

ANÁLISE ESTATÍSTICA DOS PROCESSOS DE GRAMATICALIZAÇÃO NA CLASSE RELACIONAL (CONJUNÇÕES SUBORDINATIVAS CONDICIONAIS) COM BASE NAS OCORRÊNCIAS DO BANCO DE DADOS CORPUSDOPTUGUES.ORG

Felipe de Sousa Jacob (bolsista do ICV), Kilpatrick Müller Bernardo Campelo (Orientador, Depto de Letras – UFPI)

1- INTRODUÇÃO:

O propósito fundamental de minha pesquisa é, em termos estritamente teóricos, apreender os conceitos fundamentais relativos à gramaticalização com base na moldura teórica funcionalista (CAMPELO,2007), e, em termos aplicados, recolher e analisar as ocorrências do rol das conjunções subordinativas condicionais a fim de pôr à prova a tese da existência de formas prototípicas em cada paradigma de organização da estrutura da língua. Trata-se de um trabalho de caráter funcionalista, com a finalidade de detectar quais as formas que, por causa do uso reiterado, são candidatas a promover uma recomposição do quadro gramatical e a expressar, de maneira prototípica, as características do paradigma a que pertencem. A apreensão dos conceitos relativos à gramaticalização não pode prescindir do conhecimento sobre noções gerais de Morfologia, de tipologia linguística, do modelo teórico epistemológico funcionalista e do processo de gramaticalização.

2. METODOLOGIA

Em termos metodológicos, esta pesquisa empreendeu uma seleção e uma análise satisfatória da produtividade e do desempenho das formas usualmente categorizadas como conjunções subordinativas condicionais. A recolha das ocorrências proveio do banco de dados linguísticos *corpusdoportugues.org*. Valeu-se, para o cômputo da produtividade das ocorrências e da análise das variáveis linguísticas e sociolinguísticas, de um programa estatístico. .

O *corpusdoportugues.org* é constituído por mais de 45 milhões de palavras de quase 57,000 textos em português do século XIV ao século XX. A interface permite pesquisar palavras exatas ou frases, curingas, lemas, classes gramaticais, ou qualquer combinação destes. Proporciona também a pesquisa de palavras vizinhas (colocados) com um máximo de dez palavras de cada lado. O *corpus* permite a consulta dos dados de três variáveis sociolinguísticas, a saber: a) uma de natureza diafásica, o registro; b) outra de natureza diatópica, as variedades dialetais genéricas do português europeu e do português brasileiro; c) e, finalmente, outra de natureza diacrônica, disponibilidade de textos do séc. XIV ao XX.

As conjunções subordinativas condicionais que foram levadas em consideração para a análise foram: *se; caso; contanto que; sem que; uma vez que; dado que; desde que; a menos que; a não ser que; salvo se*.

A coleta das orações e das construções linguísticas de valor condicional foram todas retiradas do banco de dados escolhido, o *corpusdoportugues.org*. Houve um esforço para analisar as ocorrências em conformidade com as variáveis linguísticas e sociolinguísticas suscitadas no projeto.

Naturalmente, a plena consecução da análise das ocorrências, com base em todas as variáveis linguísticas e sociolinguísticas levantadas, ficará a cargo de novos empreendimentos de pesquisa, vez que são muitas e complexas variáveis.

3. RESULTADOS E DISCUSSÕES:

A fase executória da pesquisa tomou forma com a coleta das orações e das construções linguísticas de valor condicional, todas retiradas do banco de dados escolhido, o corpusdoportugues.org. Após a recolha das ocorrências, sucedeu-se a análise a fim de atestar a validade das teses aqui abraçadas, qual seja, as de que a frequência de uso responde pela otimização dos traços categoriais de uma classe e/ou subclasse de palavra e de que há uma emergência contínua de formas gramaticalizadas com base, igualmente, na maior frequência de uso.

No decorrer da pesquisa, em especial, do trabalho de discreção categorial, foi possível observar quão tem sido fértil o campo das conjunções subordinadas condicionais na língua portuguesa. Dada a homonímia de alguns itens lexicais do rol de conjunções condicionais, a recolha bruta apresenta grande produtividade, o que seguramente pode enviesar o resultado ora apresentado. A fim de delimitar o universo a ser pesquisado, optou-se por fazer uma recolha mais específica. O período escolhido para a análise, devido à produtividade que tal categoria apresentou em uma sondagem inicial, foi a do século XVII que em seu total de ocorrências tivemos 41.406 ocorrências. Naturalmente, somente uma recolha detida permitirá uma depuração mais fidedigna da produtividade da subclasse relacional alvo desta pesquisa, o que poderá ser levado a cabo em outros empreendimentos de pesquisa.

