

Aspectos diacrônicos da ordem de palavras
em línguas românicas

Condicionamentos morfológicos, lexicais e sintáticos da mudança de ordem
de constituintes em textos espanhóis, italianos e portugueses
sob a perspectiva da difusão sintática

Tese apresentada ao programa de doutoramento em *Estudos Lingüísticos*
(linha B - *Estudo da Variação e Mudança Lingüística*)
da FaLe / UFMG

João Henrique Rettore Totaro

Orientador: Prof. Dr. Marco Antônio de Oliveira

27.04.2007

Agradecimentos

Um trabalho como este é produto de esforços conjugados, e minha gratidão é dirigida a todos os que colaboraram para que ele chegasse a termo.

A Deus, em todas as Suas formas de manifestação.

A meus pais, pela orientação na formação de meus valores e pelo desenvolvimento destes; a Cristiane, Catarina e Sofia, a família que constituí, a luz de meus olhos: elas vivenciaram emocionalmente cada passo deste caminho, e sua conclusão deve muito a elas. Grande parte deste trabalho foi feita nas horas que vêm depois das horas do dia, a fim de não tomar mais tempo à família, mas muitas vezes tive de valer-me de sua compreensão e de seu apoio incondicionais. A Ana Lúcia, Claudia e Frederico, que completam magistralmente essa família.

Ao Prof. Dr. Marco Antônio de Oliveira, Orientador na plena acepção da palavra, com quem o estudo e o trabalho são experiências gratificantes e extremamente enriquecedoras. Sua disponibilidade irrestrita e sua maneira exigente, crítica e criativa de refletir e de discutir em detalhe os vários aspectos de que se reveste não apenas refinaram o conteúdo como a forma do texto. Este trabalho pôde amadurecer e materializar-se graças a sua perspicácia e à pertinência de suas intervenções, da parte acadêmica, e, em termos pessoais, a seu lado profundamente humano diante das muitas pausas que nos são impostas.

À Faculdade de Letras da UFMG e a seu Programa de Pós-Graduação em Estudos Linguísticos, por terem acolhido o projeto inicial e por terem oferecido o ambiente acadêmico necessário para desenvolvê-lo; aos professores das disciplinas que cursei no doutorado, pela oportunidade única de discutir argumentos e de avaliar instrumentos imprescindíveis à elaboração deste estudo. À Biblioteca da FaLe, nas pessoas eficientes de Rosângela Bernardino e de Nina Castro.

À Pontifícia Universidade Católica de Minas Gerais, pelo subsídio oferecido a este trabalho através da redução temporária de minha carga horária em sala de aula, prevista pelo PPCD – Programa Permanente de Capacitação Docente –, e às Coordenações de curso das duas unidades em que trabalho, pelo suporte no redimensionamento de minhas atribuições.

Aos colegas e amigos da FaLe e da PUC, em especial aos profs. Drs. Rui Rothe-Neves e Gilberto Xavier da Silva, pelo companheirismo que ultrapassa os portões da Universidade. À Prof^a Dr^a Vanda de Oliveira Bittencourt, pela introdução aos estudos filológicos, cujo fruto aqui se colhe, e pela amizade de todas as horas – não apenas as canônicas..

Aos que se foram, aos que voltaram, aos que chegaram, aos que trabalham em segredo a nosso favor, e sem cuja presença eu não seria quem sou; e àqueles cuja participação foi pequena – pois é também de pequenas coisas que nos construímos: obrigado pela companhia nessa trajetória em direção à Luz, destino de todos nós.

Índice

<i>Tópico</i>	<i>pág.</i>
I - Introdução	02
II - Objetivos / formulação do problema	07
III - Justificativa e Perspectivas	10
IV - Pressupostos teórico-metodológicos	11
V - Metodologia	19
V.1 – Tratamento dos textos	20
V.2 – Montagem das bases de dados	24
V.3 – Métodos estatísticos	27
V.4 – Outros cálculos	32
V.5 – Critérios de classificação das ocorrências	35
VI.1 – Tratamento numérico das ocorrências: o <i>corpus</i> português	51
VI.2 - Tratamento numérico das ocorrências: o <i>corpus</i> espanhol	125
VI.3 - Tratamento numérico das ocorrências: o <i>corpus</i> italiano	222
VII – Comparações entre os <i>corpora</i>	289
VIII – Considerações finais	321
IX - Referências bibliográficas	331

Lista de abreviaturas

A	Adjetivo
N	substantivo - núcleo do sintagma nominal
Rel	orações relativas

NA	ordem preferencial adjetivo-substantivo
GN	ordem preferencial genitivo-substantivo
NA	ordem preferencial substantivo-adjetivo
NG	ordem preferencial substantivo-genitivo
Po	utilização preferencial de posposições
Pr	utilização preferencial de preposições

SN	sintagma nominal
SV	sintagma verbal
CN	complementos do núcleo do sintagma nominal
CV	complementos do núcleo do sintagma verbal

OV	ordem preferencial objeto-verbo
VO	ordem preferencial verbo-objeto
SCV	ordem preferencial sujeito-objeto-verbo
SVC	ordem preferencial sujeito-verbo-objeto
VSO	ordem preferencial verbo-sujeito-objeto

EIP	(línguas) espanhol(a), italiano(a) e português(esa)
-----	---

Lista de textos espanhóis

JUA - D. Juan Manuel - Libro del caballero et del escudero - Biblioteca de Autores Españoles - tomo LI - Madrid: Ediciones Atlas 1947.

ENX - El libro de los enxemplos - Biblioteca de Autores Españoles - tomo LI - Madrid: Ediciones Atlas 1947.

