

ASPECTOS DIACRÔNICOS DA ORDEM DE PALAVRAS EM LÍNGUAS ROMÂNICAS SOB A PERSPECTIVA DA DIFUSÃO SINTÁTICA

João Henrique Rettore Totaro*

Resumo: Este trabalho¹ compreende um estudo quantitativo de *corpora* românicos, com o objetivo de comparar qualitativamente algumas tendências diacrônicas de arranjo sintático nessas línguas. Vários traços foram selecionados para a construção das bases de dados, os quais também orientaram a análise dos perfis de mudança. O objetivo central do trabalho é verificar se essas alterações de ordem dos constituintes são acompanhadas de outras mudanças estruturais e se elas representam instâncias de difusão sintática. Os resultados mostram que o ambiente oracional é determinante no controle de várias ordens de mudança; contudo, nem todos os diversos movimentos aqui identificados caracterizam claramente um processo de difusão, mas exigem uma análise multinível.

Palavras-chave: Linguística românica; sintaxe; ordem dos constituintes; difusão sintática.

1. Introdução

A noção de que algumas línguas naturais tendem a colocar, de maneira consistente, elementos modificadores ou limitantes

* Pontifícia Universidade Católica de Minas Gerais.

¹ Da tese de doutoramento defendida pelo autor na Faculdade de Letras da UFMG em 2007.

antes dos modificados ou limitados, ao passo que outras se comportam de maneira oposta, é bastante difundida entre os linguistas, tendo já sido realizado um grande número de trabalhos,² que não apenas tratam descritivamente do tema, mas oferecem propostas de explicação para essas tendências.

Um ponto recorrente em obras que tratam de linguística histórica consiste em atribuir a Dante Alighieri (*De vulgari eloquentia*, DVE) a fundação dos estudos românicos no início do século XIV (Danesi, 1996). Nesse tratado (I, 8: 4,6), Dante divide as línguas europeias em três grandes famílias (grego, germânico e românico); para esta, propõe uma outra subdivisão em três grupos a partir de um critério lexical: a forma da palavra para “sim” – *si*, *oc* e *oui* – como representativa das particularidades do italiano (que chama de *Latini*) diante do francês (*Franci*) e do provençal (que chama de *Yspani*):³

² MAROUZEAU. *L'ordre des mots dans la phrase latine*; GREENBERG. Some universals of grammar with particular reference to the order of meaningful elements; KRAHE. *Grundzüge der vergleichenden Syntax der indo-germanischen Sprachen*; CANALE. Implicational hierarchies of word order relationships; WAUGH. *A semantic analysis of word order*; HAWKINS. Implicational universals as predictors of word order change; HAWKINS. *Word order universals*; entre outros.

³ “Nam alii *Oc*, alii *Oil*, alii *Sì* affirmando loquuntur, ut puta Hispani, Franci et Latini.”, a partir de GIULIANI (*Le opere latine di Dante Alighieri*, p. 29 et seq.). Não obstante o volume de trabalhos que fazem remontar a Dante o princípio dos estudos em filologia românica, DANESI (Latin vs. Romance in the middle Ages: Dante's de *Vulgari Eloquentia* revisited) observa que é o fato de Dante utilizar a expressão *ydroma tripharium* (na edição de Giuliani, *idioma trifarium*) na DVE que leva os estudiosos a considerarem a raiz de uma teoria genética para as línguas românicas.

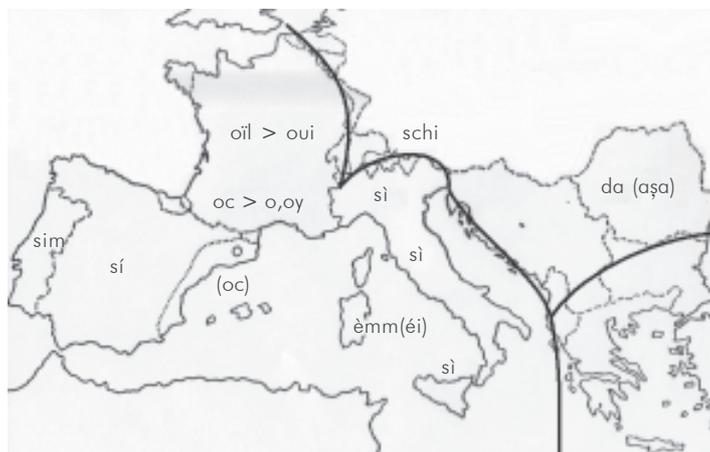


FIGURA 1: Situação das regiões românicas segundo a proposta de Dante Alighieri (1305/1878). <<http://www.fb10.uni-bremen.de/romanistik/homepages/burr/Intro/Klassifizierung.htm>>, modificado segundo Ostler.⁴

Posner⁵ afirma que a discussão das relações de parentesco entre o latim e as línguas românicas (LR) teve início no século XVI, quando os vernáculos já estavam em condições de substituir o latim. Entretanto, é apenas no século XVIII, quando a regularidade da mudança sonora começa a ser percebida, que os trabalhos clássicos sobre as origens das LR aparecem em profusão. No início do século XIX, F.-J. Raynouard distinguiu traços comuns nas LR que não poderiam ser todos atribuídos à herança latina, e postulou um ancestral comum, *la langue romane*, que teria substituído o latim antes do ano 1000 e que se assemelhava ao provençal antigo. Essa proposta, contudo, não se alinha com a de A. W. von Schlegel, um dos grandes nomes da filologia comparativa alemã, que via as LR como resultado espontâneo da decomposição do latim após a perda da pressão normativa.

⁴ OSTLER. *Empires of the word: a language history of the world*.

⁵ POSNER. *The romance languages*.

O primeiro tratado sistemático sobre a questão românica teria sido publicado em 1831 por L. Diefenbach,⁶ que sustentava que as LR teriam derivado de uma forma popular de latim, posteriormente alterada pelo contato com outras línguas – especialmente no caso do francês, que considerava “muito crioulezado”. Diefenbach teve como discípulo F. Diez, que também acreditava que o latim popular estava na origem das LR, cujas especificidades seriam devidas à ação de processos fonológicos e a ajustes funcionais; deve-se também a ele a aplicação do método histórico-comparativo às línguas românicas, a exemplo do que já tinha sido feito com sucesso por F. Bopp para as línguas indo-europeias e por J. Grimm para as germânicas.⁷ Diez dividiu o espaço românico em porções oriental e ocidental – uma proposta que foi posteriormente desenvolvida por W. von Wartburg⁸ e vários outros, como, por exemplo, G. Rohlfs, que distingue, com critérios lexicais, uma România Nuclear – Itália e Gálias – e outra, Periférica – Ibéria e Dácia; na Itália, delimita grandes regiões dialetais a partir da observação de feixes de isoglossas que passam por La Spezia e Rimini ou Pesaro (similar às linhas Massa-Senigallia, um pouco mais ao N, e Roma-Ancona.⁹

No século XX, sobressaem também outros nomes que cuidaram do tema. A. Meillet trata tanto dos aspectos estruturais das línguas indo-europeias e de seus dialetos quanto da aplicação do método histórico-comparativo em linguística histórica, e formou É. Benveniste, M. Cohen, A. Martinet, L. Tesnière e J. Vendryes. Temos, posteriormente, E. Bourciez, que adota uma visão funcionalista da mudança como resultado de uma tensão entre a economia de esforço e o desejo de comunicação efetiva (cf. explícito no §32 de sua obra¹⁰)

⁶ POSNER. *The romance languages*.

⁷ BASSETTO. *Elementos de filologia românica*.

