ou não-

pressupostas, que não foram mencionadas antes no discurso e que não são inferíveis a partir de pistas linguísticas (cf. Lambrecht, 1994).

É importante frisar que em nossa pesquisa não analisamos apenas a informatividade de alguns referentes presentes na oração (como um sujeito que pode ser recuperado através de uma anáfora pronominal), mas de toda ela. É claro que a binaridade não é suficiente para explicar como as informações são captadas pelos usuários da língua dentro de um discurso, pois isto depende de muitos fatores, tal qual o conhecimento de mundo e o contexto no qual este indivíduo está inserido.

Entretanto, para esta pesquisa, foram consideradas apenas as informações que foram expressas textualmente e de forma clara nos dados investigados, segundo a abordagem de Diessel (2013) e Lambrecht (1994). Sendo assim, se a informação trazida pela oração hipotática já havia sido expressa em algum momento prévio ao texto, essa oração era considerada pressuposta. Se este não fosse o caso, era classificada como não-pressuposta.

Tendo isto estabelecido, a seguir damos dois exemplos, um em cada língua investigada, de orações consideradas pressupostas:

(16) “Primeiro, os indivíduos da lista de espera tendem a montar lobby para garantir e acelerar sua convocação, o que gera mais custo de transação e induz a contratação de mais servidores que o necessário. Segundo, *os concursos se tornam grandes eventos*, com número elevado de candidatos, o que eleva o custo de realização dos certames. [...] **Dado que os concursos são grandes eventos, opta-se por um método de teste que facilite a correção por meio eletrônico.**” (*Corpus do Português*)

(17) “La zona archeologica, nel diritto internazionale marittimo, è quella area marina la cui ampiezza *non può superare le 24 miglia* (circa 34 km) dalla costa (o più precisamente dalle linee di base dalle quali è misurata l'ampiezza delle acque territoriali) e nella quale lo stato costiero ha giurisdizione in materia di protezione del patrimonio culturale sottomarino. [...] Basti considerare che tale formulazione si riferisce a una zona ben più ristretta della piattaforma continentale, posto che la zona contigua non può superare le 24 miglia.” (*Corpus PAISÀ*)

[A zona arqueológica, no direito marítimo internacional, é aquela área marinha cuja largura *não pode ultrapassar 24 milhas* (cerca de 34 km) da costa (ou mais precisamente das linhas de base a partir das quais se mede a largura das águas territoriais) e na qual o estado costeiro tem jurisdição sobre a proteção do patrimônio cultural subaquático. [...] Basta considerar que esta formulação se refere a uma área muito mais estreita da plataforma continental, já que a área contígua não pode ultrapassar as 24 milhas.]

No exemplo (16), o fato expressado pela oração hipotática já era esperado pelo leitor, isto porque, alguns parágrafos acima, já havia sido citado como os concursos são ou se tornam grandes eventos. Desta forma, como traz uma informação que textualmente já havia sido expressa, trata-se de uma oração pressuposta. O mesmo ocorre com o exemplo (17), em que se reafirma a quantidade de milhas que não pode ser ultrapassada. Ao retroceder algumas sentenças no texto, é possível observar que a quantidade exata de milhas já havia sido mencionada, portanto, trata-se também de uma oração com informação pressuposta.

A partir disso, apresentaremos a seguir as tendências informacionais das microconstruções analisadas. A tabela abaixo apresenta a frequência da pressuposição, ou ausência dela, nos dados coletados do português brasileiro:

Tabela 5: Status informacional das orações hipotáticas em português

	Pressuposta textualmente	Não-Pressuposta	Total
[dado que]or.	18 (20%)	72 (80%)	90 (100%)
[posto que]or.	8 (8,9%)	82 (91,1%)	90 (100%)
Total	26 (14,4%)	154 (85,6%)	180 (100%)

Fonte: Elaboração própria

Na Tabela 5, ambas orações hipotáticas ocorrem preferencialmente com status informacional do tipo não-pressuposto. Em [dado que]or as informações não-pressupostas são predominantes, constituindo 80% dos casos, enquanto 20% das informações presentes nessas orações hipotáticas já haviam sido previamente expressas. Já com [posto que]or, a diferença é mais saliente, em virtude de 91,1% dos casos são de informação não-pressuposta, ao mesmo tempo que apenas 8,9% dos casos são de informação pressuposta.

Passemos então ao panorama das orações hipotáticas em italiano em relação à pressuposição:

Tabela 6: *Status* informacional das orações hipotáticas em italiano

	Pressuposta textualmente	Não-Pressuposta	Total
[dato che]or.	6 (6,7%)	84 (93,3%)	90 (100%)
[posto che]or.	11 (12,2%)	79 (87,8%)	90 (100%)
Total	17 (9,4%)	163 (90,6%)	180 (100%)

Fonte: Elaboração própria

Na língua italiana, a situação é semelhante. As ocorrências com [dato che]or. também possuem uma tendência à não-pressuposição, apresentando 93,3% dos casos com informações que não haviam sido mencionadas *a priori* no discurso, enquanto apenas 6,7% se enquadram como orações pressupostas. O quadro se repete quando observamos a microconstrução [posto che]or em que 87,8% dos casos são de orações não-pressupostas contra 12,2% de orações pressupostas.

A diferença entre ambas as microconstruções é mínima, porém é relevante destacar que o número de orações pressupostas com [posto che]or é praticamente o dobro do que o encontrado com [dato che]or.

Como anteriormente mencionado, de acordo com a tradição assim como os estudos realizados por Diessel (2013), este fator pode estar diretamente atrelado à forma como essas orações se distribuem na ordem linear do discurso, isto é, se elas se apresentam antes, depois ou entre suas orações principais. Sendo assim, buscando mapear qual o comportamento dessas microconstruções, a próxima seção se dedica a um cruzamento entre esses dois fatores.

3.1.4 *Status* informacional x Ordenação

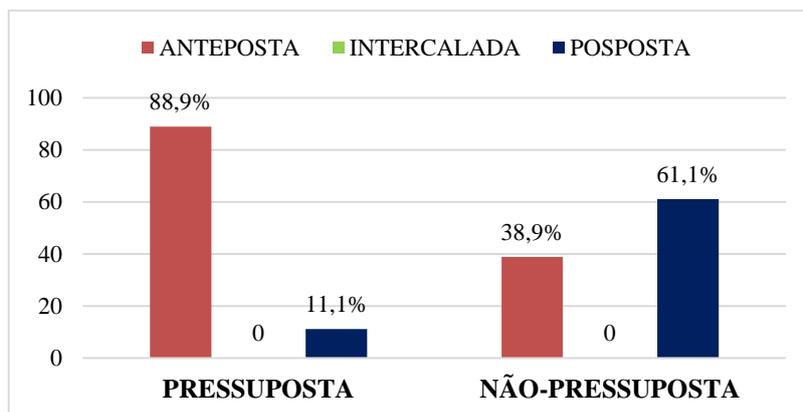
De forma geral, a literatura demonstra que orações hipotáticas que antecedem suas orações principais normalmente contêm informações que anteriormente já haviam sido ditas ou que eram de conhecimento prévio do leitor/ouvinte, enquanto orações hipotáticas posteriores à oração principal tendem a apresentar informações novas.

Como visto na subseção 3.1.2, nesta pesquisa também foi encontrada uma terceira possibilidade de colocação, apesar de pouco expressiva: a posição intercalada. Nesta, a microconstrução em estudo se aloja, geralmente, entre o sujeito de sua oração principal e seu predicado. Sobre esta ordem, o autor não faz nenhuma especulação acerca de qual seria o grau de novidade da informação codificada pela oração hipotática. Entretanto, pressupomos que, por normalmente assumirem uma função pragmática de adendo, essas orações também apresentariam informações que seriam novidade para o leitor.

Sendo assim, ponderamos que seria relevante cruzar esses dois fatores, visando observar como essas microconstruções se comportam em relação a seu *status* informacional e sua ordenação.

O gráfico abaixo apresenta o panorama referente a microconstrução [dado que]or:

Gráfico 5: Cruzamento entre *status* informacional e ordenação com [dado que]or

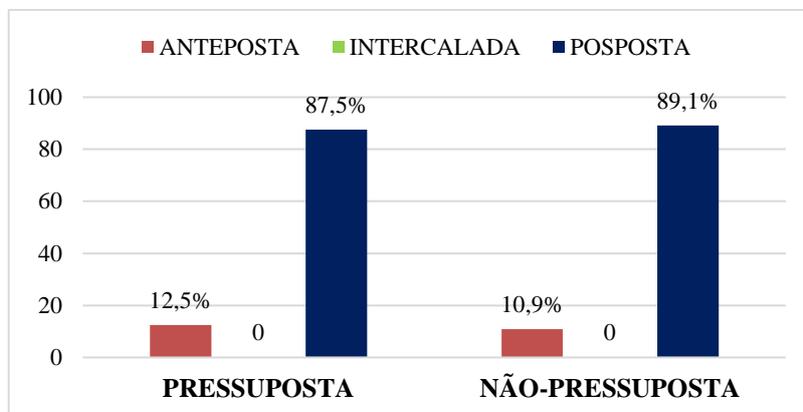


Fonte: Elaboração própria

O gráfico 5 denota uma clara preferência das orações pressupostas pela posição anteposta, visto que 88,9% das orações desse tipo se apresentam anteriores à sua oração principal, enquanto apenas 11,1% das ocorrências se enquadram como orações pospostas. Em contrapartida, os dados com informação não-pressuposta apresentam em sua maioria a posição posposta, constituindo 61,1% dos casos. Porém, a diferença de posições neste tipo de oração não é tão acentuada como no caso das pressupostas, já que 38,9% dos casos de orações não-pressupostas tenderam à anteposição. Vemos que, de um modo geral, a hipótese de Diessel acerca da relação entre ordenação e pressuposição se confirma: as orações pressupostas iniciadas por *dado que* tendem a se apresentar no antes da oração matriz, enquanto as não-pressupostas, aquelas que trazem informações novas, tendem a se apresentar pospostas à matriz.

A seguir, ilustra-se o panorama da estrutura [posto que]or. quanto a esse cruzamento:

Gráfico 6: Cruzamento entre *status* informacional e ordenação com [posto que]or

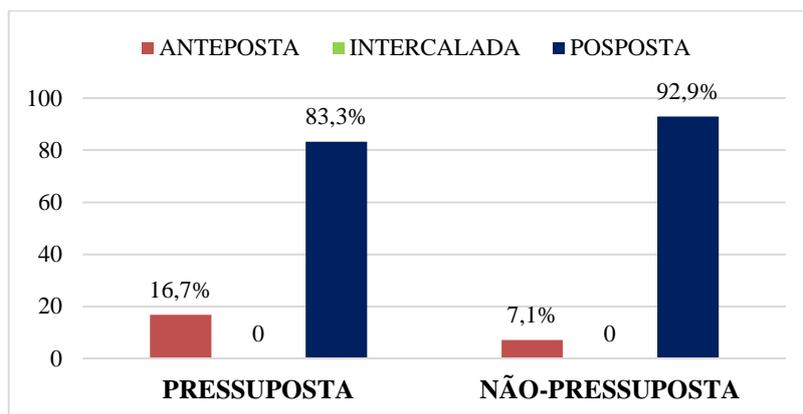


Fonte: Elaboração própria

No gráfico 6, existe uma preferência pela posição posposta independentemente do tipo de informação que é veiculada pelas orações iniciadas por [posto que], em que 87,5% das orações pressupostas se apresentaram nesta posição, assim como 89,1% das orações não-pressupostas. E, ainda que a anteposição tenha sido encontrada nos dados, sendo 12,5% dos casos de informação pressuposta e 10,9% dos casos de informação não-pressuposta, ela é minoritária quando comparada à posposição. Diferentemente, esse resultado vai contra às nossas expectativas (baseadas em Diessel, 2013), pois tanto as pressupostas quanto as pospostas com “posto que” tendem a se apresentar pospostas à matriz.