COL - Relaciones, cartas y otros documentos concernientes a los cuatro viajes que hizo el Almirante Don Cristobal Colón para el descubrimiento de las Indias Occidentales - Biblioteca de Autores Españoles - tomo LXXV- Madrid: Ediciones Atlas 1949.

DVM - Tratado en defenssa de virtuossas mugeres - Biblioteca de Autores Españoles - tomo CXVI - Madrid: Ediciones Atlas 1956.

IND - Historia general de las Indias - Primera parte - Biblioteca de Autores Españoles - tomo XXII - Madrid: Ediciones Atlas 1946.

AGR - Vida de Sor María de Ágreda - Biblioteca de Autores Españoles - tomo CIX.

CON - Alonso de Contreras - Autobiografías de soldados - Biblioteca de Autores Españoles - tomo XC - Madrid: Ediciones Atlas 1952.

JFI - José Francisco de Isla - Historia del famoso predicador Fray Gerundio de Campazas, alias Zotes - *in* Rio, A. - Antología General de la literatura española - New York: The Dryden Press 1954.

FEL - Marques de San Felipe - Comentarios - Biblioteca de Autores Españoles - tomo IC - Madrid: Ediciones Atlas 1954.

TOR – Torres Villaroel, D. – Vida (1743) – Madrid: Espasa-Calpe, 1941.

C&M - Costas y montañas - Biblioteca de Autores Españoles - tomo XCIII - Madrid: Ediciones Atlas 1953.

Lista de textos portugueses

Leal Conselheiro – o qual fez Dom Eduarte... – 1438 – Piel, J. M. (ed.) – Lisboa: Livr. Bertrand, 1942.

CEU – Zurara, Gomes Eanes de – Crônica da Tomada de Ceuta – ed. Alfredo Pimenta – Lisboa: LCE, 1964.

OLIVEIRA, Fernão de – Grammatica da lingoagem portuguesa Lisboa : e[m] casa d'Germão Galharde, 1536 – documento eletrônico (purl.pt/369/impressos.html).

GÂNDAVO, Pero Magalhães – Historia da prouincia sa[n]cta Cruz a qui' vulgarme[n]te chamam Brasil : feita por Pero de Magalhães de Gandauo, dirigida ao muito Illsre s[e]nor Dom Lionis P[ereir]a governador que foy de Malaca e das mais partes do Sul na India – Lisboa: Antonio Gonsaluez, vendense em casa de João lopez liureiro na rua noua, 1576 – documento eletrônico (purl.pt/369/impressos.html).

Leão, Duarte N. – Orthographia da lingoa portuguesa : obra vtil & necessaria assi pera bem screuer a lingoa Hespanhol como a Latina & quaesquer outras que da Latina teem origem ; Item hum tractado dos pontos das clausulas / pelo licenciado Duarte Nunez do Lião – Lisboa: João de Barreira, 1576 – documento eletrônico (purl.pt/15)

Brandão, Fr. Antônio – Crônica do Conde D. Henrique, D. Teresa e Infante D. Afonso – 1632 – Porto : Tip. e Enc. A Portuense, 1944.

Lobo, Pe. Jerônimo – Itinerário e outros escritos inéditos – 1640 – *in* Costa, Pe. M. Gonçalves (ed.) – Barcelos: Comp. Ed. do Minho, 1971.

Crônica do Sereníssimo Príncipe Dom João – 1723 – documento eletrônico (bnd.bn.pt/od).

Oliveira, Francisco X. – Viagem à Ilha do Amor – 1744 – *in* Anastácio, Vander (ed.) – Porto: Ed. Caixotim, 2001 – Col. Obras Clássicas da Literatura Portuguesa.

Queirós, Eça – Os Maias – 1888 – documento eletrônico (bnd.bn.pt/od).

Alexandre Herculano – Lendas e narrativas – 1858 – documento eletrônico (bnd.bn.pt/od).

Saramago, J. – A jangada de pedra – Lisboa: Ed. Caminho, 1988.

Lista de textos italianos

POG – Le Pogge – Comento di Jacopo di Messer Poggio sopra il triumpho della fama di Messer Francesco Petrarca,... – Roma, 1477 – documento eletrônico (gallica.bnf.fr).

MES – Libro della consolatione delle medicine semplici solutiui il quale fece Giouanni figliuolo di Mesuë – 1492 – documento eletrônico (gallica.bnf.fr).

GUE – Accolti, Benedetto – La guerra fatta da Christiani contra Barbari per la ricuperatione del sepolchro di Christo et della Giudea – Vinegia, 1549 – documento eletrônico (gallica.bnf.fr).

BAG – Minardo, Ventura – Compendio delle regole contenute ne gli Eccellentiss. Autori, che de'bagni di Caldiero nel territorio Veronese hanno scritto del modo di vsar dette acque & fango ; & d'altri particolari auisi à quelli, che vi si conducono – Venetia, 1571 – documento eletrônico (gallica.bnf.fr).

CHR – Discorso historico circa la patria, e nascita di Christoforo Colombi scopritore del Mondo Nuouo – Pietro Maria Campi, *in* Campi, P. M. – Historia universale così delle cose ecclesiastiche come secolari di Piacenza et altre città d'Italia – parte terza – 1662 – documento eletrônico (gallica.bnf.fr).

TOS – Benedetto Buommattei – Della lingua Toscana – Firenze, 1643 (3^a. impressione) – documento eletrônico (gallica.bnf.fr).