⁸ VON WARTBURG. *Evolución y estructura de la lengua*.

⁹ RENZI. *Nuova introduzione alla filologia romanza*

¹⁰ BOURCIEZ. *Éléments de linguistique romane*.

e oferece uma visão ampla da história externa e interna das línguas neolatinas; C. Tagliavini e A. Monteverdi, que utilizam critérios geográficos e estruturais para distinguir a România Oriental – Balcãs e Dácia – da Itálorromânia – toda a Itália, mais o sardo, o friulano, o rético e o ladino – e da România Ocidental – apenas as Gálias; e B. Vidos, que também divide a România em frações Oriental e Ocidental, mas isola nesta última o bálcano-românico a partir de argumentos estruturais e culturais. Ainda que sejam pertinentes as considerações feitas por Ruhlen¹¹ sobre as relações entre a afinidade genética e a continuidade dialetal entre línguas exercidas em espaços contíguos, acolhi, para efeito deste trabalho, os conceitos de *România oriental* e *ocidental* de Vidos,¹² que deixa no primeiro domínio apenas os Balcãs e a Dácia.

No plano diacrônico, diversos estudiosos¹³ ressaltam que, a partir da leitura de textos escritos em línguas diferentes e em épocas diversas, bem como de gramáticas históricas que tratam descritivamente de momentos pretéritos da evolução de sistemas linguísticos particulares, é possível observar diferenças na ordenação de seus elementos constituintes em relação aos seus respectivos usos contemporâneos. No caso da colocação do adjetivo (A) em relação ao substantivo (N) no sintagma nominal (SN) como indicativo de particularidades internas ao grupo românico, por exemplo, essas mudanças posicionais seriam motivadas por fatores de natureza semântica,¹⁴ para o português; González,¹⁵ para o

¹¹ RUHLEN. A guide to the world's languages.

¹² VIDOS. *Manual de lingüística românica*, p. 242 *et seq.*, e, em particular, p. 248.

¹³ COHEN. *Syntactic change in portuguese*: relative clauses and the position of the adjective in the noun phrase; GONZÁLEZ. La colocación del adjetivo atributivo en el español medieval: un problema metodológico e historico.

¹⁴ COHEN. *Syntactic change in portuguese*: relative clauses and the position of the adjective in the noun phrase.

¹⁵ GONZÁLEZ. La colocación del adjetivo atributivo en el español medieval: un problema metodológico e historico.

espanhol ou fonológico-morfológicas, como a perda de parte das flexões latinas devido a fatores fonológicos, e que teria, por sua vez, resultado numa fixação da ordem dos constituintes nas línguas românicas.¹⁶

Nesse contexto e do ponto de vista diacrônico, as mudanças de ordem de constituintes num sistema podem encontrar paralelos, em estágios evolutivos diversos, em línguas da mesma família, o que demandaria uma análise cuidadosa de textos que permitam abranger um largo intervalo de tempo. Para o grupo românico, entretanto, não há estudos que contemplem, de forma detalhada e comparativa, a mudança da ordem de palavras no SN em relação a outras mudanças ocorridas nessas línguas, como, por exemplo, na colocação preferencial de constituintes subordinados ao SV. Naturalmente, apesar de algumas ordens serem menos frequentes, suas proporções *relativas* podem ter sido diferentes em outros séculos, passando por um período de instabilidade tipológica mais pronunciada antes de chegar aos valores atuais. Parte da literatura teórica disponível,¹⁷ sugere que a primeira a se processar no espanhol (e, possivelmente, nas demais que são objeto deste trabalho) teria sido a referente aos constituintes S-V-C, ocorrendo, posteriormente, o realinhamento do parâmetro A-N e de outros. Se de fato assim o foi, a amostragem realizada em textos escritos nessas línguas acusará uma grande estabilidade do padrão de arranjo dos constituintes S-V-C e variações cada vez menores dos outros parâmetros.

¹⁶ TAGLIAVINI. *Le origini delle lingue neolatine*; WOLFF. *Origen de las lenguas occidentales, 100-1500 d.C.*

¹⁷Cf., especialmente, GREENBERG. Some universals of grammar with particular reference to the order of meaningful elements; e CANALE. Implicational hierarchies of word order relationships.

2. Objetivos/formulação da questão

Este trabalho pretende lançar luz sobre aspectos estruturais da mudança (ou da manutenção) da posição dos constituintes S-V-C, contrapondo esses resultados àqueles encontrados para o SN português¹⁸ e espanhol.¹⁹ Para tanto, realiza-se um estudo quantitativo tão detalhado quanto necessário ao desenvolvimento da questão, mas procurando-se dar ênfase especial à análise qualitativa dos dados linguísticos; será possível, então, comparar tendências sintático-semânticas dos arranjos sujeito/verbo/CV e A/N entre sincronias diferentes e avaliar a relevância dessas tendências para o processo de mudança, como feito por Cohen²⁰ e por Rijkhoff.²¹

Nesse diapasão, o objetivo central do trabalho compreende a investigação da possibilidade de as línguas em questão terem sofrido mudanças na ordenação de constituintes através de processos que podem servir à caracterização de fenômenos de mesma ordem no quadro românico e, em nível mais abrangente, indo-europeu.²² Isso implica estender os princípios do modelo de difusão lexical²³ ao nível sintático: a mudança posicional dos constituintes teria, então, ocorrido de forma *sintaticamente abrupta e lexical ou*

¹⁸ COHEN. *Syntactic change in portuguese*: relative clauses and the position of the adjective in the noun phrase.

¹⁹ TOTARO. *Mudança de ordem dos constituintes A e N em textos espanhóis dos séculos XIII a XX*.

²⁰ COHEN. *Syntactic change in portuguese*: relative clauses and the position of the adjective in the noun phrase.

²¹ RIJKHOFF. Explaining word order in the noun phrase.

²² Reflete-se, aqui, tanto uma preocupação com o encaixamento do fenômeno em apreço quanto à necessidade de se confirmar a aplicação da proposta de Hawkins (1979; 1983), sob a forma de um Princípio de Harmonia Transcategorial (Cross-Category Harmony Principle), segundo o qual deve haver uma tendência a colocar os elementos modificadores dos núcleos dos sintagmas sempre de um mesmo lado (antepostos ou pospostos).

²³ SHEN. *Lexical diffusion*: a population perspective and a numerical model.

sintagmaticamente gradual, subordinada a condicionamentos estruturais que responderiam por esse processo de *difusão sintática*. Naro e Lemle²⁴ caracterizaram o fenômeno a partir da observação de que, ao longo da história do português, não apenas aumentam as proporções de determinada ordem em certos ambientes oracionais, como também essa mesma ordem gradativamente atinge outros. Para a verificação das condições sob as quais a mudança é implementada, Naro e Lemle²⁵ selecionaram fatores que vão desde a saliência fônica do constituinte nominal, passando pela classe morfológica do verbo e chegando à posição de superfície do SN diante do V. Os resultados a que chegaram são corroborados pelas conclusões de Mattos e Silva²⁶ sobre as estruturas atributivas do português arcaico, as quais ilustram percursos da difusão sintática entre os séculos XIV e XV. Ainda para o português, Mollica²⁷ demonstra que outra mudança – envolvendo o *queísmo* e o *dequeísmo* – seria delineada nos moldes difusionistas; somando-se ao conjunto acima referido sobre o português, isso mostra que a mudança é, de fato e como ressaltado por Naro e Lemle,²⁸ regular e sistêmica:

In particular, we will be interested in the question of the way in which a syntactic change is realized across the relevant formal categories of a language and in the possible existence of a general principle governing this phenomenon... This [] diffusion occurs along several distinct dimensions simultaneously until eventually the whole language is entrapped.