Para que tenhamos um quadro comparativo completo, passamos aos resultados das microconstruções do italiano quanto ao cruzamento do *status* informacional e da ordenação, iniciando pela microconstrução [dato che]or.:

Gráfico 7: Cruzamento entre *status* informacional e ordenação com [dato che]or



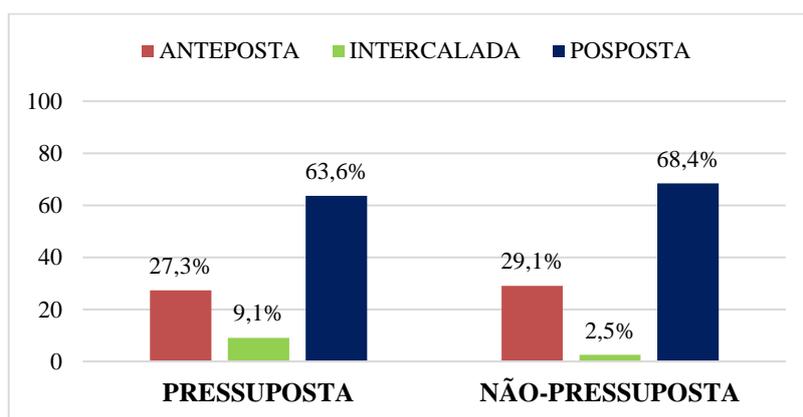
Fonte: Elaboração própria

Em relação à [dato che]or, assim como a oração anteriormente analisada em PB, o fato de a informação presente na oração hipotática ser ou não do conhecimento prévio do autor, não interfere em sua ordenação, já que em ambos os casos, a posição preferida para esta microconstrução é a posposição, havendo 83,3% de orações pospostas e pressupostas e 92,9% de orações pospostas e não-pressupostas.

Entretanto, é importante destacar que houve um número ligeiramente mais expressivo de ocorrências ante-postas em relação à pressuposição, já que 16,7% das orações pressupostas encontravam-se antes de suas orações principais.

A seguir, tem-se o cruzamento desses fatores com [posto che]or:

Gráfico 8: Cruzamento entre *status* informacional e ordenação com [posto che]or



Fonte: Elaboração própria

A respeito dessa microconstrução, os resultados se distribuem de maneira diferente. Independentemente do tipo de informação veiculada, em ambos os casos a

posposição apareceu como ordenação preferida, totalizando 63,6% dos casos em relação às pressupostas, e 68,4% conquanto às não-pressupostas.

Apesar disso, houve também um número significativo de orações antepostas, havendo 27,3% das orações pressupostas aparecendo antes de suas orações principais, e 29,1% de orações não-pressupostas ocorrendo posteriormente às suas orações principais.

Além disso, diferentemente das outras microconstruções em estudo, esta apresentou também alguns poucos casos de orações intercaladas. Neste caso, 9,1% das orações pressupostas apareceram no meio de sua oração principal, alojada entre seu sujeito e seu predicado, enquanto 2,5% das orações não-pressupostas também ocorreram nesta colocação.

Nossa hipótese era de que orações adverbiais antepostas serviriam para apresentar informações que são pragmaticamente pressupostas, enquanto orações pospostas tenderiam a apresentar informações novas. Para as orações intercaladas, pressupomos que, por frequentemente desempenharem uma função pragmática de adendo, essas orações também apresentariam informações novas para o leitor.

Quanto à pressuposição, nossa hipótese foi confirmada apenas para [dado que]or., já que a maioria dos casos pressupostos com esta construção apareceu na posição anteposta. Em contraste, as demais orações pressupostas tendem a aparecer na posição posposta. Por outro lado, em relação à não-pressuposição, todas as microconstruções foram condizentes com a hipótese de que orações que instanciam informações novas tendem à posposição.

Na próxima subseção, apresentaremos os fatores referentes ao grau de integração entre as orações, buscando mapear se existe um maior nível de integração sintática e, por consequência, semântico-pragmático, entre as orações hipotáticas e suas principais.

3.2 Elementos de integração

Para esta pesquisa também foi considerado o Princípio da Integração. Com base nos estudos de Givón (1979), acredita-se que, quanto maior a integração, maior também será a integração semântica ou pragmática entre essas orações.

Assim sendo, o nível de integração pode ser medido através dos seguintes fatores: a simultaneidade modo-temporal dos eventos das orações, em que se o tempo verbal for o mesmo tanto na oração principal quanto na hipotática, haverá um maior nível de integração; a correferencialidade de sujeitos, na qual há uma maior integração caso o sujeito da oração principal e da oração hipotática for o mesmo; e, por fim, a codificação

formal do sujeito da oração hipotática, pois acredita-se que o grau de integração é maior quando este se realiza através de anáfora, sejam ela pronominal ou zero.

Portanto, para investigar o comportamento dessas orações e se alguma delas apresentaria um maior nível de integração do que as outras, serão apresentados nas próximas subseções os resultados pertinentes a cada um desses fatores.

3.2.1 Simultaneidade Modo-Temporal

É importante pontuar que existe uma diferença em como a língua italiana e a portuguesa classificam seus tempos e modos verbais. Enquanto em italiano ainda existe o modo Condizionale (Condicional) e seus tempos verbais, Presente e Passado, em português esse modo foi integrado ao Indicativo, constituindo, dessa forma, os tempos Futuro do Pretérito Simples e Composto, respectivamente.

Sendo assim, uma oração hipotática em italiano no Condizionale Presente poderia ser considerada simultânea a uma oração principal caso no Presente do Indicativo, por exemplo. Entretanto, o mesmo não poderia ocorrer conquanto à língua portuguesa, pois a oração hipotática, de acordo com a gramática desta língua, se encontraria no Futuro do Pretérito do Indicativo simultaneamente a segunda continuaria no Presente deste mesmo modo, não havendo deste modo uma equivalência de simultaneidade temporal.

Para solucionar essa diferença, optamos por realizar esta análise tendo em consideração um esquema de equivalência entre os tempos e modos verbais do português e do italiano. No quadro abaixo consideramos apenas os tempos e modos verbais em que foram encontrados os dados de nossa pesquisa:

Quadro 2: Correspondência de tempos e modos verbais entre Português e Italiano

Português	Italiano
Presente do Indicativo	Indicativo Presente
Pretérito Imperfeito do Indicativo	Indicativo Imperfetto
Pretérito Perfeito do Indicativo	Indicativo Passato Prossimo Indicativo Passato Remoto
Pretérito Mais-que-Perfeito do Indicativo	Indicativo Trapassato Prossimo
Futuro do Pretérito Simples do Indicativo	Condizionale Presente
Futuro do Pretérito Composto do Indicativo	Condizionale Passato
Futuro do Presente do Indicativo	Indicativo Futuro Semplice
Presente do Subjuntivo	Congiuntivo Presente

Pretérito Imperfeito Simples do Subjuntivo	Congiuntivo Imperfetto
Pretérito Mais-que-Perfeito Composto do Subjuntivo	Congiuntivo Trapassato
Imperativo Presente	Imperativo Presente

Fonte: Elaboração própria

Dessa forma, foram considerados para essa análise tanto o tempo quanto o modo verbal das orações em estudo, sendo avaliada, portanto, a simultaneidade modo-temporal entre as orações hipotáticas e suas orações principais. Logo, para que duas orações fossem consideradas simultâneas, tanto o tempo quanto o modo verbal da oração hipotática deveria ser equivalente ao encontrado em sua oração principal.

Vejamos a seguir dois exemplos em que os eventos são simultâneos, um em cada língua:

a) Eventos simultâneos

(18) “Do lado ambiental, a sustentabilidade do crescimento exige o manejo sustentável dos recursos naturais, **dado que esses recursos contribuem expressivamente para renda dos pobres.**” (*Corpus* do Português)

(19) “Dunque, **posto che per ogni bambino il confronto con i coetanei avveniva già attraverso l’esibizione del possesso e l’ostentazione del proprio status sociale**, la televisione si preoccupava di orientare i consumi verso obiettivi costruttivi e valori autentici.” (*Corpus* PAISÀ)

[Portanto, **já que para cada criança o confronto com os seus pares já acontecia através da exibição dos bens e da exibição do seu estatuto social**, a televisão se preocupava em orientar os consumos para objetivos construtivos e valores autênticos.]

Como podemos observar, nos exemplos acima os eventos que ocorrem em ambas as orações, hipotática e principal, ocorrem no mesmo instante. No exemplo (18), o verbo “exigir” da oração principal e o verbo “contribuir” da oração hipotática se encontram ambos no Presente do Indicativo. Da mesma forma, no exemplo (19), os verbos “avvenire” da oração hipotática e “preoccuparsi” da oração principal se encontram no tempo Imperfetto dell’Indicativo, o equivalente ao Pretérito Imperfeito do Indicativo na língua portuguesa.

Passamos agora a exemplificação em que isto não ocorre, ou seja, em que os tempos verbais e, por consequência, os eventos presentes nessas orações não ocorrem no mesmo momento.

b) Eventos não-simultâneos

(20) “**Dado que essas despesas correntes são essenciais para garantir o funcionamento adequado dos ativos de capital, a sua escassez resultará em serviços públicos ineficientes** e, no final das contas, um peso para o país, com duvidosos efeitos sobre o crescimento econômico”. (*Corpus do Português*)

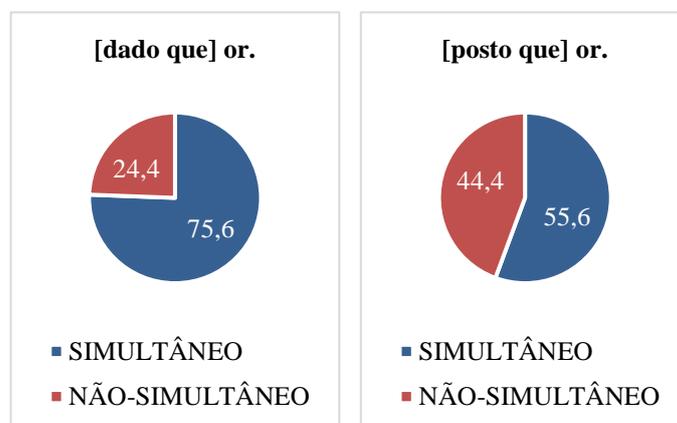
(21) “In tal caso la indicazione relativa all'ente è sufficiente ai fini degli adempimenti legislativi, posto che in tal modo sarà possibile risalire alla persona fisica cui sono destinati i precetti e le sanzioni”. (*Corpus PAISÀ*)

[Neste caso, a indicação relativa à entidade é suficiente para efeitos de obrigações legislativas, visto que desta forma será possível localizar a pessoa singular a quem se destinam os preceitos e sanções.]