SCI – Vico, Giambattista - Cinque libri di Giambattista Vico de'Principj d'una scienza nuova d'intorno alla comune natura delle nazioni – Napoli, 1730 (2^a ediz.) – documento eletrônico (gallica.bnf.fr).

ITA – Schott, Francesco – Itinerario d'Italia – Roma, 1747 – documento eletrônico (gallica.bnf.fr).

ETN – Borsari, Ferdinando – Geografia etnologica e storica della Tripolitania, Cirenaica e Fezzan : con cenni sulla storia di queste regioni e sul silfio della Cirenaica – Napoli, 1888 – documento eletrônico (gallica.bnf.fr).

UBC – Lombroso, Cesare – L'uomo bianco e l'uomo di colore : letture su l'origine e la varietà delle razze umane – Turin, 1892 – documento eletrônico (gallica.bnf.fr).

ECO – Eco, Umberto – Il nome della rosa – Milano: Gr. Ed. Fabbir, 1980.

I - Introdução

A noção de que algumas línguas naturais tendem a colocar, de maneira consistente, elementos modificadores ou limitantes antes dos modificados ou limitados, ao passo que outras se comportam de maneira oposta, é bastante difundida entre os lingüistas, tendo já sido realizado um grande número de trabalhos (Marouzeau 1922; Bolinger 1952; Greenberg 1966; Krahe 1972; Canale 1976; Waugh 1977; Hawkins 1979 1983; Rijkoff 1990; Dryer 1992; Whaley 1997 etc.) que não apenas tratam descritivamente do tema, mas oferecem propostas de explicação para essas tendências.

Um ponto recorrente em obras que tratam de lingüística 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 européias em três grandes famílias (grego, germânico e românico); para este último, 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*)¹:

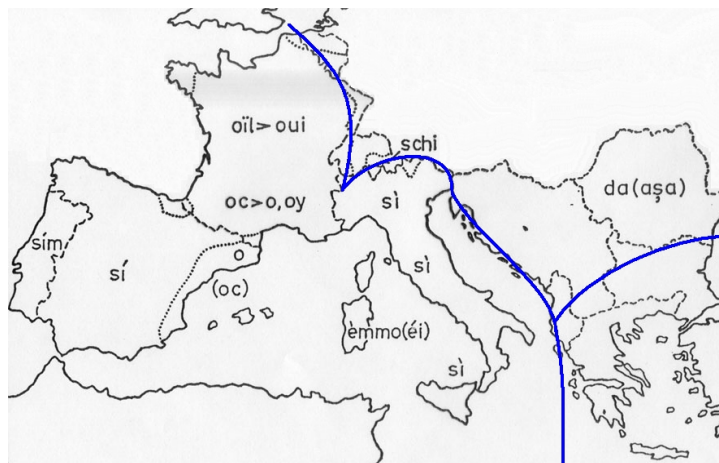


Fig. FR1 – situação das regiões românicas segundo a proposta de Dante Alighieri (1305/1878) [<http://www.fb10.uni-bremen.de/romanistik/homepages/burr/Intro/Klassifizierung.htm>, modificado segundo Ostler 2005].

¹ “Nam alii *Oc*, alii *Oil*, alii *Si* affirmando loquuntur, ut puta Hispani, Franci et Latini.”, a partir de Giovanni 1878:29ss. Não obstante o volume de trabalhos que fazem remontar a Dante o princípio dos estudos em filologia românica, Danesi *op.cit.* observa que é o fato de Dante utilizar a expressão *ydioma tripharium* (na edição de Giovanni *op. cit.*, *idioma trifarium*) na DVE que leva os estudiosos a considerarem a raiz de uma teoria genética para as línguas românicas; sob essa perspectiva, Danesi propõe uma releitura e reavaliação dessa obra – não como o fundamento da filologia românica ou da lingüística histórica, mas como uma obra sobre a linguagem poética.

Posner 1996 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. Rainouard 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.

O primeiro tratado sistemático sobre a questão românica teria sido publicado em 1831 por L. Diefenbach (Posner *op. cit.*), 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 crioulizado'. 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-européias e por J. Grimm para as germânicas (Bassetto 2001). Diez dividiu o espaço românico em porções oriental e ocidental – uma proposta que foi posteriormente desenvolvida por W. von Wartburg 1966 e vários outros, como, p.e., 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 (Renzi 1994)).

Ao final do século XIX e no início do XX, os franceses deixavam a postura regularista e reconstrucionista em filologia românica para se concentrarem mais na dialetologia e nas abordagens sociológica e pragmática das línguas individuais; entretanto, a despeito das observações de Saussure sobre a relevância da filologia românica, poucas foram as adições importantes à Romanística na primeira metade do século XX (Posner *op.*

cit.). Nesse contexto, sobressai A. Meillet, que trata tanto dos aspectos estruturais das línguas indo-européias e de seus dialetos quanto da aplicação do método histórico-comparativo em lingüística histórica, e que formou É. Benveniste, M. Cohen, A. Martinet, L. Tesnière e J. Vendryes.

Temos, posteriormente, E. Bourciez 1956, 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 em seu §32) 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 Itallorromânia – 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 balcano-românico a partir de argumentos estruturais e culturais (1996:228-229). Ainda que sejam pertinentes as considerações feitas por Ruhlen 2000 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 (1996: 242ss e, em particular, 248), que deixa no primeiro domínio apenas os Balcãs e a Dácia.

No plano diacrônico, diversos estudiosos (Cohen 1986/1990 1988 1995; González 1989) ressaltam que, a partir da leitura de textos escritos em línguas diferentes e épocas diversas, bem como de gramáticas históricas que tratam descritivamente de momentos pretéritos da evolução de sistemas lingüísticos 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 (Cohen, *op. cit.*, para o português; González, *op. cit.*, para o 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 (Tagliavini 1964; Wolff 1971).