Esse modelo contempla justamente a possibilidade de uma determinada mudança sintática espalhar-se (ou não) através

²⁴ NARO; LEMLE. Syntactic diffusion.

²⁵ NARO; LEMLE. Syntactic diffusion.

²⁶ MATTOS E SILVA. *O português arcaico*: fonologia, morfologia e sintaxe.

²⁷ MOLLICA. *Queísmo e dequeísmo em português*.

²⁸ NARO; LEMLE. Syntactic diffusion, p. 259-267.

do sistema linguístico – por exemplo, de um item lexical para outro ou entre ambientes sintáticos. Dessa forma, privilegiamos a caracterização e a comparação dos recortes sincrônicos, procurando observar o alcance interno das modificações estruturais enquanto as mesmas se processam. Essa posição exclui, naturalmente, a busca *a priori* de uma justificativa externa para a mudança de sistemas linguísticos derivados de uma única fonte,²⁹ ainda que esse movimento se mostre convergente entre as línguas estudadas. Ademais, partindo-se do pressuposto de que a mudança pode ser identificada em qualquer e em todos os níveis de cada sistema linguístico,³⁰ estendem-se à morfologia e à sintaxe os princípios da difusão lexical. Ao mesmo tempo, a apreciação dos diferentes perfis sintáticos dessas línguas permite não apenas estudarem-se as possíveis combinações históricas de ordens intra- e inter-sintagmáticas de constituintes, como também aventarem-se possibilidades de motivação interna para a maior (ou menor) estabilidade de algumas dessas ordens em ambientes sintáticos determinados e para todo o grupo românico.

3. Metodologia

Foram selecionados textos em prosa escritos originalmente nas línguas em estudo, cobrindo o mais amplo intervalo temporal possível para essas línguas até o século XX. Utilizam-se dois textos de cada século por língua a fim de se ter uma amostragem mais representativa da tendência de colocação dos constituintes nesses períodos.

²⁹ SAPIR. *Language: An introduction to the study of speech*.

³⁰ Cf. THOMASON; KAUFMAN. *Language contact, creolization and genetic linguistics*, p. 9 *et seq.*

3.1 Construção das bases de dados

A coleta de ocorrências foi feita em textos espanhóis dos séculos XIII a XX, portugueses dos séculos XV a XX e italianos dos séculos XV a XX; gramáticas históricas apontaram as estruturas sintáticas correntes à época em que os textos foram escritos, considerando-se com mais cuidado, por certo, a classificação de casos menos prototípicos. Na seleção dos textos, utilizaram-se os mesmos critérios que orientaram em Totaro (1998) a busca de fontes filologicamente adequadas para a amostragem de sintagmas nominais. O procedimento de estruturação das bases de dados acompanha o disposto por Leech *et al.*³¹ e Biber, Conrad e Reppen³² sobre a marcação sintática de *corpora*.

Os excertos das obras que se prestaram à prospecção de ocorrências foram: a) digitalizados com o auxílio de *scanner* e convertidos em formato.TXT através da utilização de um sistema OCR da HP, nos casos em que as bibliotecas visitadas dispunham de exemplares para consulta; ou b) descarregadas diretamente das bibliotecas eletrônicas em formato PDF ou JPG através de protocolos servidor-cliente e tratadas manualmente.

3.2 Métodos estatísticos

O número mínimo de ocorrências a ser colhido em cada texto foi determinado a partir da metodologia exposta em Woods, Fletcher e Hughes³³ e em Tesitelová.³⁴ As diferenças individuais existentes entre os textos de um mesmo século foram normalizadas tomando-se a *média* dos valores de cada texto. O procedimento mostra que as mudanças apontadas pelas proporções relativas

³¹ LEECH *et al.* *Recommendations for the syntactic annotation of corpora*.

³² BIBER; CONRAD; REPPEN. *Corpus linguistics: investigating language structure and use*.

³³ WOODS; FLETCHER; HUGHES. *Statistics in language studies*.

³⁴ TESITELOVÁ. *Quantitative linguistics*.

das classes ultrapassam os limites da simples variação individual ou dialetal presente nos textos, refletindo, de fato, um aspecto evolutivo do sistema. Observe-se, ainda, que essa evolução se dá segundo um perfil sigmoidal, que é justamente o esperado como representante da mudança não catastrófica de um parâmetro³⁵ decorrente da competição gramatical entre variantes.³⁶ O mesmo perfil é apresentado por Batali,³⁷ em simulações feitas com redes neurais, corroborando a viabilidade e a propriedade do tratamento numérico como elemento de análise das ocorrências reais.

As bases de dados foram construídas no MS Excel considerando-se a portabilidade dos arquivos.XLS e o grande número de funções que o programa oferece para o tratamento estatístico de dados. Procedeu-se à criação de um arquivo de ocorrências, codificações e tratamento para cada língua, separando-se em pastas individuais os dados de cada texto; depois dessas pastas, seguem as referentes ao tratamento estatístico – uma para cada fator e seus desdobramentos (como, por exemplo, o cruzamento com outros e os gráficos correspondentes aos resultados numéricos), na mesma sequência em que constituem cada pasta: *ordem de constituintes, peso de S e de C, ambiente oracional, regência do verbo, tempo do verbo, classe semântica do verbo, agentividade, conjugação verbal e item lexical*.

Para a contagem das ocorrências, o programa permite a associação de um filtro (*Data>Filter>AutoFilter*) a um intervalo de células qualquer: acionado manualmente através da seleção de um dos valores compreendidos no intervalo de amostragem,

³⁵ LIGHTFOOT. *How to set parameters*: arguments from language; SHEN. *Lexical diffusion*: a population perspective and a numerical model; NIYOGI; BERWICK. *The logical problem of language change*; GOERTZEL. *Chaotic logic*; GOERTZEL. *The miraculous mind attractor*; KIRBY. *Function, selection and innateness: the emergence of language universals*.

³⁶ KROCH. Morphosyntactic variation.

³⁷ BATALI. Computational simulations of the emergency of language.

esse filtro retorna o número de *tokens* que detêm a característica selecionada, além de permitir isolar as linhas que contém esses traços. Os filtros podem, por fim, ser concatenados, combinando-se os critérios de classificação das ocorrências.

3.3 Critérios de classificação das ocorrências

As hipóteses que orientam a identificação e a seleção dos fatores pertinentes são fruto da observação das ocorrências encontradas nos textos. A classificação destas foi feita segundo critérios que permitissem: a) definir combinações de fatores estruturais indicadoras de um controle da mudança; b) determinar a magnitude dessas alterações, quando as houvesse; e c) definir a tendência geral de comportamento dos sistemas linguísticos amostrados. Os critérios morfológicos foram sugeridos a partir de trabalhos que contemplam a influência de traços dessa natureza e do peso do sintagma na mudança;³⁸ os semânticos, a partir de Waugh,³⁹ Lehrer⁴⁰ e de Beaugrande.⁴¹

Para o tratamento das ocorrências e a análise dos resultados preliminares, veja-se que o cruzamento de fatores individuais foi feito aos pares, tomando-se como fator dependente a posição dos modificadores em relação a seus respectivos núcleos. Os gráficos gerados a partir do tratamento numérico permitem avaliar a consistência das hipóteses iniciais sobre a natureza desses critérios de classificação das ocorrências, já que apresentam perfis claros de mudança.