No exemplo (20), enquanto o verbo “ser” da oração hipotática se apresenta no Presente do Indicativo, o verbo “resultar” da oração principal apresenta-se no Futuro do Presente do Indicativo, instanciando, portanto, diferentes momentos temporais. Da mesma forma, no exemplo (21), o verbo “ser” da oração principal também ocorre no Presente do Indicativo, simultaneamente o mesmo verbo ocorre no Futuro Semplice, equivalente ao Futuro do Presente do Indicativo, na oração hipotática.

No gráfico a seguir, estão indicados os valores percentuais da relação simultânea e não-simultânea nas orações do português:

Gráfico 9: Simultaneidade modo-temporal das orações hipotáticas em português

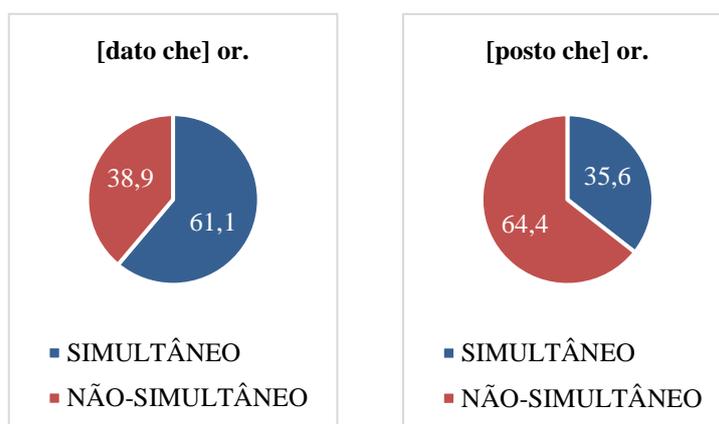


Fonte: Elaboração própria

A construção [dato que]or. costuma ser mais usada em contextos em que há simultaneidade, apresentando 75,6% dos dados nesta condição, em vista de apenas 24,4% dos dados ilustrarem eventos não-simultâneos. Da mesma forma, [posto que]or. também apresenta uma preferência pela simultaneidade, visto que 55,6% de suas ocorrências possuem orações hipotáticas e principais com o mesmo tempo e modo verbal, em detrimento da não-simultaneidade, que totalizou apenas 44,4% dos casos.

Passemos então ao cenário na língua italiana:

Gráfico 10: Simultaneidade modo-temporal das orações hipotáticas em italiano



Fonte: Elaboração própria

A microconstrução [dato che]or. também se encontra em cenários preferencialmente simultâneos, totalizando 61,1% de seus casos com uma equivalência modo-temporal entre os verbos das orações hipotáticas e suas principais, do que não-simultâneos, que somaram apenas 38,9% das ocorrências. Semelhantemente, [posto che]or. apresenta a mesma preferência por cenários simultâneos, havendo 64,4% nesta condição, enquanto foram contabilizados 35,6% de casos não-simultâneos.

Para continuar em busca dos aspectos integrantes aos conectivos neste estudo, dedicamos a seguinte subseção ao segundo fator: a codificação formal do sujeito da oração hipotática.

3.2.2 Codificação formal do sujeito da oração hipotática

O próximo fator a ser descrito foi a codificação formal do sujeito das orações hipotáticas. Como mencionado na seção de Pressupostos Teóricos, acredita-se que as orações hipotáticas que apresentam seus sujeitos através de uma anáfora pronominal ou uma anáfora zero possuem um maior grau de integração com suas orações principais

(Givón, 1979). Desta forma, nos dedicamos a identificar as possíveis formas que o sujeito da oração hipotática pode se apresentar, analisando o comportamento das duas línguas e averiguando qual das construções apresentaria um maior número de anáforas.

A partir disso, após a análise de nossos resultados, os sujeitos das orações em estudo se expressaram de cinco diferentes formas, como podemos observar nos exemplos abaixo, nos quais os sujeitos se encontram sublinhados:

a) Sujeito como Sintagma Nominal

(22) “Ao concluir o julgamento e depois de satisfazer os elementos da Seção 234 do Código, o Sr. Lallion foi condenado por homicídio. O Tribunal de Primeira Instância não tinha discricionariedade para impor uma sentença contra ele **posto que a pena é automática**, de acordo com a lei de Granada.” (*Corpus* do Português)

(23) “La distanza è un concetto fondamentale **dato che gli algoritmi di clustering raggruppano gli elementi a seconda della distanza** e quindi l'appartenenza o meno ad un insieme dipende da quanto l'elemento preso in esame è distante dall' insieme.” (*Corpus* PAISÀ)

[A distância é um conceito fundamental **dado que os algoritmos de clustering agrupam os elementos de acordo com a distância** e, portanto, o pertencimento ou não a um conjunto depende de quanto o elemento em análise está distante do conjunto.]

b) Sujeito como Sintagma Verbal

(24) “**Dado que apreender a realidade de uma maneira incorreta é a causa raiz das nossas dificuldades na vida**, iremos nos referir ao não-apercebimento neste contexto como “a confusão sobre a realidade”. (*Corpus* do Português)

(25) “A parte tutto, poveri Canino e Amadeus... non vorremmo proprio essere nei loro panni questa notte, che li vedrà atterriti dai tarli del flop in agguato e dall'incubo della soppressione in agguato (**posto che credere nei miracoli è ancora lecito**)”. (*Corpus* PAISÀ)

[Além de tudo, pobres Canino e Amadeus... nós realmente não gostaríamos de estar no lugar deles esta noite, que os verá aterrorizados pelos vermes do flop à espreita e pelo pesadelo da supressão à espreita (**já que acreditar em milagres ainda é lícito**).]

c) Sujeito como Anáfora Zero

(26) “Contudo, isto é complicado, **dado que não sabemos se a criança será feliz ou não**, nem fazemos ideia do quanto a criança poderia vir a ser capaz de superar as suas dificuldades.” (*Corpus* do Português)

(27) “Presto Robba cominciò a guadagnare una sua reputazione e fu premiato da commesse prestigiose da parte di ecclesiastici, aristocratici e borghesi. [...] Nel 1745 fu nominato “ingegnere di stato” della Carniola. Durante tutto questo tempo non perse mai contatto con Venezia, **posto che compì parecchie visite alla sua città natale.**” (*Corpus PAISÀ*)

[Logo Robba começou a ganhar uma reputação e foi premiado com ordens de prestígio de eclesiásticos, aristocratas e burgueses. [...] Em 1745 foi nomeado “engenheiro estadual” de Carniola. Durante todo este tempo não perdeu nunca contato com Veneza, **visto que fez diversas visitas à sua cidade natal.**]

d) Sujeito como Anáfora Pronominal

(28) “O termo “abandono afetivo” não representa corretamente o problema em análise, **posto que o situa apenas no âmbito subjetivo do afeto.** Melhor seria utilizar uma designação que trouxesse melhor a ideia do que há de objetivo na questão, como abandono de cuidado paterno-filial.” (*Corpus do Português*)

(29) “Purtroppo il premio non attirò molti matematici seri, **dato che questi erano ben consci dell'estrema difficoltà del problema** e quindi non produsse reali ricadute nel campo della matematica, ma ebbe il merito di rendere famoso al grande pubblico il problema dell'ultimo teorema di Fermat.” (*Corpus PAISÀ*)

[Infelizmente, o prêmio não atraiu muitos matemáticos sérios, **dado que estes eram bem conscientes da extrema dificuldade do problema** e, portanto, não produziram reais consequências no campo da matemática, mas teve o mérito de tornar famoso para o grande público o último problema do teorema de Fermat.]

e) Sujeito Impessoal

(30) “In quell'epoca l'offerta di premi per la soluzione di enigmi matematici era una pratica comune e molte accademie la seguivano, **dato che si poteva indirizzare la ricerca di molti scienziati in aree specifiche del sapere.**” (*Corpus PAISÀ*)

[Naquela época, a oferta de prêmios para a resolução de enigmas matemáticos era uma prática comum e muitas academias a seguiam, **dado que se poderia endereçar a pesquisa de muitos cientistas para áreas específicas do saber.**]

Nos exemplos (22) e (23), os sujeitos das orações hipotáticas se apresentam como sintagmas nominais, em que é um substantivo que constituem seus núcleos (“pena” e “algoritmi”, respectivamente). Enquanto isso, os sujeitos dos exemplos (24) e (25) possuem como seu núcleo um verbo (“apreender” e “credere”, respectivamente), sendo, portanto, codificados como sintagmas verbais.

Em sequência, nos exemplos (26), (27), (28) e (29) possuímos sujeitos que se materializam, ou não, através do fenômeno de anáfora. Nos exemplos (26) e (27) temos o caso de anáfora zero. No primeiro caso, o sujeito pode ser identificado através da desinência do verbo “sabemos”, em que se identifica, portanto, a primeira pessoa do plural e o sujeito oculto “nós”. Ao passo que o sujeito ocultado no exemplo (27) é contextual, visto que se trata da biografia de Francesco Robba, um famoso escultor italiano do período Barroco, mencionado ao longo do texto, sendo, portanto, altamente inferível.

Conquanto aos exemplos (28) e (29), o fenômeno da anáfora acontece de forma material, havendo, desta forma, um pronome explícito para retomada do sujeito. No primeiro exemplo, o pronome direto “o” retoma o termo “abandono afetivo”, expressado pela oração principal. Concomitantemente, no exemplo (29), o pronome demonstrativo “questi”, presente na oração hipotática italiana, resgata o sintagma “molti matematici seri”, expresso na oração principal.

Por fim, também foram encontrados alguns poucos casos de sujeitos impessoais, aparentes apenas na língua italiana. No exemplo (30), é demonstrado um desses casos, em que o sujeito é alguém ou um conjunto de pessoas genérico. Esse tipo de sujeito se materializa através da partícula “se” em conjunto com o verbo na terceira pessoa do singular e, em italiano, este fenômeno é conhecido como “si impersonale”. Neste exemplo, não é explícito e determinado quem seria a pessoa (ou as pessoas) que endereçaria(m) a pesquisa, sendo assim, portanto, uma oração com sujeito impessoal.

Tendo isto estabelecido, apresentamos os valores percentuais de cada um desses tipos de sujeito com cada conectivo do português:

Tabela 7: Codificação formal dos sujeitos das orações hipotáticas em português

Construção	Sintagma Nominal	Sintagma Verbal	Anáfora zero	Anáfora pronominal	Sujeito Imp.	Total
[dado que]or.	65 (72,2%)	1 (1,1%)	14 (15,6%)	10 (11,1%)	0	90 (100%)
[posto que]or.	45 (50%)	0	33 (36,7%)	12 (13,3%)	0	90 (100%)
Total	137	1	90	42	0	180

	(50,7%)	(0,4%)	(33,3%)	(15,6%)		(100%)
--	---------	--------	---------	---------	--	--------

Fonte: Elaboração própria

Na tabela acima, podemos constatar que os sujeitos de [dato que]or preferencialmente se apresentaram em forma de sintagma nominal, totalizando 72,2% dos casos, paralelamente apenas 15,6% dos dados apresentaram sujeitos em forma de anáfora zero e 11,1% como anáforas pronominais. Além disso, foi apenas com esta microconstrução que foi registrada uma ocorrência de sujeito como sintagma verbal, totalizando (1,1%) de seus casos.

Conquanto a [posto que]or, a forma como os sujeitos das orações hipotáticas se apresentam possui um diferente cenário, pois os casos de sintagmas nominais são menos expressivos que os da outra microconstrução, totalizando 50% dos casos, enquanto 36,7% dessas orações apresentaram sujeitos que se manifestam como anáforas zero e 13,3% como anáforas pronominais.