No que respeita à determinação das características sintáticas do grupo românico, observamos, inicialmente, que Perini *et alii* (*mimeo* 1994), numa exposição incisiva e

detalhada dos caracteres sincrônicos do SN português, também reconhecem e discutem a relevância de uma série de restrições sintáticas, semânticas e discursivas que, em conjunto, caracterizam, a distribuição diferencial dos elementos dentro do que chamam de “área esquerda” (pré-nominal) e “área direita” (pós-nominal) do SN; a análise da primeira seria facilitada pelo fato de muitas funções serem exercidas por itens lexicais especializados e pertencentes a classes fechadas, ao passo que os itens lexicais que ocorrem na área direita são altamente polivalentes e pertencem a classes abertas (Perini 1996: 101). Fazendo uma descrição da estrutura sincrônica do SN português contemporâneo, Perini (1996: 95-106, 233-234) distingue “nove posições fixas, estritamente ordenadas”, cuja alteração de ordem corresponde a uma mudança de função. A lista dessas posições / funções é aqui reproduzida:

- Determinante - *o, este, algum, etc.*
- Possessivo - *meu, seu, nosso, etc.*
- Reforço - *mesmo, próprio, certo.*
- Quantificador - *poucos, vários, diversos, muitos, único, primeiro, etc.*
- Pré-Núcleo Externo - *mero, pretenso, meio, suposto, reles, inesquecível, ilusório, simples, bom, velho, etc. (classe aberta)*
- Pré-Núcleo Interno - *mau, novo, velho, claro, grande, etc.*
- Núcleo
- Modificador Interno - (classe aberta)
- Modificador Externo - pode incluir não apenas palavras isoladas como sintagmas maiores.

Nesse contexto e do ponto-de-vista diacrônico, é possível observar em textos portugueses (Cohen 1986/1990), espanhóis (Totaro 1998) e catalães (Lima 2003) uma alteração na posição relativa de adjetivos simples em relação ao N. Em Totaro *op. cit.*, chamam-se complementos adjetivais simples àqueles compostos unicamente de adjetivos intransitivos, ou seja, que não exigem complementos preposicionados; assim, são simples *grandes ejércitos, formidables escaramuzas e bonetes colorados*, ao contrário de *hombres ansiosos por batallar* e *bateles llenos de cerveza*. Várias ocorrências mostram que o pronome possessivo segue o núcleo do sintagma nominal, enquanto a colocação preferencial de adjetivos simples é anterior ao mesmo; em outras, as posições de diversos

satélites do núcleo do SN são algo mais flexíveis em relação aos contemporâneos (Tabela A). Esse comportamento pode mesmo encontrar paralelos em estágios evolutivos diversos das outras línguas neolatinas, o que demandaria uma análise mais cuidadosa de textos que permitam abranger um largo intervalo de tempo. Entretanto, partindo do latim vulgar até as línguas espanhola, italiana e portuguesa (EIP) modernas, 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.

Tomem-se, a título de ilustração, as estruturas constantes da Tabela B abaixo, colhidas nos textos que serviram de base à constituição do *corpus* espanhol em Totaro *op. cit.*. Apesar de as mesmas serem menos freqüentes do que o cânone SVC, tanto nesses textos como no espanhol contemporâneo, suas proporções *relativas* podem ter sido diferentes em outros séculos a partir do latim SCV, passando por um período de instabilidade tipológica mais pronunciada antes de chegar aos valores atuais. Parte da literatura teórica disponível (cf., especialmente, Greenberg 1966 e Canale 1976) 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 outros². 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.

<i>texto</i>	<i>século</i>	<i>pág-parágr-linha</i>	<i>ocorrência</i>
JUA	XIV	237-2-04	et non haber otra manera ninguna buena
COL	XV	103-02-30	y después partir para otra isla grande mucho
DVM	XV	57-02-23	El espíritu bueno tuyo me trairá en la carrera derecha
DVM	XV	61-01-55	Engañado por mostrar el arte tuyo luxorioso
IND	XVI	161-02-52	la tierra que se habita es isla cercada toda del Océano
AGR	XVII	227-01-01	Vasallos del Rey de los siglos inmortal
CON	XVII	087-02-46	hoy día tiene una casa bien buena , labrada a mi costa
FEL	XVIII	189-02-27	las confirmaba con nunca intermitentes instancias
FEL	XVIII	193-01-43	sacó una no leve herida
C&M	XIX	329-02-10	Exponiendo el sin rival calzado al filo de las rocas

Tabela A - exemplos de ordens não-usuais de colocação dos satélites do N em textos históricos espanhóis.

² Há que se considerar, naturalmente, a imobilidade de SPs e de orações relativas, associados, na classificação de Perini (1996) acima apresentada, ao grupo de *modificadores externos*.

ENX	XIV	447-01-04	E de que <u>esto</u> <u>vió el confessor</u> , díjole	<u>CVS</u>
CON	XVII	77-02-41	En suma, se <u>dió tan buena maña el padre</u> , que	<u>VCS</u>
JFI	XVIII	35-01-01	<u>Parió</u> , pues, <u>la tía Cartuja un niño</u> como unas flores	<u>VSC</u>
JFI	XVIII	36-01-06	<u>dió el niño grandes señales</u> de lo que había de ser	<u>VSC</u>

Tabela B - exemplos de ordens não-canônicas dos constituintes S-V-C em textos históricos espanhóis.