Com respeito à ordenação básica dos constituintes sujeito-verbo-complemento, as ocorrências foram classificadas de acordo com os critérios listados abaixo, de a a i, e seus desdobramentos

³⁸ WILMET. *La détermination nominale*; MARTINS. The loss of IP-scrambling in portuguese: clause structure, word-order variation and change.

³⁹ WAUGH, L. *A semantic analysis of word order*.

⁴⁰ LEHRER. *Semantic fields and lexical structure*.

⁴¹ DE BEAUGRANDE. *New foundations for a theory of text and discourse*.

(acompanhados por exemplos, aqui, apenas em a). Adota-se, aqui, uma linha de trabalho semelhante à de Costa,⁴² que situa, em primeiro lugar, a posição do elemento central em apreço, e, em seguida, os fatores que possivelmente contribuem para, ou restringem,⁴³ os fenômenos de mudança de ordem observados no sistema.

- α. Posição dos constituintes - SVC, SCV etc.; AN, NA (apenas para o *corpus* italiano)
- β. Natureza do sujeito e do complemento – pleno, pronominal ou oracional.
- γ. *Status* morfológico do pronome sujeito ou complemento;
- δ. Ambiente pronominal.
- ε. Ambiente oracional.
- ζ. Regência e tempo do verbo.
- η. Classe semântica do verbo⁴⁴
- θ. Agente e Paciente.
- ι. Frequência de ocorrência do item lexical.

⁴² COSTA. *Word order variation: a constraint-based approach*.

⁴³ VINCENT. *Competition and correspondence in syntactic change: null arguments in latin and romance*; KROCH. *Syntactic change*.

⁴⁴ DE BEAUGRANDE. *New foundations for a theory of text and discourse*.

QUADRO 1
 Ordenação básica dos constituintes sujeito-verbo-complemento

<u>S</u> <u>V</u> <u>C</u>	por huma bahia asima onde <u>os turquos</u> <u>tem</u> <i>huma forsa bem fraqua de pedra</i> [<i>Pe. Jerônimo Lobo</i>]
<u>S</u> <u>C</u> <u>V</u>	et p{er}o <u>l</u> angelo <u>l</u> adimando piena digratia quasi mutandole ilnome p{er} [<i>Expositione del Reverendo</i>]
<u>V</u> <u>S</u> <u>C</u>	sabedes <u>vos todos</u> <i>que se guarnesçen agora ellos quanto pueden contra nos</i> [<i>Historia Troyana</i>]
<u>V</u> <u>C</u> <u>S</u>	aos quaaes nom deve <u>fazer perjuizo</u> <u>o que outras fezerom em contrairo</u> [<i>Leal Conselheiro</i>]
<u>C</u> <u>S</u> <u>V</u>	mas de que <u>lo vos vierdes</u> [<i>Libro del caballero y del escudero</i>]
<u>C</u> <u>V</u> <u>S</u>	<i>Molto auara</i> mi è <u>V.S.</u> delle sue lettere, da me [<i>Viaggi del pellegrino</i>]

Cabem ainda duas palavras sobre os perfis elaborados a partir das tabelas de dados numéricos: em primeiro lugar, notam-se, de maneira geral, tendências claras e inegáveis de aumento, de manutenção ou de decréscimo das médias de frequências por século, mas não de todas ao mesmo tempo – o que indica, com elevado grau de segurança, a propriedade das classes e a regularidade de sua aplicação. Por outro lado, alguns pontos excêntricos diminuem o grau de ajuste das curvas polinomiais,⁴⁵ mas é natural que haja um ou outro dado que aumente a dispersão do grupo. Esse efeito individual, que é muito mais sensível nos limites do intervalo de amostragem, pode ser reduzido graças à ampliação longitudinal (para alcançar outros séculos) ou transversal (utilizando-se outros textos como

⁴⁵ Retas de regressão linear e curvas que se ajustam aos pontos dos gráficos são traçados automaticamente por programas de tratamento estatístico, o que permite adequar o grau de refinamento das curvas de tendência. As retas de regressão e curvas de segundo grau (do tipo $ax^2 + bx + c = 0$) mostram mais claramente as tendências do conjunto de dados, mas podem não se ajustar perfeitamente aos valores numéricos, desprezando as oscilações representadas por pontos que escapam à tendência do grupo; por sua vez, curvas mais complexas (do tipo $ax^3 + bx^2 + cx + d = 0$) se ajustam melhor aos dados, mas sua construção enfatiza aspectos pontuais, e não o comportamento do conjunto.

fonte de ocorrências) da base de dados, o que leva a uma diluição das proporções particulares a cada texto e a uma aproximação de uma tendência de conjunto.

4. Tratamento numérico das ocorrências

Os gráficos gerados pelas bases de dados permitem observar não apenas as tendências dos sistemas individuais, como comparar os perfis de mudança.⁴⁶

A base de dados de SNs italianos compreende cerca de 2000 ocorrências⁴⁷ colhidas nos seis séculos da amostragem, perfazendo pouco mais de 150 SNs por texto. São exemplos de ocorrências encontradas neste estudo:

QUADRO 2
Exemplos de ocorrências de AN e de NA nos textos do *corpus* italiano

Século	Texto	Ocorrência
XV	SERM	Incominciano i deuotissimi Simoni del diuo Bernardo Abbate di sicontiene la sôma di tutte le uirtute necessarie a ciascheduno
	PUNG	& torno a penitentia somma malignita e & diabolica crudelta rinfacciare della contentione e peccato dia/bolicho & grauissimo & stoltissi/mo & cagione
XVI	HIST	erano conosciute in gran parte da molti:& tra gli altri perniciosissimo à Italia,et à tutta la Republica Christiana:
		col ritrouarsi in tutte quelle importanti espeditioni ,che si faranno a i suoi di, ogni
	PIGN	Et sopra il tutto douerà far ogni opera per interuenire ne' consigli secreti.

(continua)

⁴⁶ Cf. metodologia proposta por BRILL *et al.* Deducing linguistic structure from the statistics of large corpora; ELMAN, Finding structure in time; McMAHON; SMITH, A review of statistical language processing techniques.

⁴⁷ Seria interessante estudar também os condicionadores estruturais da variação entre adjetivos e formas encabeçadas por preposição, como “regioni **del maré**” X “regioni **maritime**”.

(continuação)

XVII	GALI	che elle siano, prestano subito assenso , & applauso: & all'incontro, quelle
		ma con isdegno , & ira acerbissima , e taluno di costoro spinto dal furore
	CITT	Contra di lui insorse Celio Re de gli Albani con grosso esercito : e giunto
		verilmente gli era uscito contra con l' esercito romano , si conuenero
XVIII	PRIN	tre principali punti; che si dia Provvedenza divina ; che l'anime umane sien'
		e che si debbano moderare l' umane passioni con la Giustizia, e da quella sì
	LETT	raramente sono giudici competenti , illuminati, e caldamente spassionati
		Mille ostacoli impediscono il retto giudizio degli altri; ma, freddato interamente

As proporções individuais de AN e de NA foram somadas século a século, e as médias foram lançadas no Gráfico 1.