Por fim, também é importante destacar que, dos dados investigados nesta língua, não houve nenhuma ocorrência de sujeitos impessoais em ambas orações hipotáticas.

Então, passando para o cenário na língua italiana, a tabela abaixo indica os valores percentuais da codificação de sujeitos em relação a [dato che]or. e [posto che]or:

Tabela 8: Codificação formal dos sujeitos das orações hipotáticas em italiano

Construção	Sintagma Nominal	Sintagma Verbal	Anáfora zero	Anáfora pronominal	Sujeito Imp.	Total
[dato che]or.	53 (58,9%)	0	29 (32,2%)	6 (6,7%)	2 (2,2%)	90 (100%)
[posto che]or.	66 (73,3%)	2 (2,2%)	15 (16,7%)	6 (6,7%)	1 (1,1%)	90 (100%)
Total	119 (66,1%)	2 (1,1%)	44 (24,4%)	12 (6,7%)	3 (1,7%)	180 (100%)

Fonte: Elaboração própria

Na tabela 8, podemos verificar que os sujeitos de [dato che]or. possuem uma preferência pela forma de sintagma nominal, totalizando 58,8% de suas ocorrências. Em segundo lugar, o tipo de codificação mais frequente é a anáfora zero, representando 32,2% dos casos. Em números menos expressivos, ocorreram também casos de anáfora pronominal, que totalizam 6,7% dos casos, e sujeitos impessoais, que são 2,2% das ocorrências com esta microconstrução. Além disso, não foi encontrado nenhum caso de sujeito como sintagma verbal.

Em relação à [posto che]or, existe uma delimitação mais concreta sobre as preferências de codificação formal de seus sujeitos, já que 73,3% desses se apresentaram como sintagmas nominais. Em sequência, em números bem menores, houveram casos de anáfora zero, que totalizaram 16,7% dos casos, e também anáfora pronominal, por sua vez, sendo 6,7% dos casos, assim como na construção anterior. Ainda é importante destacar que esta microconstrução apresentou 2,2% de seus sujeitos como sintagma verbal, o que não havia ocorrido com [dato che]or. Por fim, foi encontrada também uma única ocorrência de sujeito impessoal, que contabiliza apenas 1,1% dos casos.

Tendo isso, estabelecido, passemos a subseção a seguir, na qual verificaremos em que medida o sujeito da oração hipotática se assemelha ao sujeito da oração principal.

3.2.3 Correferencialidade dos sujeitos da oração hipotática e da oração matriz

O último fator analisado em relação à integração entre as orações hipotáticas e suas orações principais é a correferencialidade entre os sujeitos de ambas as orações. É postulado que quando a oração hipotática e sua principal apresentam o mesmo sujeito, existe, portanto, uma maior integração sintática e semântico-pragmática entre elas (cf. Givón, 1979).

Assim sendo, apresentamos alguns dados que apresentaram essa correferencialidade que integra ambas as orações:

(31) “**Dado que não encontramos em nenhum documento anterior a fórmula completa da profissão de fé, não podemos assegurar que era idêntica ao nosso Credo.**”
(*Corpus* do Português)

(32) “Un agente per funzionare non necessita di una descrizione completa dell'ambiente anche perché una descrizione realmente esaustiva dell'ambiente è impossibile, dato che richiederebbe troppi dati da memorizzare e troppi sensori.”
(*Corpus* PAISÀ)

[Um agente para funcionar não necessita de uma descrição completa do ambiente também porque uma descrição realmente exaustiva do ambiente é impossível, dado que exigiria muitos dados a serem memorizados e muitos sensores.]

Tanto o exemplo (31) quanto o exemplo (32) apresentam sujeitos correferenciais que se expressam através do fenômeno da anáfora zero. No primeiro deles, temos a presença de um sujeito desinencial, que pode ser facilmente identificado através dos verbos “encontramos”, na oração hipotática, e “podemos assegurar”, na oração principal, evidenciando, portanto, que o sujeito destas orações se trata de “nós”, a primeira pessoa do singular. Enquanto isso, no exemplo (32), o sujeito “una descrizione realmente esaustiva dell’ambiente”, presente na oração principal, é omitido na oração hipotática, mas é facilmente recuperado através do contexto.

Sendo assim, demonstramos as tendências gerais da língua portuguesa relativas à correferencialidade na tabela abaixo:

Tabela 9: Correferencialidade entre os sujeitos das orações matrizes e hipotáticas em português

Construção	Correferente	Não- correferente	Total
[dado que]or.	9 (10%)	81 (90%)	90 (100%)
[posto que]or.	36 (40%)	54 (60%)	90 (100%)
Total	45 (25%)	135 (75%)	180 (100%)

Fonte: Elaboração própria

Na tabela 9, é possível perceber que a não-correferencialidade é mais expressiva com ambas as orações, em que 90% dos sujeitos que ocorrem com [dado que]or. não são correspondentes aos de sua oração principal. Da mesma forma, porém com um número menor de ocorrências, isso ocorre com [posto que]or, já que 60% dos dados também apresentam sujeitos não-correferentes. Vemos, com isso, que as orações iniciadas por dado que tendem a ser mais integradas à matriz do as orações com “posto que”.

A seguir, vejamos como esse fator se comporta na língua italiana:

Tabela 10: Correferencialidade entre os sujeitos das orações matrizes e hipotáticas em italiano

Construção	Correferente	Não- correferente	Total
[dato che]or.	31 (34,4%)	59 (65,6%)	90 (100%)
[posto che]or.	19 (21,1%)	71 (78,9%)	90 (100%)
Total	50 (27,8%)	130 (72,2%)	180 (100%)

Fonte: Elaboração própria

Na tabela acima, novamente há uma preferência pelo uso da não-correferencialidade, no entanto, esse fenômeno se apresenta de forma menos expressiva do que no português. Podemos perceber que [dato che]or. parece ter uma menor integração de relação intrasujeitos se comparado à outra microconstrução, já que apenas 65,6% dos casos são não-correferentes. Enquanto isso, [posto che]or. parece ser um pouco mais integrado, havendo 78,9% de seus sujeitos sendo correspondentes aos sujeitos de suas orações principais.

Nesta e nas subseções anteriores, observamos que alguns aspectos distintivos entre as microconstruções são de diversos tipos e níveis linguísticos, o que nos indica um quadro complexo do funcionamento das orações hipotáticas tanto no português brasileiro quanto no italiano. Há outros pontos que podem ser úteis para continuar a busca por uma característica distintiva, sendo eles o tipo e o número de ocorrências dos verbos que aparecem na oração hipotática, como veremos a seguir.

3.3 Frequência de tipo e de ocorrência dos itens verbais da oração hipotática

Os estudos de Bybee (2010) demonstram, por meio de diversos exemplos na língua inglesa, o caráter conservador e inovador da frequência de uso dentro de um sistema linguístico. De acordo com a autora, os itens existentes em uma língua são fortalecidos a cada vez que uma nova frequência de uso é mapeada.

Desta forma, exemplares que ocorrem com mais frequência são considerados mais fortes dentro da rede de construções do falante e, por consequência, tendem a ser mais

facilmente acessados pelos usuários dessa língua. Portanto, a frequência de uso é um fator muito importante, que tem o poder tanto de conservar formas linguísticas já existentes quanto de criar novos itens dentro de um sistema.

Para isso são levados em conta dois tipos de frequência: a de tipo e a de ocorrência. A frequência de tipo (*type*) diz respeito à variedade de itens lexicais que um padrão linguístico pode ter ou, no caso de construções mais esquemáticas, quais diferentes itens podem preencher os *slots* existentes. Enquanto isso, a frequência de ocorrência (*token*) corresponde ao número de vezes que um item em particular ocorre.

Considerando a importância da frequência, verificamos quais verbos aparecem em cada microconstrução e com qual frequência, iniciando pelo português:

Tabela 11: Frequência de tipo e de ocorrência dos itens verbais da oração hipotática em português

<i>Types verbais</i>	[dado que]or. <i>Tokens</i>	[posto que]or. <i>Tokens</i>	Total de <i>tokens</i>
1. Ser	23	27	50
2. Estar	7	2	9
3. Ter	6	3	9
4. Existir	4	0	4
5. Envolver	2	1	3
6. Representar	1	2	3
7. Saber	2	1	3
8. Trabalhar	2	1	3
9. Ver	2	1	3
10. Chamar	0	2	2

Total de itens verbais da tabela	Total de itens verbais com “dado que” na tabela	Total de itens verbais com “posto que” na tabela	Total de ocorrências na tabela
10	9	9	89
Outros itens verbais	Outros itens verbais com “dado que”	Outros itens verbais com “posto que”	Outras ocorrências
84	40	48	91
Total de itens verbais	Total de itens verbais com “dado que”	Total de itens verbais com “posto que”	Total de ocorrências
94	49	57	180

Fonte: Elaboração própria

Na tabela 11, apresentamos apenas os 10 primeiros verbos mais frequentes considerando as duas microconstruções em português. Vale a pena apontar que se apresentaram 94 verbos distintos que se distribuíram de forma desigual em 180 ocorrências totais. Em específico, foram identificados 49 diferentes itens verbais com [dado que]or e 57 com [posto que]or.

Ainda que os verbos não se apresentem na mesma quantidade de vezes nos dados (frequência de ocorrência), a variedade de verbos (frequência de tipo) parece ser bem equilibrada entre as duas microconstruções em análise, mesmo com uma leve predominância quantitativa de itens verbais da oração hipotática introduzida por [posto que], em razão de esta oração possuir 8 itens verbais a mais que [dado que]or.

Ainda mais especificamente, investigamos que 11 itens verbais são compartilhados por estas duas microconstruções. Sendo assim, podemos constatar que 37 itens verbais acontecem exclusivamente apenas com [dado que]or e 45 com [posto que]or.

Em relação à frequência de ocorrência, a microconstrução [dado que]or. foi a que apresentou o maior índice quantitativo na tabela, com 49 ocorrências, enquanto [posto que]or teve 40 ocorrências. A partir disso, também foram consideradas as ocorrências dos outros itens verbais que não aparecem nesta tabela. Dessa forma, podemos inferir que existem outras 41 ocorrências com [dado que]or e 50 ocorrências com [posto que]or. Além disso, existe uma predominância significativa do verbo SER nas duas estruturas, havendo 23 ocorrências em [dado que]or e 27 em [posto que]or.

Após o verbo primeiro colocado em frequência de ocorrência, as microconstruções se assemelham um pouco quanto a ordenação, sendo SER, ESTAR, TER, EXISTIR, ENVOLVER, SABER, TRABALHAR, VER e REPRESENTAR para [dato que]or; e SER, TER, ESTAR, REPRESENTAR, CHAMAR, ENVOLVER, SABER, TRABALHAR e VER para [posto que]or.