II - Objetivos / formulação do problema

Com base no exposto acima, 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 português (Cohen 1986/1990) e espanhol (Totaro 1998). 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 lingüísticos; 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 Cohen *op. cit.* e por Rijkhoff 1990.

Nesse diapasão, os objetivos centrais do trabalho foram i) a construção e a codificação das bases de dados segundo os critérios de análise propostos na Seção V.3³; ii) a investigação da possibilidade de as línguas em questão terem sofrido mudanças na ordenação de constituintes semelhantes àquelas apontadas por Cohen *op. cit.*, para a língua portuguesa, Totaro 1998, para a língua espanhola e Lima 2003, para o catalão; e iii) testar a hipótese de que, em um ou mais desses sistemas, essa mudança não foi independente, inserindo-se num quadro mais geral de evolução das línguas neolatinas.

O ponto ii) implica estender os princípios do modelo de difusão lexical (Shen 1991) ao nível sintático, *mutatis mutandis*. No caso das línguas neolatinas aqui estudadas, pretende-se estudar se a mudança posicional dos constituintes se processou de forma *sintaticamente abrupta e lexical ou sintagmaticamente gradual*, e que condicionamentos estruturais responderiam por esse processo de *difusão sintática*. Naro & Lemle 1976

³ Reflete-se, aqui, tanto uma preocupação com o encaixamento do fenômeno em apreço quanto a necessidade de se confirmar a aplicação da proposta de Hawkins (1979, 1983), sob a forma de um Princípio de Harmonia Trans-Categorial (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).

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 & Lemle *op. cit.* 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 Silva 1992 sobre as estruturas atributivas do português arcaico, as quais ilustram percursos da difusão sintática entre os séculos XIV e XV. Ainda para o português, Mollica 1992 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 & Lemle *op. cit.*, 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.

Naro & Lemle 1976:259, 267

Minha proposta neste trabalho envolve o estudo diacrônico das línguas EIP, privilegiando os aspectos sintáticos da associação dos elementos do SV – sem perder de vista a relação entre a sintaxe e os demais níveis de organização da língua. Na análise dos dados numéricos, os resultados serão confrontados com aqueles oferecidos, por exemplo, a) por Cohen 1986/1990 e por Totaro 1998 para os adjetivos no SN português e espanhol, respectivamente, e b) por Bittencourt 1995, Davies 1992 e Davies 2000, sobre as estruturas causativas em português e em português e espanhol, respectivamente, a fim de verificar a existência de semelhanças e dessemelhanças tipológicas instaladas ao longo do tempo nesses sistemas geográfica, lingüística e culturalmente próximos⁴.

⁴ A esse respeito, veja-se, por exemplo, Paiva Boléo (1974-1975: 355 – 398), sobre as influências mútuas entre as línguas portuguesa e espanhola, e Thomason & Kaufman 1991, sobre as relações e interferências mútuas entre línguas provenientes de uma raiz comum.

Para tanto, são tomados textos escritos em um largo intervalo de tempo a fim de observar, de maneira mais patente, quaisquer mudanças ocorridas na colocação desses elementos na modalidade escrita dessas línguas, conforme a metodologia de base empírica descrita nas seções IV e V adiante e que reflete a postura de Bynon 1977, nominalmente transcrita abaixo:

... it is possible to abstract the grammatical structure of the language of each period from the documents and in this way a series of synchronic grammars may be set up and compared. The differences in their successive structures may then be interpreted as reflecting the historical development of the language.

(Bynon, *op. cit.*: 2; grifos meus)

... it is only by comparison via written records of different stages of a language that [the historical linguist] can obtain an idea of the nature of diachronic rules. It can, in fact, be said that there is an optimal time-lapse of say four or five centuries which is most favourable for the systematic study of change. This is so because on the one hand the differences between successive language states are then sufficiently large to allow the statement in the form of rules of completed changes and on the other continuity is not at stake - one is clearly still dealing with 'the same language'.

(Bynon, *op. cit.*: 6; grifos meus)

Conforme ressaltado acima, todos os resultados da análise diacrônica conduzida neste trabalho terão suporte na coleta e na descrição de dados reais de língua produzidos em diversos recortes sincrônicos sucessivos, procurando-se determinar a existência (ou não) de fases de transição tipológica, situando-as no tempo (cf. proposto por Adrados 1999 e Schmidt 1999).

O desenvolvimento e a conclusão deste trabalho visam, numa perspectiva mais abrangente, (i) fornecer argumentos diversos para pesquisas futuras, atendendo ao suprimento de dados, hipóteses e análises relativos à mudança (ou manutenção) da posição dos complementos nominais nos sintagmas nominal e verbal ao longo do tempo nas línguas em questão, além de procurar detectar motivações para que estas tenham sido estruturadas da maneira como o foram e as possíveis conseqüências de tal estruturação; (ii) coletar e sistematizar dados sobre o comportamento dos sistemas lingüísticos estudados.

III - Justificativa e perspectivas

As gramáticas históricas e os trabalhos sobre a história das línguas românicas consultados (Aldrete, 1606; Hanssen 1945; Lapesa 1962, 2000; Pidal 1949, 1956; Alonso 1972; Company 1992; Llorach 1995 e Penny 1995, para o espanhol; Brunot 1947; Brunot & Bruneau 1949; von Wartburg 1966; Bauer 1995, para a língua francesa; Moll 1952; Lima 2003, para o catalão; Vidos 1968, 1996, para o grupo românico) fazem referência generalizada à ordem CN/N⁵ como sendo dominante nas mesmas em certos períodos de sua evolução. Entretanto, limitando-se à caracterização sucinta desses sistemas lingüísticos no período medieval e não se permitindo um aprofundamento em tópicos específicos, essas obras não oferecem ao leitor uma apreciação diacrônica (sob a perspectiva de uma língua específica) e comparada (relacionando duas ou mais línguas) da questão da anteposição ou posposição dos CNs e CVs aos seus respectivos núcleos.