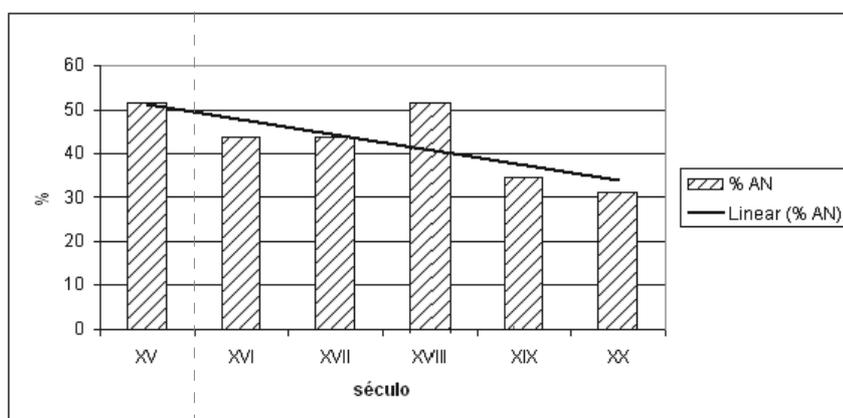


GRÁFICO 1: Evolução das médias gerais de AN por século nos textos do *corpus* italiano utilizados para amostragem

No que respeita à posição relativa dos constituintes A e N internos ao SN, o *corpus* italiano exhibe, portanto, a mesma

tendência à ordem NA apresentada por seus pares português⁴⁸ e espanhol.⁴⁹ Além disso, vê-se que essa mesma reta intercepta a linha de 50% próximo à passagem do século XV para o XVI (no gráfico, a reta pontilhada); esse resultado é interessante, uma vez que o francês teria cruzado esse mesmo patamar no século XVI;⁵⁰ o espanhol o fez no século XVII;⁵¹ e o português, no século XVIII.⁵²

A partir dos resultados numéricos, compreendo que, da maneira como os fatores foram apresentados e utilizados para classificar as ocorrências, o método é eficiente para localizar fenômenos associados à mudança sintática nas línguas estudadas. Entretanto, a própria definição dos critérios de marcação das ocorrências impõe uma série de limitações que demandam investigação mais detalhada do peso de cada fator e de seus desdobramentos (subtipos oracionais, classes semânticas etc.).

O fator *ambiente oracional*, que se relaciona diretamente à questão da difusão sintática, merece destaque em virtude de determinar os movimentos dos demais fatores, o que lhe confere um peso maior na análise. Assim é que os outros fatores são dependentes da variável *ambiente oracional* de forma semelhante, ou seja, os perfis dos gráficos que relacionam essas variáveis são similares – o que leva à consideração de que a difusão entre ambientes sintáticos não envolve, por exemplo, apenas os itens lexicais mais ou menos frequentes, mas sim, níveis estruturais diversos.⁵³ É

⁴⁸ COHEN. *Syntactic change in portuguese: relative clauses and the position of the adjective in the noun phrase.*

⁴⁹ TOTARO. *Mudança de ordem dos constituintes A e N em textos espanhóis dos séculos XIII a XX.*

⁵⁰ VON WARTBURG. *Evolución y estructura de la lengua francesa.*

⁵¹ TOTARO. *Mudança de ordem dos constituintes A e N em textos espanhóis dos séculos XIII a XX.*

⁵² COHEN. *Syntactic change in portuguese: relative clauses and the position of the adjective in the noun phrase.*

⁵³ Cf. também proposto por HAWKINS. Implicational universals as predictors of word order change; HAWKINS. *Word order universals*; HAWKINS. Seeking motives for change in typological variation.

difícil, portanto, sustentar a tese da difusão generalizada entre esses ambientes sintáticos a partir dos fatores selecionados para o tratamento dos *corpora* utilizados neste estudo; o que se observa, sem sombra de dúvida, é uma série de perfis simultâneos e complementares de mudança – mas que são de ordem diferente daqueles apontados por Naro e Lemle⁵⁴ e Davies.⁵⁵ Os dados aqui apresentados indicam que o *ambiente oracional* deve ser seriamente contemplado numa análise da mudança sintática, e que é preciso estudar mais refinadamente as relações entre os fatores *ordem de constituintes*, *item lexical* e *ambiente oracional*, e entre esses três e outros ainda.

Vejamos, então, alguns aspectos particulares da aplicação dos diferentes critérios de análise. Inicialmente, a ordem SVC é, em todo o intervalo de amostragem, relativamente mais comum que qualquer outro ordenamento dos três constituintes. Portanto, a orientação dos *corpora* no sentido de uma mudança SCV – SVC de Canale⁵⁶ aconteceu fora do intervalo de amostragem; conseqüentemente, analisando as bases de dados aqui apresentadas e descartando as projeções gráficas, não é possível situar cronologicamente essa mudança no SV nem relacioná-la à mudança posterior AN – NA nessas línguas,⁵⁷ como proposto por Canale.⁵⁸ Isso leva a uma avaliação do intervalo em que se processaram essas alterações nos sistemas EIP, no sentido de que a mudança no nível S-V-C demanda um tempo maior que o necessário a muitos outros ajustes sintáticos.

Outros movimentos, contudo, acompanham as oscilações de frequência das ordens que envolvem os constituintes S-V-C no intervalo de amostragem e é possível, portanto, situá-los em relação

⁵⁴ NARO; LEMLE. Syntactic diffusion.

⁵⁵ DAVIES. Syntactic diffusion in spanish and portuguese infinitival complements.

⁵⁶ CANALE. Implicational hierarchies of word order relationships.

⁵⁷ VON WARTBURG. *Evolución y estructura de la lengua francesa*.

⁵⁸ CANALE. Implicational hierarchies of word order relationships.

à mudança dos constituintes A-N. O fator *ambiente oracional*, por exemplo, compreende um avanço das frequências absolutas de orações principais (especialmente em espanhol, cujo perfil ultrapassa os 50% no século XX e um recuo nas frequências de orações subordinadas nos três sistemas para baixo dos 50%, segundo a sequência *espanhol* (século XVIII) – *português* (século XVIII-XIX) – *italiano* (século XX). Isso faz com que, nos casos ibéricos, a mudança de um modelo oracional predominantemente hipotático ceda lugar às orações principais e coordenadas pouco depois da mudança $AN > NA$; para o italiano, há um intervalo maior entre esses movimentos.

Uma comparação entre as frequências absolutas das ordens que têm os três constituintes S-V-C explícitos (Gráficos 2 a 4) mostra peculiaridades de cada sistema – como a tendência à diminuição generalizada de ocorrência simultânea de S e de C em português, diante do comportamento contrário nos *corpora* espanhol e italiano.

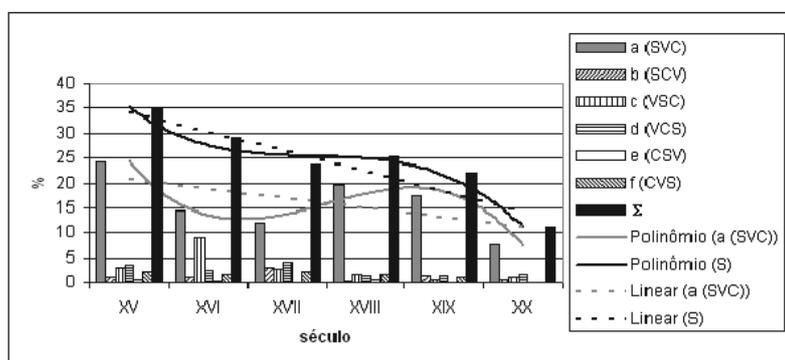
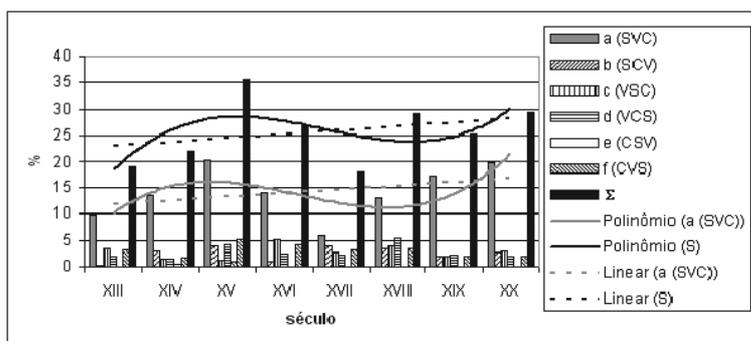
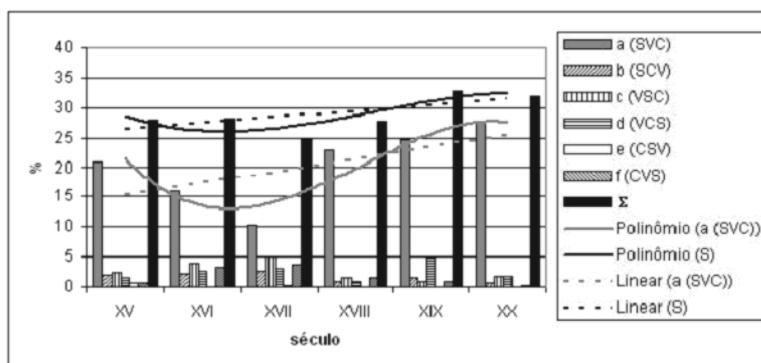