Passemos então a observar como essas frequências se comportam com as microconstruções [dato che]or e [posto che]or:

Tabela 12: Frequência de tipo e de ocorrência dos itens verbais da oração hipotática em italiano

<i>Types verbais</i>	[dato che] or. <i>Tokens</i>	[posto che]or. <i>Tokens</i>	Total de <i>tokens</i>
1. Essere (Ser/Estar)	26	32	58
2. Avere (Ter)	2	2	4
3. Fare (Fazer)	2	2	4
4. Utilizzare (Utilizar)	4	0	4
5. Accedere (Acessar)	2	0	2
6. Assicurare (Assegurar)	1	1	2
7. Avvenire (Acontecer)	1	1	2
8. Consentire (Consentir)	2	0	2
9. Considerare (Considerar)	1	1	2
10. Creare (Criar)	2	0	2
Total de itens verbais da tabela 10	Total de itens verbais com “dato che” na tabela 10	Total de itens verbais com “posto che” na tabela 6	Total de ocorrências na tabela 82

Outros itens verbais 88	Outros itens verbais com “dato che” 45	Outros itens verbais com “posto che” 48	Outras ocorrências 98
Total de itens verbais 98	Total de itens verbais com “dato que” 55	Total de itens verbais com “posto che” 54	Total de ocorrências 180

Fonte: Elaboração própria

Na tabela acima, da mesma forma, expomos apenas os 10 primeiros verbos mais frequentes considerando as duas microconstruções em italiano. Neste caso, foram encontrados 98 distintos itens verbais no total, havendo 55 diferentes tipos com [dato che]or e 54 com [posto che]or.

Ao realizar uma nova filtragem, constatamos que existem 11 verbos em comum entre [dato che]or e [posto che]or, havendo, por consequência, 44 itens verbais que ocorrem apenas com [dato che]or e 43 com [posto che]or. Ainda assim, isto não se mostra percentualmente relevante para uma distinção entre a frequência de tipo entre essas duas microconstruções.

Conquanto à frequência de ocorrência, a microconstrução [dato che]or. apresentou o maior índice quantitativo na tabela, com 43 ocorrências, enquanto [posto che]or. teve 39 ocorrências. A partir disso, também foram consideradas as ocorrências dos outros itens verbais que não aparecem nesta tabela. Dessa forma, podemos inferir que existem outras 47 ocorrências com [dato che]or e 51 ocorrências com [posto che]or.

Assim como na língua portuguesa, existe um alto número de ocorrências do verbo ESSERE nas duas estruturas, havendo 26 ocorrências em [dato che]or e 32 em [posto che]or. Porém, é importante destacar que [dato che]or. ocorre com todos os dez verbos listados na tabela, em contrapartida [posto che]or. aparece com apenas 6 desses itens verbais.

Após o verbo primeiro colocado em frequência de ocorrência, as microconstruções se assemelham um pouco quanto a ordenação, sendo ESSERE, UTILIZZARE, AVERE, FARE, ACCEDERE, CONSENTIRE, CREARE,

ASSICURARE, AVVENIRE e CONSIDERARE para [dato che]or; e ESSERE, AVERE, FARE, ASSICURARE, AVVENIRE e CONSIDERARE para [posto que]or.

Na seção a seguir, retomaremos as principais tendências dessas quatro microconstruções em estudo, realizando uma revisitação aos nossos objetivos e hipóteses, além de uma comparação final entre ambas as línguas.

4 REFLEXÕES SOBRE A ANÁLISE COMPARATIVA: CONSIDERAÇÕES FINAIS

Nossa pesquisa tinha como objetivo geral mapear o comportamento das quatro microconstruções em estudo, tanto intralinguisticamente, comparando os conectivos em variação dentro do sistema linguístico de cada idioma, quanto entre português e italiano, visando evidenciar as principais semelhanças e diferenças entre os conectivos.

Sendo assim, para um comparativo final, elaboramos a seguinte tabela, que demonstra a sintetização dos resultados encontrados na análise:

Tabela 13: Comparação entre os usos das orações hipotáticas em português e italiano

Fatores Analisados		[dado que]or.	[posto que]or.	[dato che]or.	[posto che]or.
Valor Semântico	Causal	94,4%	93,3%	96,7%	67,8%
	Causal-Cond.	5,6%	4,4%	3,3%	0
	Condicional	0	0	0	32,2%
	Concessivo	0	2,3%	0	0
Ordenação	Anteposta	48,9%	11,1%	7,8%	28,9%
	Intercalada	0	0	0	3,3%
	Posposta	51,1%	88,9%	92,2%	67,8%
Ordenação x Valor Semântico	Anteposta Causal	45,8%	8,3%	5,8%	29,5%
	Intercalada Causal	0	0	0	1,7%
	Posposta Causal	54,2%	91,7%	94,2%	68,8%
	Anteposta Causal-Cond.	100%	25%	66,7%	0
	Intercalada Causal-Cond.	0	0	0	0
	Posposta Causal-Cond.	0	75%	33,3%	0

Ordenação x Valor Semântico	Anteposta Condicional	0	0	0	27,5%
	Intercalada Condicional	0	0	0	6,9%
	Posposta Condicional	0	0	0	65,6%
	Anteposta Concessiva	0	100%	0	0
	Intercalada Concessiva	0	0	0	0
	Posposta Concessiva	0	0	0	0
Pressuposição	Pressuposta	20%	8,9%	6,7%	12,2%
	Não Pressuposta	80%	91,1%	93,3%	87,8%
Ordenação x Pressuposição	Anteposta Pressuposta	88,9%	12,5%	16,7%	27,3%
	Intercalada Pressuposta	0	0	0	9,1%
	Posposta Pressuposta	11,1%	87,5%	83,3%	63,6%
	Anteposta N. Pressuposta	38,9%	10,9%	7,1%	29,1%
	Intercalada N. Pressuposta	0	0	0	2,5%
	Posposta N. Pressuposta	61,1%	89,1%	92,9%	68,4%

Simultaneidade	Simultâneo	75,6%	55,6%	61,1%	35,6%
Modo-Temporal	N. Simultâneo	24,4%	44,4%	38,9%	64,4%
Codificação Formal do Sujeito	S. Nominal	72,2%	50%	58,9%	73,3%
	S. Verbal	1,1%	0	0	2,2%
	Anáfora Zero	15,6%	36,7%	32,2%	16,7%
	Anáfora Pron.	11,1%	13,3%	6,7%	6,7%
	Sujeito Imp.	0	0	2,2%	1,1%
Correferencialidade entre Sujeitos	Correferente	10%	40%	34,4%	21,1%
	N. Correferente	90%	60%	65,6%	78,9%
Itens verbais mais frequentes		Ser, estar, ter	Ser, ter, estar	Essere, utilizzare, avere	Essere, avere, fare
Frequência type		49	57	55	54

Fonte: Elaboração própria

Em relação ao valor semântico, tínhamos como objetivo analisar semanticamente as orações, descrevendo as relações instauradas por cada uma das construções. A partir da revisão bibliográfica realizada, tínhamos a hipótese de que o conectivo “dado que” poderia apresentar os valores condicional, de acordo com as gramáticas tradicionais, ou causal, de acordo com os estudos de Neves (1999), Batista (2013) e Nascimento (2022). Já “posto que”, em relação às gramáticas tradicionais, poderia ser encontrado com o valor concessivo, ou ainda com os valores causal ou causal-condicional, segundo a pesquisa realizada por Batista (2013) e Nascimento (2022), entre outros autores citados no primeiro capítulo desta dissertação.

Conquanto à língua italiana, baseando-se na gramática de Dardano e Trifone (1999) e na enciclopédia Treccani, tinha-se a hipótese de que o valor encontrado com a microconstrução iniciada por “dato che” seria o causal, enquanto “posto che” seria encontrado com o valor condicional.

No tocante aos nossos resultados, alguns desses conectivos demonstraram mais de um valor adverbial possível ou ainda um valor semântico não previsto pelas gramáticas consultadas. Entretanto, de forma geral, o valor causal foi aquele que apareceu com maior frequência em todos os conectivos analisados, constituindo 94,4% dos casos de orações iniciadas por “dado que”, 93,3% com “posto que”, 96,7% com “dato che” e 67,8% com “posto che”. Sendo assim, é importante destacar que, apesar de ter um quantitativo maior de orações com o valor semântico causal, [posto che]or. apresentou um número de ocorrências causais consideravelmente menor, se comparado às demais construções.

No que se refere ao valor semântico condicional, foram encontrados apenas alguns casos com a microconstrução [posto che]or., como era previsto pela tradição italiana, constituindo 32,2% dos casos. Ainda assim, este pode ser considerado um valor abaixo do que esperávamos, considerando que este seria o valor prototípico dessa microconstrução de acordo com a bibliografia consultada. Para além disso, cabe ressaltar que, apesar do valor pleno não ter sido encontrado com as demais microconstruções, houve alguns casos de valor híbrido, mais especificamente do valor causal-condicional, que ocorreram com as outras três orações em estudo. Este valor apareceu em 5,6% dos casos de [dado que]or., 4,4% de [posto que]or. e 3,3% de [dato che]or.

Por fim, no que diz respeito ao valor concessivo, apenas [posto que]or. apresentou algumas poucas ocorrências correspondentes a este sentido, em que apenas 2,3% dos dados demonstraram esta circunstância, diferindo-se, portanto, daquilo esperado pelas gramáticas tradicionais, que postulam que este conectivo se manifestaria com o valor de concessão.

Desta forma, podemos fazer algumas considerações acerca do sentido que estas microconstruções assumem em ambas as línguas. Enquanto em português existe uma diferença ínfima entre [dado que]or. e [posto que]or., já que ambas microconstruções apresentam um grande quantitativo que tende à causalidade, alcançando mais de 90% de seus casos, em italiano o cenário é diferente. Isto porque existe um número considerável de ocorrências de [posto che]or. com o valor condicional, o que não ocorre com a microconstrução [dato che]or., que, assim como as microconstruções em língua portuguesa, possui a maior parte de seus dados como orações causais.

Além disso, por consequência, as microconstruções [posto que]or. e [posto che]or. se diferem, já que parece haver particularidades de uso no sistema linguístico de cada idioma, pois ao mesmo tempo que uma se aproxima quase inteiramente do polo das causais, a outra microconstrução apresenta um número significativo de casos pertencentes

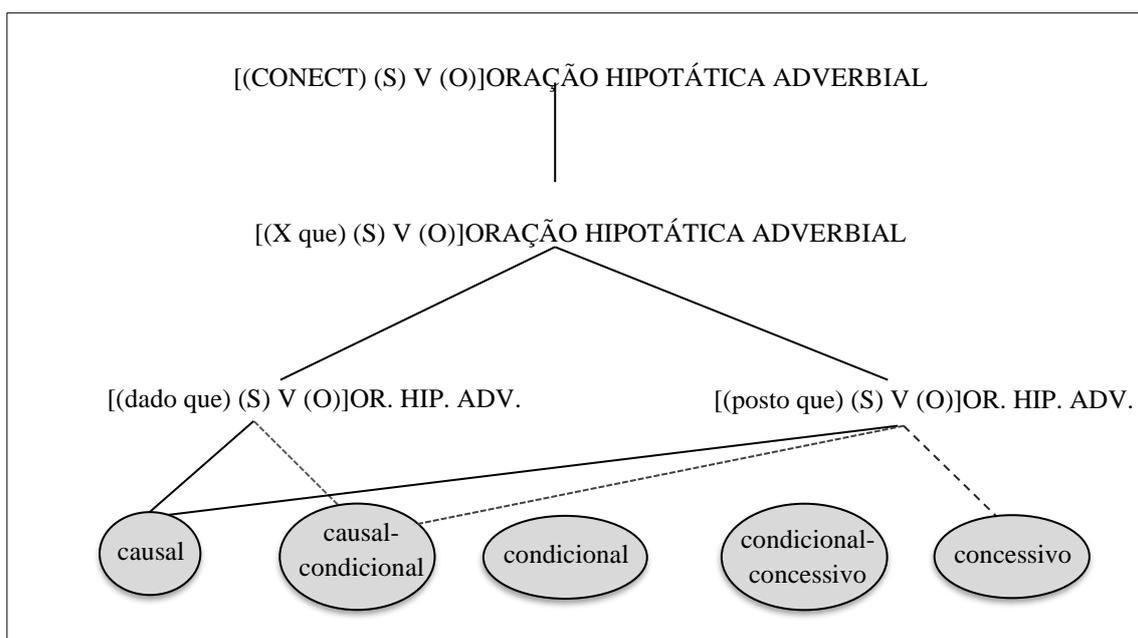
ao polo das condicionais. Ademais, cabe destacar a presença do valor concessivo em português, que, por sua vez, não ocorre na língua italiana.