Em termos da tipologia lingüística discutida por Greenberg 1966 e associada à ordem dos constituintes como acima mencionado, a definição das tendências gerais de posicionamento dos constituintes S-V-C / A-N e a caracterização de sua evolução por um período relativamente longo do tempo real podem contribuir para um melhor entendimento do comportamento das línguas e dos fatores que podem atuar em alguns processos de mudança lingüística. Daí surgiu a proposta de que se reveste este trabalho, que pretende investigar se, de fato e a partir da combinação de fatores de diversas ordens, é possível identificar um movimento de difusão sintática nas línguas românicas.

Além de trazer alguma contribuição ao conhecimento que se tem hoje da evolução particular das línguas em questão e, numa perspectiva mais ampla, das línguas românicas, este trabalho pretende enfatizar a necessidade de se conduzirem desdobramentos e novos estudos sobre o assunto. O estudo da relevância de possíveis condicionamentos da mudança também pode auxiliar na compreensão dos processos e caminhos segundo os quais a alteração da ordem de constituintes se dá (cf. observações feitas a esse respeito por Mariz e Pádua 1960). Com esse propósito, contemplam-se de forma comparativa as línguas amostradas, a fim de se investigar seu comportamento em relação à colocação dos constituintes S-V-C numa abordagem diacrônica de base empírica e segundo um modelo

⁵ v. Nota 2 acima.

semelhante ao proposto por Shen 1991 para a difusão lexical, qual seja o de Naro & Lemle 1976.

IV - Pressupostos teóricos

Greenberg 1966 apresenta uma tipologia sintática que pretende responder pelas tendências principais de organização dos sistemas lingüísticos, aqui exposta de maneira extremamente resumida: segundo esse autor, as línguas naturais podem ser classificadas de acordo com a organização de seus elementos modificadores e modificados, o que se expressa tanto em termos do uso de preposições ou de posposições (Pr/Po), quanto de relações posicionais para os elementos sujeito/verbo/objeto (S-V-O), adjetivo/substantivo (A-N no SN) e possessivo/núcleo (G-N), dentre outros constituintes. A partir desses elementos e de sua análise combinatória, haveria cerca de 24 tipos lógicos possíveis para as línguas; quatro deles, entretanto, responderiam pela grande maioria das línguas (Comrie 1981), a saber: VSO/Pr/NG/NA, SVC/Pr/NG/NA, SCV/Po/GN/AN e SCV/Po/GN/NA. Em vista do grande número de inconsistências decorrentes de uma formulação desse tipo para sistemas naturais, os trabalhos que seguem a diretriz lançada por Greenberg (*op. cit.*) tratam esses universais lingüísticos em termos de *tendências*, e não como absolutos.

Partindo desse ponto, Lehmann 1973 e Vennemann 1974 defendem, respectivamente, um *Princípio Estrutural da Linguagem* e um *Princípio de Serialização Natural*, segundo os quais os modificadores são colocados, em relação a um elemento sintático básico (o termo modificado), no lado oposto ao de seu concomitante primário (V); em outras palavras, se a ordem não-marcada numa língua é SVC, deve-se prever, por exemplo, que a mesma seja NA/NG/Pr, e que, se a ordem for SCV, a língua deve apresentar um padrão AN/GN/Po. De maneira menos rígida, a implicação entre as ordens de colocação de constituintes diversos é reavaliada por Hawkins 1979, 1983, que propõe haver, nas línguas, uma *tendência* a respeitar uma mesma proporção de anteposição / posposição de operadores em relação a seus respectivos operandos. Deve-se, ainda, considerar a exposição feita por Canale 1976, na qual o autor argumenta que as mudanças processadas no SN ou no SV não têm, *necessariamente*, implicações externas a esses sintagmas, mas acarretam apenas modificações internas aos mesmos - o que também merece verificação no intervalo de amostragem aqui proposto.

Por outro lado, Comrie 1981, apesar de reconhecer que, em linhas gerais, os parâmetros S-V-O, N-A, N-G e Pr/Po são logicamente independentes uns dos outros, observa, ao comentar a tipologia lingüística proposta por Greenberg 1966 e seus seguidores e as contribuições feitas por Lehmann 1972, que à ordem SVC, ao contrário das demais, não são facilmente correlacionáveis quaisquer outros parâmetros (Comrie, *op. cit.*) - o que equivale a dizer que os graus de liberdade em sistemas SVC são muito maiores do que nas demais combinações, ainda que a utilização de parâmetros numa determinada ordem deva ser harmônica com a de outros.

Também a partir da análise de Greenberg 1966, Canale 1976 afirma ser possível esperar que uma série de hierarquias implicacionais de ordem de constituintes ligados ao SN e ao SV estariam refletidas na seqüência cronológica de mudança sintática na passagem de uma conformação dominante SCV para SVC. Segundo o autor, seria impossível que essas mudanças ocorressem em massa dentro do SN antes que o primeiro dos padrões ligados ao SV mudasse, e vice-versa; entretanto, ao contrário de Givón 1984, Lehmann 1973 e Vennemann 1974, Canale 1976 sustenta que as inovações de ordens de palavras dentro do SN não têm efeitos nas ordens de palavras dentro do SV e vice-versa, restringindo-se quaisquer influências aos respectivos limites desses sintagmas.