Depois de diversas especificações e da classificação dos itens que realmente pertenciam à classe relacional condicional, verificamos que, provavelmente, a conjunção **se** configura-se como a forma curinga no aspecto das conjunções subordinativas condicionais, já que, nos dois diagnósticos primários tivemos o seguinte registro de ocorrências: Entre século XIII e o século XX das 440.721 ocorrências, com a seguinte incidência bruta: Se (242.828), caso (18412), contanto que (430), sem que (2642), uma vez que (2185), dado que (384), desde que (2785), a menos que (210), a não ser que (210), salvo se)

Com a escolha do século XVII como alvo tivemos 41.406 ocorrências. Destas, a imensa maioria, 40.212 das ocorrências, são do item **SE**. Ante a sua maior produtividade, escolheram-se cem sentenças com a incidência da conjunção mais produtivas a fim de levar a cabo análise mais específicas. Para as demais variáveis, houve um inelutável recorte, que, a nosso ver, não acarreta prejuízo para comprovar a teoria e as hipóteses regentes da pesquisa.

4 – CONCLUSÕES

Os resultados alcançados no decorrer dessa pesquisa são incipientes. Significa dizer que se empreendeu uma investigação dos fenômenos de gramaticalização incidentes nas conjunções

subordinadas condicionais, mas há necessidade de um aprofundamento teórico e metodológico a fim de alcançar resultados mais conclusivos.

No entanto, apesar empecilhos de todo cariz enfrentados no trajeto da pesquisa, pode-se afirmar que os resultados obtidos podem ser considerados satisfatórios. E faz-se necessário salientar que o aprofundamento da análise desses dados seria de grande importância, a fim de comprovar a tese da proformalização e de ajustar o ensino de língua materna aos achados acadêmicos. Em primeiro lugar, porque questiona aspectos deixados em aberto pela gramática tradicional. Em segundo lugar, porque promove o refinamento do debate acadêmico. Por fim, é fundamental para aprofundar a discussão relativa ao ensino de língua materna, em especial o ensino de descrição da língua.

Decerto a união de recursos teóricos com a quantificação estatística proverá uma descrição mais acurada, o que pode proporcionar ensino de língua materna mais qualificado, em especial de gramática (Morfologia e Sintaxe, principalmente). Uma metodologia que alie análise sistêmica e observância do uso pode, em tese, prover ao aprendiz um melhor entendimento das variações e da constituição linguística. Eventuais insuficiências descritivas poderão ser satisfatoriamente supridas com o cabedal teórico proposto neste trabalho. Da mesma forma, os achados de caráter prático poderão contribuir para que o ensino se embase em ocorrências circulantes e não em exemplários sem representatividade no uso, o que, não raro, sucede com os exemplos das gramáticas ditas tradicionais.

Esse estudo acaba por produzir efeitos benéficos que se estendem para além da iniciação científica. O que o faz se tornar um instrumento de disseminação de conhecimento científico de inestimável valor cultural e social.

5- BIBLIOGRAFIA:

CAMARA Jr, Joaquim Matoso. **Estrutura da língua portuguesa**. Petrópolis: Vozes, 2004.

CAMPELO, Kilpatrick. **O estatuto conceitual e funcional das proformas**. Pronome: o protótipo das proformas. Tese de Doutorado. Fortaleza, Universidade Federal do Ceará, 2007.

FERREIRA, Vanessa Pernas. **A conjunção subordinativa quando na perspectiva funcional-discursiva**. Rio de Janeiro, Faculdade de Letras/UFRJ, 2008. 130 p., mimeo. Dissertação de Mestrado em Língua Portuguesa.

GONÇALVES, Sebastião Carlos Leite; LIMA-HERNANDES, Maria Célia; CASSEB-GALVÃO, Vânia Cristina. **Introdução à gramaticalização**. São Paulo: Parábola Editorial, 2007.

NEVES, Maria Helena de Moura. **A gramática: história, teoria e análise, ensino**.

ROSA, Carlota Maria. **Introdução à Morfologia**. São Paulo: Contexto, 2000.

PALAVRAS – CHAVE: GRAMATICALIZAÇÃO. ESCALARIDADE. CONDICIONAIS.