Este cenário permite pressupor que [posto que]or. e [posto che]or. se encontram em diferentes momentos de mapeamento semântico: [posto que]or. parece mais consolidado na língua portuguesa, apresentando pouca variação em seu sentido, já [posto che]or. parece possuir mais mobilidade semântica na língua italiana.

Diferentemente, este não parece ser o caso das outras duas microconstruções. Tanto em português quanto em italiano, foram encontrados os mesmos valores semânticos com [dado que]or. e [dato che]or. Ambas aparecem com mais de 90% de seus casos com o valor causal enquanto as de valor causal-condicional possuem um índice percentual abaixo de 6%.

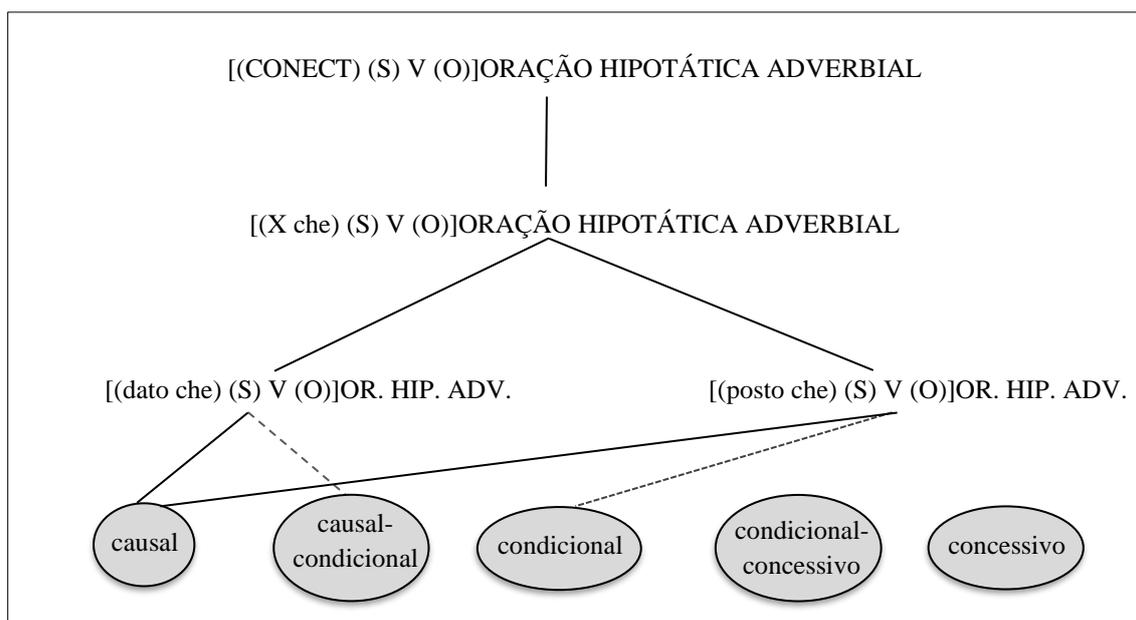
A seguir, elencamos esquemas que buscam demonstrar como essas construções se organizam acerca dos valores semânticos encontrados com essas línguas:

Figura 8: Rede semântica dos conectivos em português



Fonte: Elaboração própria

Figura 9: Rede semântica dos conectivos em italiano



Fonte: Elaboração própria

Apresentamos, acima, as redes semânticas das orações hipotáticas em português e em italiano. A partir do esquema mais geral e abstrato [(CONNECT) (S) V (C)]or. é instanciado o subesquema das hipotáticas adverbiais introduzidas pelo conectivo [X que], em português, e [X che], em italiano. Estes, por sua vez, instanciam as microconstruções esquemáticas estudadas nessa pesquisa: [dado que]or. e [posto que]or. em língua portuguesa e [dato che]or. e [posto che]or. em língua italiana.

Todas essas microconstruções, como explicitado nesta pesquisa, se relacionam com o campo semântico da causalidade, seja ela afirmada (orações causais), hipotetizada (orações condicionais) ou negada (orações concessivas), conforme os estudos realizados por Neves (1999). Porém, o elo entre essas microconstruções e esses valores semânticos não acontece da mesma forma.

No caso do português, [dado que]or. se relaciona fortemente com o domínio da causalidade, já que, na maioria dos dados, esta microconstrução aparece com este valor semântico. De maneira menos atenuada, como evidenciada pela linha pontilhada, [dado que]or. também se vincula ao valor causal-condicional. Da mesma forma, isto ocorre com [posto que]or., que também possui um elo forte com a causalidade e um *link* mais fraco com o eixo das causal-condicionais. Entretanto, além disso, essa microconstrução apresenta uma relação com o domínio da concessão, sendo o elo mais fraco devido ao baixo número de ocorrências.

Conquanto ao italiano, o cenário de [dato che]or. é extremamente semelhante à sua correspondente em língua portuguesa, havendo também um forte elo com o valor causal e um elo mais enfraquecido com o eixo das orações causal-condicionais. Todavia, [posto che]or. possui uma característica que a distingue de seu par em português. Apesar de também haver um elo forte com o valor causal, esta microconstrução apresenta um vínculo com o domínio da condicionalidade, havendo um número significativo de ocorrências com este valor semântico.

Outro de nossos objetivos era observar as preferências colocacionais de cada construção, considerando a ordem das orações hipotáticas em relação às orações matrizes. Em relação a este fator, todas as microconstruções apresentaram uma preferência pela posição posposta. Entretanto, a distribuição desta condição se dá de forma diferente entre as microconstruções. Enquanto com [posto que]or. e [dato che]or. essa condição ultrapassa ou se aproxima de 90% dos casos, com [dado que]or. foram contabilizadas apenas 51,1% das ocorrências nesta posição, ao passo que em relação à [posto che]or., apenas 67,8% dos casos ocorrem após suas orações matrizes.

Sendo assim, [dado que]or., em português, possui certa mobilidade de posição, apresentando resultados com valores percentuais quase equivalentes em relação à quantidade de antepostas (48,9%) e pospostas (51,1%), [dato che]or. possui uma clara preferência pela posposição (92,2%). Da mesma forma, [posto que]or. possui uma nítida preferência pela posição posposta (88,9%), em contrapartida o número com [posto che]or. é significativamente menor, havendo 67,8% de casos ocorrendo nesta posição.

Além disso, é cabível destacar que, dentre todas as microconstruções em estudo, apenas [posto che]or. apresentou uma outra possibilidade de ordenação, além da anteposição e posposição: a posição intercalada, que ocorre entre o sujeito e o predicado da oração principal, representando 3,3% de seus casos.

Como abordado no capítulo que trata dos pressupostos teóricos utilizados para esta pesquisa, existe uma relação estreita entre ordem e vínculo semântico. Portanto, um de nossos objetivos foi analisar a relação entre esses dois fatores, tendo como base os estudos de Diessel (2013) e Neves (1999). Nossa hipótese, guiada por essas pesquisas, era de que as orações causais geralmente se posicionariam após a sua oração principal, ao passo que orações condicionais tenderiam a preceder à oração principal. Por sua vez, a ordenação de orações concessivas seria ambígua, pois depende do propósito comunicativo do interlocutor.

Em relação às orações causais, apesar de todas as quatro microconstruções corroborarem com a hipótese de que a causalidade estaria ligada à posposição, visto que, canonicamente, se anuncia primeiro o efeito para depois se anunciar a causa, houveram algumas diferenças percentuais significativas. Enquanto [posto que]or. e [dato che]or. apresentaram mais de 90% de seus dados causais tendo como preferência a posição posposta, em [dado que]or. a diferença foi bem menos significativa, em virtude de apenas 54,2% de suas ocorrências aparecerem após suas orações principais. De forma um pouco menos significativa, isso se deu também com [posto che]or., que apresentou 68,8% de seus dados causais pospostos.

Portanto, ao mesmo tempo que em português [dado que]or. parece ter uma maior flexibilidade colocacional, havendo números percentuais de anteposição e posposição muito próximos, em italiano, com [dato che]or., existe uma evidente preferência pela posposição. De forma semelhante, em português [posto que]or. aparece majoritariamente após suas orações principais, ao passo que, em italiano, em relação à [posto che]or., um número considerável dessas orações ocorre, na verdade, antes de suas orações principais, o que pode ser outro indicativo que, apesar das semelhanças, essas construções variam de uma língua para outra.

Outro valor encontrado foram as orações causal-condicionais, até então não previstas em nossas hipóteses. Considerando o que a tradição postula sobre esses dois valores semânticos, a causalidade tenderia à posposição, em oposição a condicionalidade, por sua vez, tenderia à anteposição. Não havendo, portanto, uma hipótese concreta sobre qual colocação seria a preferida para este valor semântico, os resultados acerca da mescla desses dois sentidos ocorreram de maneira distinta para as microconstruções em estudo.

Em relação à [dado que]or. todos os casos encontrados com este valor se apresentaram na posição anteposta. Da mesma forma, a microconstrução correspondente em italiano, [dato che]or., apresentou mais orações antepostas com este valor causal, porém, em um percentual menos significativo, totalizando 66,7% de suas ocorrências. Já [posto que]or. não apresentou nenhuma ocorrência com este sentido, em contrapartida [posto che]or., obteve um maior número de orações causal-condicionais na posição posposta, totalizando 75% de seus casos e diferenciando-se das outras microconstruções supracitadas.

Como anteriormente mencionado, estudos linguísticos precedentes postulam que a posição preferida das orações condicionais é anteposição, na qual primeiro se enuncia a ocorrência de uma condição, expressa pela oração hipotática, que pode ou não ser

satisfeita, para depois ser enunciada a oração que depende da concretização desta condição, expressa pela oração principal.

Apenas [posto che]or. apresentou alguns casos de orações plenamente condicionais. Neste caso, não houve uma corroboração de nossa hipótese, em função de essas orações aparecerem preferencialmente pospostas às suas orações principais. Porém, apesar da posposição ter sido a colocação preferida das orações com esse valor semântico, totalizando 65,6% dos dados, houve uma quantidade significativa de orações antepostas, contabilizando 27,5% dos casos. Além disso, este sentido ainda apresentou uma outra possibilidade de ordenação, apresentando 6,9% de seus casos na posição intercalada.

Em relação à concessão, apenas a [posto que]or. apresentou casos com este valor semântico. Sobre este sentido específico, não havia uma hipótese específica, pois as orações concessivas teriam um posicionamento linear ambíguo, a depender de seu propósito comunicativo e também seu grau de informatividade, como será melhor abordado. Em nossos resultados, não houve uma variação colocacional, uma vez que 100% das ocorrências se apresentaram na posição anteposta. Neste caso, primeiro se refuta uma possível ou previsível objeção do interlocutor, codificada pela oração concessiva, para só depois se fazer uma asseveração, expressada pela oração principal (cf. Neves, 1999).