A fim de acompanhar esses movimentos, trabalhamos com uma definição de núcleo e de complemento que procura seguir os trabalhos de Rijkhoff 1990 e de Dryer 1992, enquanto a consideração de interdependência entre os constituintes segue a linha de Waugh 1977; esta autora considera que a sintaxe é determinada por fatores concorrentes e de naturezas diversas, desde a tendência tipológica correspondente ao estágio evolutivo do sistema até a interação semântica entre os constituintes. Orientação similar é dada por Wilmet 1986, que distingue quatro abordagens principais para a especificação da posição de constituintes no SN: semântica, estilística, estética e filosófica (subjeto - modificador anteposto, ou objetivo - modificador posposto), mas reconhece o concurso de fatores lexicais, semânticos, morfológicos e de frequência na probabilidade de ocorrência de complementos antepostos ou pospostos.

Ainda em termos formais, E. Koktová 1999 oferece uma perspectiva diferente de descrição e de tratamento dos fenômenos sintáticos, além de uma tipologia alternativa de OC, baseada nas dicotomias 'ordem profunda X ordem de superfície' e 'ordem fixa X

ordem livre'; a autora sugere cinco tipos de ordem de palavras a partir da relação entre esses pares, sempre enfatizando a sintaxe como um sistema auto-regulado de maneira *sui generis*.

No caso específico do latim, vários autores (Maurer Jr. 1959; Harris 1980, dentre outros) descrevem um sistema transicional, de CV para VC, verificando-se um realinhamento generalizado dos operadores – mas, infelizmente, não é clara a seqüência de mudança desses modificadores. Observa-se, por exemplo, nesses estudos, que o latim vulgar se transforma em um sistema NG / NA / NRel / Pr, mas que, antes de chegar a essa configuração, os adjetivos denominais já eram pós-nominais; além disso, o pronome possessivo e o advérbio eram colocados *após* seus respectivos operandos. A descrição do latim vulgar e do romance como predominantemente NA encontra eco em Maurer Jr (1959). No entanto, a determinação empírica do caráter AN dominante nas línguas portuguesa (Cohen 1986/1990) e espanhola (Totaro 1998) em seus estágios iniciais leva a um questionamento lógico: ou a atribuição de uma ordem NA ao latim vulgar foi precipitada, feita com base em apenas alguns tipos de adjetivos, ou existe uma flutuação desse traço sintático no grupo românico. Cabe, portanto, perguntar como as classes de modificadores nominais e verbais se comportaram desde o latim até a conformação contemporânea das línguas EIP, e, naturalmente, se as mudanças implantadas em duas ou mais dessas línguas são coetâneas ou não e qual a sua relação com a perda flexional latina nas mesmas.

... Lo qual se vee manifiesta-
mente en muchos modos de hablar Latinos, que ajustada-
mente vienen con los Romance, i no por esso podra con-
razon dezirse, que son desta lengua. Porque si se dixesse, afir-
mariamos, que Terentio, Ciceron, Horacio, i aun el antiguo Plauto, hablaron en
Castellano, porque en ellos hallamos
estos modos de dexir mui propios del Romance. Pondre, al-
gunos exemplos breuemente. Dezimos en Romance. No lo
e viſto oi en todo el dia. Terencio **Quem ego hodie toto non_*
vidi die. Dezimos, bien grande, bien rico, bien adinerado,
bien largo, bien nacido. Tambien los vsa Ciceron. **Obuiam*
cum bene magna caterua sua uenit. el miſmo escriuiendo a At-
tico, *bene longae literae*, i en una oracion **bene numatus*. Hora-
cio **Dedecorant bene natae culpae*. Por la semejança llama-
mos braços de mar, como tambien en Italiano, i Frances, i en
Latin lo dixio Ouidio. *

. *nec brachia longo*
Margine terrarum porreberat Amphitrite
Dezimos tambien braço de rio vsado de Tito Liuiio & om-

nium primo brachio fluminis eos excludit | Plínio* dixo braços
de montes, de arboles Virgílio*
In medio ramos, annosaque brachia pandit
Vlmus opaça ingens
Aun lo que es viciofo en nueftra lengua se halla en la Lati-
na, como dezir, mas maior, Plauto **Qui dederit magis maio-
res nuqas egerit.*

Aldrete, 1606: 193-194

O emprego de palavras como instrumentos gramaticais em lugar de desinências e sufixos é, ao mesmo tempo, causa e conseqüência da ruína do sistema flexional [latino] antigo: causa, porque o aparecimento destas formas expressivas e claras antecedeu, de algum modo, à perda da flexão; conseqüência, porque, à medida que a flexão se enfraquecia, o seu uso se tornava mais necessário para a enunciação do pensamento.

Maurer Jr. 1959: 191

Todas las lenguas romances son más conservadoras que el francés... El italiano es, después del sardo, la más conservadora de las lenguas románicas y, en general, es la que menos se ha alejado del latín. ... De las lenguas iberrománicas, el portugués y el español son más conservadores que el catalán, y el portugués es más conservador que el español.

Vidos 1968: 361

A escolha de línguas situadas em limites opostos da România Ocidental, somando-se a todos os fatores externos que contribuíram para a formação das mesmas (substratos diversos, época e intensidade diferenciais de romanização, invasões subseqüentes, continuidade dialetal etc.), refletiu a expectativa de se encontrarem, entre esses sistemas, semelhanças e distinções no que respeita à mudança sintática. Em outras palavras, ainda que se considerasse a possibilidade de identificar movimentos similares em alguns aspectos da mudança, esperava-se que a distância entre o italiano e o português resultasse em tendências mais díspares do que as observadas entre os sistemas ibéricos.