GRÁFICO 2: Distribuição das ordens de constituintes S-V-C no *corpus* português

No *corpus* espanhol, as frequências absolutas de ocorrência das ordens que têm os três constituintes S-V-C explícitos mostra uma tendência semelhante: o padrão mais recorrente é também SVC, e a frequência de ocorrência dessa ordem é crescente. A curva S, que reflete a somatória de todas as frequências, tem inclinação positiva, o que aponta a probabilidade cada vez maior de se encontrarem ordens que envolvam simultaneamente S e C explícitos em textos espanhóis.

GRÁFICO 3: Distribuição das ordens de constituintes S-V-C no *corpus* espanhol

Os dados dos textos italianos apontam para um comportamento semelhante ao do sistema espanhol: as proporções de ocorrências SVC, dominantes em todo o intervalo de amostragem, tendem a aumentar em termos relativos. Em decorrência da diminuição das frequências dos demais tipos sintáticos, a inclinação da reta que resolve a distribuição S do conjunto é mais baixa, mas ainda francamente positiva. Esses perfis mostram, novamente, uma individualização do português diante dos outros dois sistemas, uma vez que este se mostra menos sensível às pressões que levam a um aumento de SVC em espanhol e em italiano.

GRÁFICO 4: Distribuição das ordens de constituintes S-V-C no *corpus* italiano

No Gráfico 5, apresentam-se as curvas referentes à distribuição, nos três sistemas, de estruturas SVC, VC e do conjunto de todas as ordens em que o verbo é seguido de seu complemento. Observa-se facilmente que há uma espécie de setorização dos feixes de curvas relativas a um traço específico, especialmente na região medial do gráfico: as curvas SVC ficam na faixa dos 10 – 30%, as VC ocupam a faixa dos 30 – 50%, e as relativas ao grupo [VC] correm entre os 50 – 70%. A dispersão das curvas no final do intervalo coberto pela amostragem sugere a possibilidade de os três sistemas se individualizarem nos últimos pontos do gráfico; entretanto, as regiões-limite dos gráficos devem ser consideradas com cautela, pelos motivos expostos acima.

Em relação aos resultados fornecidos pelo tratamento das ocorrências segundo os critérios de análise selecionados, verifica-se que as alterações nas frequências das classes de dados favorecem os traços prototípicos através das estruturas não marcadas (corroborando, portanto, as previsões de Naro e Lemle⁵⁹ nesse sentido): sobem as proporções de orações principais, paralelamente ao avanço generalizado das ordens [VC], das classes *sujeito pleno* e *complemento verbal pleno*, de *verbos transitivos diretos* e de *verbos de I conjugação*. Nota-se, portanto, que, a despeito de essas mudanças de frequências serem claras, elas não indicam tendência a alteração na orientação dos sistemas, mas, sim, reforçam as características adquiridas em fases anteriores não cobertas pela amostragem. Nesse sentido, os movimentos detectados correspondem, no âmbito da tipologia das línguas românicas e do grupo indo-europeu, a) a um alinhamento de vários traços estruturais cuja dependência mais estreita deve ser melhor estudada, ou b) a oscilações cuja amplitude não implica mudança tipológica. Esses traços *default* comuns aos três sistemas estudados podem, portanto, ser tratados sob uma única perspectiva, assumindo-se a ausência de transformações para a produção deles em relação aos traços não prototípicos.

⁵⁹ NARO; LEMLE. Syntactic diffusion.

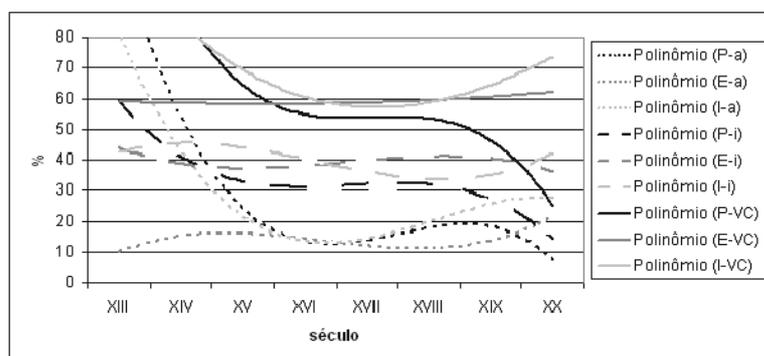


GRÁFICO 5: Sobreposição das curvas de frequência absoluta de ocorrência de estruturas SVC, VC e do grupo VC.

Exemplifique-se: não há necessidade de funções que operem uma conversão de orações principais, mas há uma operação (morfo-)sintática que relaciona coordenadas e subordinadas às principais. Ainda que o estudo detalhado dessa hipótese de trabalho esteja além da abrangência desta investigação, é impossível deixar de notar a propensão à ocorrência concomitante desses traços *default* (nos moldes da harmonia trans-categorial de Hawkins,⁶⁰ a despeito de haver uma flutuação em suas frequências absolutas. O argumento não exclui, obviamente, a possibilidade de se verificar grande estabilidade em sistemas caracterizados por outros arranjos de constituintes e combinações diversas de fatores de várias ordens, como demonstrado por Fodor⁶¹ e retomado por Hawkins,⁶² Lightfoot,⁶³ Dryer,⁶⁴ e Kaufmann;⁶⁵ entre outros. Os fatores que contribuem para essa estabilidade de combinações não *default* ainda são uma questão em aberto.

⁶⁰ HAWKINS. Implicational universals as predictors of word order change; HAWKINS. *Word order universals*.

⁶¹ FODOR. *The rate of linguistic change*.

⁶² HAWKINS. *Word order universals*.

⁶³ LIGHTFOOT. *How to set parameters: arguments from language change*.

⁶⁴ DRYER. The Greenbergian word order correlations.

⁶⁵ KAUFMANN. *The origins of order: self-organization and selection in evolution*.