Outro de nossos objetivos era observar a informatividade das orações hipotáticas, mapeando se tais construções veiculam (ou não) informações pressupostas pragmaticamente. Conquanto a este fator, as orações não evidenciaram muitas diferenças entre si, em razão que todas apresentaram um valor percentual maior de orações não-pressupostas pragmaticamente. Entretanto, é passível de destaque que, diferentemente de [posto que]or. e [dato che]or., que contabilizaram mais de 90% de suas ocorrências como não-pressupostas, em que as informações trazidas pelas orações hipotáticas não haviam sido expostas anteriormente no texto, o cenário é ligeiramente diferente em [dado que]or. e [posto che]or. Embora essas duas microconstruções apresentem mais de 80% de seus casos como orações não-pressupostas, existe um número de orações considerável de orações pressupostas, 20% com [dado que]or. e 12,2% com [posto che]or.

Sendo assim, ao comparar as microconstruções correspondentes em cada língua, [dado que]or. possui um número de orações pressupostas triplamente maior que [dato che]or. Enquanto isso, a diferença percentual entre [posto que]or. e [posto che]or. é menos latente, sendo 8,9% de casos pressupostos com a primeira e 12,2% com a segunda.

Ainda com o objetivo de analisar a relação entre vínculo semântico, ordem e pressuposição das orações hipotáticas (cf. Diessel, 2013; Neves, 1999), realizamos também um cruzamento entre o grau de informatividade das orações e sua ordenação. Tínhamos como hipótese que orações adverbiais antepostas teriam a função de apresentar informações que são pragmaticamente pressupostas, assim como que orações pospostas tenderiam a apresentar informações novas, isto é, não-pressupostas. Em relação às orações intercaladas, pressupomos que, por normalmente assumirem uma função pragmática de adendo, essas orações também apresentariam informações que seriam novidade para o leitor.

Em relação à pressuposição, nossa hipótese foi corroborada apenas por [dado que]or., pois 88,9% de seus casos apareceram na posição anteposta. Ainda assim, é passível de destaque que [posto che]or. apresentou um número considerável de orações pressupostas antepostas, havendo 27,3% de seus casos nesta condição. Ao passo que [posto que]or. e [dato che]or. apresentaram mais de 80% de seus dados pressupostos na posição posposta.

Em contraponto, todas as microconstruções corroboram a hipótese de que orações não-pressupostas tenderiam à posposição. Entretanto, é importante destacar que, tanto [dado que]or. quanto [posto che]or. apresentaram um número considerável de orações antepostas que não poderiam ser previstas pelo interlocutor, havendo, respectivamente, 38,9% e 29,1% de seus casos nesta condição.

Portanto, no mesmo momento em que as orações pressupostas com [dado que]or. tendem à anteposição, aquelas com [dato che]or. tendem a posposição. Da mesma forma, as orações pressupostas com [posto que]or. aparecem mais na posição posposta, em contrapartida, o número em relação à [posto che]or. é significativamente menor. Da mesma forma, essa dinâmica se repete a respeito das orações não-pressupostas. Sendo assim, pode-se salientar que os pares correspondentes possuem diferentes usos em cada sistema linguístico.

Como foi evidenciado, o fator de ordenação tem um grande grau de importância para essa pesquisa. Em alguns casos, o comportamento das microconstruções se altera completamente a depender da posição linear que ela ocupa. Sendo assim, em pesquisas futuras, seria interessante observar e discutir se as orações [V_{pp} que] aqui estudadas constituem diferentes nós linguísticos dentro da rede de construções dos usuários dessas línguas quando precedem, sucedem ou perpassam sua oração principal; além de analisar

em que nível elas se diferem tendo como ponto de partida a sua ordenação linear e, a partir disso, propor uma rede taxonômica que evidencie essa possível diferença.

Mais um de nossos objetivos era medir o grau de integração entre as orações hipotáticas e suas matrizes, identificando algumas de suas preferências sintáticas, como a codificação formal do sujeito, a correferencialidade entre os sujeitos de ambas as orações (hipotática e matriz) e a simultaneidade modo-temporal. Nossa hipótese era que se o tempo e o modo verbal forem o mesmo tanto na oração principal quanto na hipotática, haverá um maior nível de integração, assim como se houvesse correferencialidade entre os sujeitos da oração hipotática e da principal. Além disso, de acordo com essa proposta, acreditava-se que o grau de integração era maior quando o sujeito da oração se realizava através de anáfora, seja ela pronominal ou zero.

Em relação à simultaneidade modo-temporal, [dado que]or. foi a que obteve um maior número de orações hipotáticas que ocorrem no mesmo tempo e modo que suas orações principais, totalizando 75,6% de seus casos. Em segundo lugar, [dato che]or. possui 61,1% de suas orações hipotáticas e principais com tempo e modo compatíveis. Passando a terceira posição, os valores entre simultâneos e não-simultâneos se aproximam bastante, visto que 55,6% dos dados investigados com [posto que]or. apresentam uma correspondência entre ambas as orações. Enquanto isso, em número significativamente menor, [posto che]or. apresenta apenas 35,6% de simultaneidade modo-temporal com suas orações principais.

Em relação ao segundo fator de integração, consideramos a codificação formal do sujeito. Como dito anteriormente, quanto maior o número de anáforas, mais integração haveria entre a oração hipotática e sua principal. Sendo assim, concentrando-se apenas nas anáforas encontradas em nossos dados, tanto pronominais quanto zero, [posto que]or. foi o que apresentou um maior número de ocorrências. Somando o número de anáforas pronominais (13,3%) e zero (36,7%), se constata que 50% de seus sujeitos são codificados como anáforas.

Em segundo lugar, tem-se [dato che]or., contabilizando 38,9% de seus sujeitos como anáforas, 6,7% como anáforas pronominais e 32,2% como anáforas zero. Em sequência, [dado que]or. apresenta 26,7% de seus sujeitos como anáforas, 15,6% como anáforas pronominais e 11,1% como anáforas zero. Por fim, em último lugar, tem-se [posto che]or., que soma 23,4% de seus sujeitos com esta codificação, havendo 6,7% como anáforas pronominais e 16,7% como anáforas zero.

Se concentrando então no último fator de integração entre orações, temos a correferencialidade entre os sujeitos da hipotática e sua principal. Acredita-se que, quando os sujeitos de ambas as orações são os mesmos, tem-se, por consequência, uma maior integração entre essas orações. A microconstrução que apresentou um maior número de sujeitos correferentes foi [posto que]or., havendo 40% de suas ocorrências com sujeitos correspondentes aos de suas orações principais. Em sequência, [dato che]or. contabiliza 34,4% de seus casos como sujeitos correferentes. Em terceiro lugar, segue-se [posto che]or., com 21,1% de seus dados tendo o mesmo sujeito que a oração principal. E, por fim, em número percentual muito menor, [dado que]or. apresenta apenas 10% de sujeitos correferentes.

Tendo esses resultados em consideração, elaboramos abaixo um quadro que sintetiza, em ordem de relevância percentual, como essas microconstruções se organizam acerca dos fatores de integração:

Quadro 3: Fatores de integração entre orações com as construções em análise

Posição	+Simultaneidade modo-temporal	+Sujeitos anafóricos	+Sujeitos correferentes
1º	[dado que]or.	[posto que]or.	[posto que]or.
2º	[dato che]or.	[dato che]or.	[dato che]or.
3º	[posto que]or.	[dado que]or.	[posto che]or.
4º	[posto che]or.	[posto che]or.	[dado que]or.

Fonte: Elaboração própria

Ao levar em conta os três fatores de integração, [posto que]or. foi aquele que apresentou, conjuntamente, o maior grau de integração entre a oração hipotática e a oração principal, seguido por [dato che]or., [dado que]or. e [posto che]or. Isso permite também observar que existe uma diferença de uso entre as línguas ao comparar [posto que]or. com sua correspondente em italiano, já que [posto che]or., diferentemente, apresentou um menor grau de integração dentre todas as microconstruções.

Por fim, nosso último objetivo era analisar a produtividade dos itens verbais, a partir da proposta de Bybee (2010), verificando a frequência que eles aparecem com cada

construção. Em relação aos itens verbais mais frequentes, foi observado, que, apesar de haver uma diferença quantitativa entre eles, os verbos utilizados por essas microconstruções não variam muito. Tanto em língua portuguesa quanto em língua italiana, os verbos “ser”, “estar” e “ter”, e seus correspondentes, “essere” (que possui dois usos principais em língua italiana, “ser” e “estar”) e “avere”, aparecem entre os mais utilizados nessas orações.

Ainda assim, o uso de verbos como “utilizzare” (utilizar) e “fare” (fazer), com [dato che]or. e [posto che]or. respectivamente, ambos os itens completamente fora do escopo semântico dos demais verbos encontrados, pode incitar que existe sim uma diferença de uso entre as duas línguas.

Porém, devido a isso, cabe também destacar que a representação SVO não parece ser a prototípica de nenhuma dessas microconstruções, já que a maioria dos itens verbais que preenchem o *slot* V da oração hipotática são verbos de ligação, como tinha sido previamente hipotetizado com base no trabalho anteriormente realizado por Nascimento (2022). Portanto, um futuro estudo classificando semanticamente os itens verbais que preenchem esse *slot* pode ser uma outra forma de diferenciar e mapear os usos dessas microconstruções em competição.

Em relação à análise da frequência *type* pudemos observar que [posto que]or. possui um maior nível de produtividade, já que apresentou 57 diferentes itens verbais. Em sequência, [dato che]or. aparece com 55 itens verbais, [posto che]or. com 54 itens verbais e, por último e em menor quantidade, [dado que]or. com 49 itens verbais.

A importância destas considerações é que exemplares que ocorrem com mais frequência são considerados mais fortes dentro da rede de construções do falante e, por consequência, tendem a ser mais facilmente acessados pelos usuários dessa língua. Sendo assim, se uma microconstrução apresenta uma variabilidade maior de preenchimento de seus *slots*, neste caso, do *slot* V, consequentemente, ela será mais facilmente acessada e utilizada pelos usuários daquele sistema linguístico.

Ainda é importante destacar que as microconstruções estudadas apresentam diferentes graus de produtividade, pois no contexto real de uso é possível que o falante priorize determinada microconstrução em detrimento de outra devido ao exemplar mais forte na memória. No caso desta pesquisa, como foi delimitado um número específico e igual para todas as microconstruções, esta tarefa não se torna viável. Sendo assim, em investigações posteriores, seria interessante a busca desses conectivos em *corpora* mais

bem delimitados, medindo, de maneira mais eficaz, a produtividade e a competição entre essas construções.

Pesquisas futuras podem incluir outras construções oracionais hipotáticas com outros conectivos como “visto que” e “visto che” para que o mapeamento da rede gramatical de cada língua possa ficar mais completo. Além disso, podem ser pensadas etapas futuras a partir de diferentes testes empíricos e computacionais, como a Análise Collostrucional (Stefanowitsch, 2013).

A partir dessas considerações, é perceptível que, embora as orações hipotáticas introduzidas por esses conectivos tenham comportamentos e se insiram em contextos semelhantes, seus usos não são completamente iguais, o que corrobora com o Princípio da Não-Sinonímia (Goldberg, 1995), que guia esta pesquisa. Portanto, não existem formas diferentes que evidenciem significados iguais, assim sendo, se há diferenças na forma, haverá também, em algum grau, uma função comunicativa diferente.

Levando tudo isso em consideração, nesta pesquisa foi realizada uma análise detalhada, a partir dos pressupostos teóricos e metodológicos da Linguística Funcional Centrada no Uso, de duas construções oracionais em português e duas em italiano, além de uma extensa comparação de como cada uma dessas línguas mapeia os usos de construções com valores virtualmente semelhantes. Embora tradicionalmente cada par de construção oracional em cada língua em questão seja considerada sinônima, constatamos que há diferenças discursivas e formais importantes, que apenas uma análise baseada em dados reais pode constatar.

REFERÊNCIAS BIBLIOGRÁFICAS

ARENA, A. B. **Construcionalização do conector “daí que” em perspectiva funcional centrada no uso**. Tese (Doutorado em Estudos de Linguagem), Universidade Federal Fluminense, Niterói, 2015.

AMORIM, F. da S. **Notas sobre a gramaticalização de conectores causais X-que**. Revista da ABRALIN, [S. l.], v. 20, n. 2, p. 1-21, 2021. Disponível em: <https://revista.abralin.org/index.php/abralin/article/view/1834>. Acesso em: 01 de agosto de 2024.

BARLOW, W. ; KEMMER, S. **Usage-Based Models of Language**. Stanford: Stanford University Press, 2000.

BATISTA, P. E. M. **Orações causais introduzidas por conjunções de base participial**. Dissertação (Mestrado em Estudos Linguísticos), Universidade Federal do Mato Grosso do Sul, Três Lagoas, 2013.

BECHARA, E. **Moderna gramática portuguesa**. Rio de Janeiro: Ed. Nova Fronteira, 2009.

BYBEE, J; EDDINGTON, D. **A usage-based approach to Spanish verbs of ‘becoming’**. *Language* 82. p. 323-355. 2006.

BYBEE, J. **Language, usage and cognition**. Cambridge: Cambridge University Press, 2010.

BYBEE, J. L. Usage-based theory and exemplar representations of constructions. In: HOFFMANN, T.; TROUSDALE, G. (Eds.). **The Oxford Handbook of Construction Grammar**. Oxford: Oxford University Press, 2013. p. 49-69.

CEZARIO, M. M.; FURTADO, M. A. **Linguística Centrada no Uso: uma homenagem a Mário Martelotta**. Rio de Janeiro, Mauad-X, 2013. CeZARIO

CEZARIO, M. M. C.; SANTOS SILVA, T.; SANTOS, M. **Formação da Construção [XQUE]CONNECT no Português**. *Revista E-escrita*, v. 6, p. 229-243, 2015.

CROFT, W. **Radical construction grammar: syntactic theory in typological perspective**. Oxford: Oxford University Press, 2001.

CUNHA, C.; CINTRA, L. F. L. **Nova gramática do português contemporâneo**. Rio de Janeiro: Ed. Lexikon, 2008.

DARDANO, M.; TRIFONE, P. **Grammatica Italiana con Nozione di Linguistica**. Bolonha: Ed. Zanichelli, 1995.

DARDANO, M.; TRIFONE, P. **Manuale di linguistica e di grammatica italiana**. 2. ed. Torino: UTET Università, 2011.

DIESSEL, H. Usage-based construction grammar. In: DABROWSKAD, E.; DIVJAK, D. (eds.). **Handbook of Cognitive Linguistics**. Berlin: Mouton de Gruyter, 2015.

DIESSEL, H. Adverbial subordination. In: LURAGHI, S.; PARODI, C. (org.). **Bloomsbury Companion to Syntax**. Londres: Bloomsbury Academic, 2013. p. 341-353.

- DIESSEL, H. **The Grammar Network**: How language structure is shaped by language use. Cambridge: University Press, 2019.
- DIESSEL, H. The ordering distribution of main and adverbial clauses: A typological study. In: *Language*, Volume 77, nº 3. Nova Iorque: Linguistic Society of America, p. 433- 455, 2001.
- FILLMORE, C. J.; KAY, P.; O'CONNOR, C. **Regularity and idiomaticity in grammatical constructions: the case of let alone**. *Language*, 63, 3, 1988. pp. 501-538.
- FURTADO DA CUNHA, M. A.; COSTA, M. A.; CEZARIO, M. M. Pressupostos teóricos fundamentais. In: FURTADO DA CUNHA, M. A.; OLIVEIRA, M. R.; MARTELOTTA, M. E. (orgs.). **Linguística Funcional**: teoria e prática, 1 ed., São Paulo: Parábola Editorial, 2015, p. 21-47.
- GIVÓN, T. **On Understanding Grammar**. Nova Iorque: Academic Press, 1979.
- GIVÓN, T. **Functionalism and grammar**. Amsterdã: John Benjamins Publishing Company, 1995
- GOLDBERG, A. E. **Constructions**: a construction grammar approach to argument structure. Chicago: The University of Chicago Press, 1995.
- GOLDBERG, A. E. Surface Generalizations: An Alternative to Alternations. In: CULICOVER, P.; NOWAK, A. (Ed.). **The Nature of Explanation in Linguistic Theory**. Stanford: CSLI Publications, 2002. p. 327-355.
- GOLDBERG, A. E. **Constructions at Work**: The Nature of Generalization in Language. New York: Oxford University Press, 2006.
- GRIES, S. Th. Behavioral Profile: A fine-grained and quantitative approach in corpus-based lexical semantics. In: LIBBEN, G.; JAREMA, G.; WESTBURY, C. (Eds.), **Methodological and analytic frontiers in lexical research**. Hoboken, NJ: John Benjamins, 2012, pp. 57-80.
- HASPELMATH, M. The converb as a cross-linguistically valid category. In: M. HASPELMATH & E. KÖNIG (eds.) **Converbs in cross-linguistic perspective**. Berlin: Mouton de Gruyter, 1995, 1-55.
- HILPERT, M. **Construction Grammar and its Application to English**. Edimburgo: Edinburgh University Press, 2014.
- HILPERT, M. **Ten Lectures on Diachronic Construction Grammar**. Leiden/Boston: Brill Publisher, v.26, 2021.
- HOPPER, P.; TRAUGOTT, E. C. **Grammaticalization**. Cambridge: Cambridge University Press, 1993.
- HOPPER, P. J. **A primer of grammaticalization**. Amsterdam: Benjamins, 1998.
- ISTITUTO DELLA ENCICLOPEDIA ITALIANA FONDATA DA GIOVANNI TRECCANI. **La Grammatica Italiana Treccani**. Torino: Verba Volante, 2012.
- LANGACKER, R. W. **Cognitive grammar**. Basic Readings, 2008, 29.

- MARTELOTTA, M. E. **Processos de gramaticalização no português do Brasil**: uma abordagem diacrônica. São Paulo: Editora Humanitas, 2003.
- MARTELOTTA, M. E. **Mudança linguística**: uma abordagem baseada no uso. São Paulo: Cortez, 2011.
- NASCIMENTO, J. B. **Construções hipotáticas introduzidas por visto que, dado que e posto que no português brasileiro**: uma análise baseada no uso. 2022. 54 f. Monografia (Licenciatura em Letras - Português e Italiano) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2022.
- NASCIMENTO, J. B.; CASTANHEIRA, D. As orações hipotáticas introduzidas por “visto que”, “dado que” e “posto que”. **LINGÜÍSTICA Baseada no Uso**: Explorando Métodos, 2020, 161.
- NEVES, M. H. M. (org.). **Gramática do português falado**. São Paulo: Humanitas/FFLCH/ USP; Campinas: Editora da UNICAMP, v. 7, 1999. p. 545-591.
- NEVES, M. H. M. **Gramática de usos do português**. São Paulo: Unesp, 2011
- OLIVEIRA, T. P.; CLEMENTE, C. G. C. **Esquematicidade e produtividade na reconfiguração da rede de conectores condicionais**. Revista do GEL, 2022, 19.3: 58-84.
- PIERREHUMBERT, J. B. Exemplar dynamics: Word frequency, lenition and contrast. In: BYBEE, J.; HOPPER, P. **Frequency effects and the emergence of linguistic structure**. Amsterdam: John Benjamins, 2001.
- PINHEIRO, D. Um modelo gramatical para a linguística funcional-cognitiva: da Gramática de Construções para a Gramática de Construções Baseada no Uso. In: ALVARO, P. T.; FERRARI, L. (Orgs.). **Linguística Cognitiva**: dos bastidores da cognição à linguagem. Campos: Brasil Multicultural, 2016.
- POLITO, A. G. **Michaelis Italiano Gramática Prática**. 3.ed. São Paulo: Melhoramentos, 2016.
- PRANDI, M., DE SANTIS, C. **Le regole e le scelte**: Manuale di Linguistica e di Grammatica Italiana. Milano: UTET Università, 2011
- PRINCE, H. Toward a taxonomy of give-new information. In: COLE, P. (ed.) **Radical Pragmatics**. E.U.A., Nova Iorque: Academic Press, 1981, p. 223-255.
- PRINCE, E.F. The ZPG letter: subjects, definiteness and information-status. In: MANN, W.C.; THOMPSON, S.A. **Discourse description**. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1992. p. 295-325.
- RENZI, L.; SALVI, G.; CARDINALETTI, A. (orgs.). **Grande grammatica italiana di consultazione**: II. I sintagmi verbale, aggettivale, avverbiale. La subordinazione. Bologna: Il Mulino, 2001.
- ROCHA LIMA, C. H. **Gramática Normativa da Língua Portuguesa**. Rio de Janeiro: Ed. José Olympio, 1985. RoODR

RODRIGUES, V. V. **Usos de conectores**: uma abordagem funcional-discursiva. Diadorim, Rio de Janeiro, vol. 20 – Especial, p. 535-560, 2018.

SERIANNI, L.; CASTELVECCHI, A. (colab.). **Grammatica italiana**: italiano comune e lingua letteraria. Milano: UTET Università, 1997.

STEFANOWITSCH, A. Collostruction analysis. In: HOFFMANN; TROUSDALE. **The Oxford Handbook of Construction Grammar**. Oxford: OUP USA, 2013.

TRAUGOTT, E. C. Subjectification in grammaticalisation. In: STEIN, D.; WRIGHT, S. (Ed.). **Subjectivity and subjectivisation**: linguistic perspectives. Cambridge: Cambridge University Press, 1995. p. 31-54.

TRAUGOTT, E. C.; BUCHSTALLER, I. **The lady was demonyak**: Historical aspects of adverb ALL. *English Language and Linguistics*, v. 10, 2006.

TRAUGOTT, E.; TROUSDALE, G. **Construcionalization and Constructional Changes**. Oxford: Oxford University Press, 2013.

TRAUGOTT, E. Grammaticalization, constructions and the incremental development of language: Suggestions from the development of degree modifiers in English. In: ECKARDT, R.; JÄGER G.; VEENSTRA, T. (eds.). **Variation, Selection, Development Probing the Evolutionary Model of Language Change**. Berlim, Nova Iorque: Mouton de Gruyter, 2008, p. 219-250.

TRAUGOTT, E. C.; DASHER, B. **Regularity in semantic change**. *Cambridge Studies in Linguistics*, 97, 2002.