Em linhas mais gerais, os resultados deste trabalho levam à conclusão de que não se pode falar num *drift* românico generalizado, apesar de as tendências de alguns traços morfossintáticos serem coincidentes entre as línguas estudadas. Por outro lado, em vez de um enfoque configuracional como o previsto pelo modelo sapiriano, que privilegia a alteração teleológica de a para b, a análise sob a perspectiva difusionista aponta aspectos processuais importantes para essas mesmas línguas, como o controle de várias mudanças através do tipo sintático – nos termos deste trabalho, do fator *ambiente oracional*. Além disso, é preciso investigar até que ponto a variação das frequências de uma classe qualquer é influenciada pela oscilação de outra (*default* ou não) ou de uma combinação de proporções de mais de uma categoria – i.e., se há um valor mínimo de X que, associado a um outro mínimo de Y, possa disparar uma mudança em Z. Portanto, a possibilidade de algumas mudanças dependerem de uma variedade de fatores é tema que merece investigação sob o prisma da dinâmica de sistemas complexos.

Para Heine e Kuteva,⁶⁶ a transmissão de caracteres estruturais entre sistemas linguísticos em contato – denominada, por eles, “replicação gramatical” – obedece a restrições severas; apesar de poder atingir qualquer domínio da estrutura linguística, esse processo não envolve preferencialmente estruturas morfossintáticas. Segundo essa proposta, portanto, a sequência de mudanças sintáticas internas a cada sistema tratado neste trabalho não representa um caso de replicação gramatical; tratar-se-ia, mais propriamente, da instalação de uma nova estrutura em sistemas geneticamente relacionados e geograficamente contíguos, a qual teria ocorrido mesmo sem o contato entre as respectivas comunidades de falantes. A proximidade dos sistemas linguísticos apenas reforçaria ou aceleraria uma mudança estrutural em andamento, e não criaria irregularidades nas línguas.⁶⁷ Essa última afirmação está em

⁶⁶ HEINE; KUTEVA. *Language contact and grammatical change*, p. 68 et seq.

⁶⁷ Como exemplificado por *Herbstzeit* > *Zeit des Herbstes* na Bélgica Oriental, por influência do francês *temps d'automne* (RIEHL 2001 *apud* HEINE; KUTEVA. *Language contact and grammatical change*, p. 55).

consonância com a identificação dos fenômenos de mudança aqui estudados com o mecanismo de extensão, e não de reanálise: um mesmo processo – a extensão – operaria, então, internamente a cada sistema e entre sistemas.

Ao discutirem uma série de alterações ocorridas nas línguas europeias ocidentais, Heine e Kuteva⁶⁸ sugerem: a) que essas mudanças primárias sejam o resultado de uma combinação de fatores sistêmicos e externos – causação múltipla –; b) que elas são limitadas à disponibilidade de estruturas semelhantes nas línguas envolvidas; e c) que a difusão entre sistemas seja consequência do prestígio cultural atribuído à comunidade-fonte das mudanças. Essa posição, em consonância com a proposta de Harris e Campbell⁶⁹ e com as observações de Nielsen⁷⁰ sobre o grupo germânico, prevê a operação de mecanismos semelhantes na difusão de uma inovação através do modelo laboviano (para uma comunidade monolíngue) e em situações de contato entre comunidades que falam línguas distintas, geneticamente aparentadas ou não.⁷¹

Sugere-se, por fim, a ampliação das bases de dados de maneira a contemplar outras categorias sintáticas, outros níveis estruturais e outros sistemas linguísticos (a começar pelo francês, no grupo românico, e pelas línguas germânicas, conforme proposto por Nielsen.⁷² A condução desta proposta de trabalho

⁶⁷ Como exemplificado por *Herbstzeit* > *Zeit des Herbstes* na Bélgica Oriental, por influência do francês *temps d'automne* (RIEHL 2001 *apud* HEINE; KUTEVA. *Language contact and grammatical change*, p. 55).

⁶⁸ HEINE; KUTEVA. *Language contact and grammatical change*, p. 94.

⁶⁹ HARRIS; CAMPBELL. *Historical syntax in a cross-linguistic perspective*.

⁷⁰ NIELSEN. *The germanic languages: origins and early dialectal interrelations*.

⁷¹ HOCK; JOSEPH. *Language history, language change and language relationship: an introduction to historical and comparative linguistics*; DIXON. *The rise and fall of languages*; HARRIS; CAMPBELL. *Historical syntax in a cross-linguistic perspective*; HEINE; KUTEVA. *Language contact and grammatical change*; HEINE; KUTEVA. *The changing languages of europe*.

⁷² NIELSEN. *The germanic languages: origins and early dialectal interrelations*.

pode não apenas colaborar decisivamente para a determinação das relações temporais e causais entre as mudanças morfossintáticas sistêmicas, como também para a avaliação da relevância dos contatos linguísticos e de outros fatores externos envolvidos nessas mesmas mudanças.

Resumen: Este trabajo comprende un estudio cuantitativo de *corpora* románicos, con el objetivo de comparar cualitativamente algunas tendencias diacrónicas de organización sintáctica en esas lenguas. Varios trazos fueron seleccionados para la construcción de las base de datos, los cuales también orientaron el análisis de los perfiles de cambio. El objetivo central del trabajo es verificar si esas modificaciones de orden de los constituyentes están acompañadas de otros cambios estructurales y si ellos representan ejemplos de difusión sintáctica. Los resultados muestran que el ambiente oracional es determinante en el control de varios órdenes de cambio; pero ni todos los diversos movimientos aquí identificados caracterizan claramente un proceso de difusión, aun que exigen un análisis de muchos niveles.

Palabras-llave: Lingüística románica; sintaxis; orden de los constituyentes; difusión sintáctica.

Referências

ALIGHIERI, D. De vulgari eloquentia. In: GIULIANI, G. (Ed.) *Le opere latine di Dante Alighieri*. Firenze: Successori Le Monnier, 1878. Disponível em: <gallica.bnf.fr, doc. 84326>.

BASSETTO, B. F. *Elementos de filologia românica*. São Paulo: Edusp, 2001.

BATALI, J. Computational simulations of the emergency of language. In: HURFORD, J. R. et al. (Ed.). *Approaches to the evolution of language*. Cambridge: Cambridge University Press, 1998.

- BIBER, D., CONRAD, S.; REPPEN, R. *Corpus linguistics: investigating language structure and use*. Cambridge: CUP, 1998.
- BOURCIEZ, E. *Éléments de linguistique romane*. Paris: Klincksieck, 1956
- BRILL, E. *et al.* Deducing linguistic structure from the statistics of large corpora. In: PROCEEDINGS of the DARPA Speech and Natural Language Workshop, 1990.
- CANALE, M. Implicational hierarchies of word order relationships. In: CHRISTIE JR., W. M. (Ed.). *Current progress in historical linguistics*. Amsterdam: North-Holland Pub. Co. 1976.
- COHEN, M. A. A. M. *Syntactic change in portuguese: relative clauses and the position of the adjective in the noun phrase*. 1990. Tese (Doutorado) – Unicamp, Campinas, 1990. Inédita.
- COSTA, J. *Word order variation. A constraint-based approach*. 1998. Tese (Doutorado) – Universidade de Leiden, Leiden, 1998.
- DANESI, M. Latin vs. romance in the middle Ages: Dante's De Vulgari Eloquentia revisited. In: WRIGHT, R. (Ed.). *Latin and the romance languages in the early middle ages*. University Park: Penn State University Press, 1996.
- DAVIES, M. Syntactic diffusion in Spanish and Portuguese Infinitival Complements. In: DWORKIN, S. N.; WANNER, D. (Ed.). *New approaches to old problems: Issues in Romance historical linguistics*. Philadelphia: John Benjamins Publ. Co., 2000.
- DE BEAUGRANDE, R. *New foundations for a theory of text and discourse*. Cópia eletrônica, 1997.
- DIXON, R. M. W. *The rise and fall of languages*. 4. ed. Cambridge: CUP, 2002.
- DRYER, M. S. The Greenbergian word order correlations. *Language*, v. 68, n. 1, 1992.
- ELMAN, J. L. Finding structure in time. *Cognitive Science*, n. 14, p. 179-211, 1990. Disponível em: <ucsd.edu//pub/neuralnets/fsit.ps.gz>
- FODOR, I. *The rate of linguistic change*. The Hague: Mouton & Co., 1965.

GONZÁLEZ, J. G. La colocación del adjetivo atributivo en el español medieval: un problema metodológico e histórico. In: XIX CONGRESO INTERNACIONAL DE LINGÜÍSTICA E FILOLOXÍA ROMÁNICAS. Volume de Resumes - Santiago: Universidade de Santiago de Compostela, 1989.

GOERTZEL, B. *Chaotic logic*. 1995a. Disponível em: <altair.comoc.com/People/ben/ben/chlog0.html>. Acesso em: 29 jul. 1997.

GOERTZEL, B. *The miraculous mind attractor*. 1995b. Disponível em: <altair.comoc.com/People/ben/ben/cm0.html>. Acesso em: 29 jul. 1997.

GREENBERG, J. Some universals of grammar with particular reference to the order of meaningful elements. In: GREENBERG, J. (Ed.). *Universals of language*. MIT Press, 1966.

HARRIS, A.; CAMPBELL, L. *Historical syntax in a cross-linguistic perspective*. Cambridge: CUP, 1995.

HAWKINS, J. A. Implicational universals as predictors of word order change. *Language*, v. 55, n.3, p. 618-648, 1979.

HAWKINS, J. A. *Word order universals*. New York: Academic Press, 1983.

HAWKINS, J. A. Seeking motives for change in typological variation. In: W. CROFT; K. DENNING; S. KEMMER (Ed.). *Studies in typology and diachrony for Joseph H. Greenberg*, 95-108 (Typological Studies in Lgg, 20). Amsterdam: John Benjamins, 1990.

HEINE, B.; KUTEVA, T. *Language contact and grammatical change*. Cambridge: CUP, 2006a.

HEINE, B.; KUTEVA, T. *The changing languages of Europe*. Oxford: OUP, 2006b.

HOCK, H. H.; JOSEPH, B. D. *Language history, language change and language relationship: an introduction to historical and comparative linguistics*. New York: Mouton, 1996.

KAUFMANN, S. *The Origins of order: self-organization and selection in evolution*. New York: Oxford UnPress, 1993.

KIRBY, S. *Function, selection and innateness: the emergence of language universals*. University of Edinburgh, 1996 (PhD Thesis). Disponível em: <<http://www.ling.ed.ac.uk/%7Esimon/publications.html>>. Acesso em: 7 dez. 2006.

KRAHE, H. *Grundzüge der vergleichenden Syntax der indo-germanischen Sprachen*. Innsbruck: UnivInnsbruck, 1972.

KROCH, A. Morphosyntactic variation. In: BEALS, K. (Ed). *Papers from the 30th Regional Meeting of The Chicago Linguistics Society*. 1994.

KROCH, A. *Syntactic change*. 2000. Disponível em: <ftp://ling.upenn.edu/papers/faculty/tony_kroch/papers/diachronic-syntax-99.pdf>. Acesso em: 15 dez. 2006.

LEECH, G. et al. *Recommendations for the syntactic annotation of corpora*. 1996. Disponível em: <ftp.ilc.pi.cnr.it/pub/eagles/corporasag1.ps.gz>

LEHRER, A. *Semantic fields and lexical structure*. Amsterdam: North-Holland Publ. Co., 1974.

LIGHTFOOT, D. *How to set parameters: arguments from language change*. Cambridge: MIT Press, 1991.

MAROUZEAU, J. *L'ordre des mots dans la phrase latine*. Paris: Champion 1922. v. 1: Les groupes nominaux.

MARTINS, A. M. The Loss of IP-Scrambling in Portuguese: clause structure, word-order variation and change. In: LIGHTFOOT, D. (Ed.). *Syntactic effects of morphological change*. Oxford: OUP, 2002.

MATTOS E SILVA, R. V. *O português arcaico: fonologia, morfologia e sintaxe*. São Paulo: Contexto, 1992.

McMAHON, J. & SMITH, F.J. A review of statistical language processing techniques. *Artificial Intelligence Review*, v. 12, n. 5, pp. 347-391, oct. 1998. Disponível em: <cs.qub.ac.uk/~J.McMahon/paper3.ps>

MOLLICA, M. C. M. *Queísmo e dequeísmo em português*. Rio de Janeiro: UFRJ, Faculdade de Letras, 1992. Tese de Doutorado. Inédita.

NARO, A.; LEMLE, M. Syntactic diffusion. *Ciência e Cultura*, v. 29, n. 3, 1976.

NIELSEN, H. F. *The germanic languages: origins and early dialectal interrelations*. Tuscaloosa: University Alabama Press, 1989.

NIYOGI, P.; BERWICK, R. C. *The logical problem of language change*. MIT, 1995. (Artificial Intelligence Laboratory Memo 1515) Disponível em: <ftp://publications.ai.mit.edu/ai-publications/1500-1999/AIM-1516.ps.Z > Acesso em: 29 jul. 1997.

OSTLER, N. *Empires of the word: a language history of the world*. New York: Harper Collins PubCo, 2005.

POSNER, R. *The romance languages*. Cambridge: CUP, 1996.

RIJKHOFF, J. Explaining word order in the noun phrase. *Linguistics*, n. 28, 1990.

RENZI, L. *Nuova introduzione alla filologia romanza*. Milano: IL Mulino, 1994.

RUHLEN, M. *A Guide to the world's languages*. 3rd ed. Stanford: Stanford UnPress, 2000. v. 1: Classification.

SAPIR, E. *Language: an introduction to the study of speech*. New York: Harcourt & Brace, 1920.

SHEN, Z. *Lexical diffusion: a population perspective and a numerical model*. Berkeley: University of California, 1991.

TAGLIAVINI, C. *Le origini delle lingue neolatine*. 4. ed. Bologna: Casa Ed. Prof. Ricc. Pàtron, 1964.

TESITELOVÁ, M. D. *Quantitative linguistics*. New York: John Benjamins, 1992.

THOMASON, S. G.; KAUFMAN, T. *Language contact, creolization and genetic linguistics*. Berkeley: University California Press, 1991.

TOTARO, J. H. R. *Mudança de ordem dos constituintes A e N em textos espanbóis dos séculos XIII a XX*. Belo Horizonte: Faculdade Letras da UFMG, 1998. (Dissertação, Mestrado). Inédita.

VIDOS, B. E. *Manual de lingüística românica*. Rio de Janeiro: Eduerj 1996.

VINCENT, N. Competition and correspondence in syntactic change: null arguments in latin and romance. In: PINTZUK, S.; TSOULAS, G.; WARNER, A. (Ed.). *Diachronic syntax: models and mechanisms*. Oxford: Oxford University Press, 2000.

VON WARTBURG, W. *Evolución y estructura de la lengua francesa*. Madrid: Ed. Gredos, S.A., 1966.

WAUGH, L. *A semantic analysis of word order*. Leiden: E. J. Brill, 1977.

WILMET, M. *La détermination nominale*. Paris: Presses Universitaires de France, 1986.

WOLFF, P. *Origen de las lenguas occidentales, 100-1500 d.C.* Madrid: Ed. Guadarrama 1971.

WOODS, A., FLETCHER, P. & HUGHES, A. *Statistics in language studies.* Cambridge: CUP, 1986.

Recebido para publicação em fevereiro de 2009.

Aprovado em abril de 2009.