O modelo de difusão sintática (Naro & Lemle 1976) que utilizamos, por sua vez, contempla justamente a possibilidade de uma determinada mudança sintática espalhar-se (ou não) através do sistema lingüístico – por exemplo, de um item lexical para outro ou entre ambientes sintáticos. Dessa forma, privilegiamos a caracterização e a comparação dos recortes sincrônicos representados pelos textos históricos, procurando observar de maneira bastante refinada o alcance interno das modificações estruturais enquanto as mesmas se processam – em vez de procurar estabelecer *a priori* uma justificativa externa para a mudança de sistemas lingüísticos derivados de uma única fonte (Sapir 1920), 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 lingüístico (cf. Thomason & Kaufman 1991: 9ss), estendem-se, *mutatis mutandis*, à morfologia e à sintaxe os princípios da difusão lexical, conforme, aliás, já referido acima (Wallace 1981; Davies, 2000); 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.

De maneira mais abrangente, a delimitação teórica que suporta este trabalho é oferecida por Harris & Campbell 1995 (assinalada como HC nesta Seção) e refinada por Harris 2005, que correlaciona diversos aspectos da mudança de ordem de constituintes de maneira a permitir falar em sintaxe diacrônica.

HC definem, inicialmente, o arcabouço de uma teoria de mudança sintática através da caracterização de universais da sintaxe diacrônica que dependem crucialmente de um *corpus* interlingüístico consideravelmente extenso. Essa proposta, retomada em Harris 2005, é dirigida tanto à caracterização precisa da mudança sintática e dos universais que a delimitam quanto à expressão de uma teoria sintática que possa analisar a mudança de ordem de constituintes. Os elementos que comporiam essa teoria de mudança sintática seriam (HC 1995:50) resumidamente apresentados como segue, e implicariam a existência de a) apenas três mecanismos gerais de mudança, que compreendem a reanálise, a extensão e o empréstimo; b) um conjunto definido de operações diacrônicas gerais, as quais são implementadas por meio de um desses mecanismos; c) um conjunto de princípios gerais que interagem com essas operações, e d) um conjunto de construções sintáticas que são parte de uma gramática universal, no sentido de que elas sempre podem ser referidas em estruturas alternativas.

O primeiro desses pontos sugere, a partir da comparação inter-lingüística, três estratégias universais de mudança. A reanálise é definida como “um mecanismo que altera a estrutura subjacente de um padrão sintático e que não envolve nenhuma modificação de sua manifestação de superfície” (Harris *op. cit.*: 532). Salvi 2000 (sobre a retenção de formas latinas na morfossintaxe românica), Gutiérrez 1997 (sobre o futuro perifrástico em

quechua) e Oliveira 2004 (sobre a voz passiva em português) apresentam exemplos de reanálise que envolvem diretamente o italiano, o espanhol e o português, respectivamente.

Em seguida, temos a extensão, que nos interessa mais particularmente, e que seria “um mecanismo que resulta em mudanças na manifestação superficial de um padrão e que não envolvem alterações imediatas ou intrínsecas da estrutura subjacente a ele associada” (Harris *op. cit.: idem*). Poder-se-ia relacionar esse mecanismo de generalização de regras, por exemplo, à mudança dos recursos de marcação de funções sintáticas observada entre a gramática do latim vulgar e as das línguas românicas. Um dos aspectos distintivos desses mecanismos é o fato de que a reanálise introduz, freqüentemente, complexidade sintática na gramática, ao passo que a extensão pode eliminar exceções e irregularidades ao trazer nova perspectiva de análise diante do restante da gramática existente (Harris *op. cit.: 97*).

O terceiro desses processos é o empréstimo lingüístico, que é “um mecanismo de mudança pelo qual a réplica de um padrão sintático é incorporada à língua beneficiária através da influência de um padrão originalmente encontrado na língua de contato”. Tomem-se, para efeito de fornecer a descrição de casos mais específicos, os trabalhos de Dillon 1945 e de Haugen 1950 sobre as evidências históricas do empréstimo; de De Marco 2005 e de Bencini & Manetti 2005, sobre a evolução do léxico italiano; de Patterson 1986 e de Silva-Corvalán 2001, sobre os aspectos envolvidos nos empréstimos feitos pelo espanhol; e de Harmon 1994, de Lamberti 2004 e de tantos outros sobre os empréstimos no léxico português.

As operações a que se referem HC 1995:57ss correspondem a um conjunto de mudanças potenciais, que não se aplicam necessariamente em todas as línguas e nem mesmo são obrigatórias quando as condições para sua aplicação se verificam. Em vez disso, elas constituem uma série de mudanças recorrentes em padrões oracionais encontrados em línguas não-relacionadas geneticamente, mas que estão universalmente disponíveis e que são derivadas a) dos mecanismos de mudança sintática, b) da disponibilidade de determinadas construções e c) dos princípios gerais que regem a mudança lingüística. HC *op. cit.: 58* citam o movimento de pronomes relativos e a alteração do *status* verbal (produzindo-se auxiliares a partir de verbos mais ‘altos’) como representativos de operações diacrônicas universalmente viáveis.

Os princípios universais propostos por HC *op. cit.*: 57, 113ss são restrições impostas aos mecanismos de mudança expostos acima; no caso específico da extensão, o princípio determina que esse mecanismo sempre remova condições da regra sobre a qual atua; não se exclui, contudo, a possibilidade de intervenção concomitante de restrições específicas sobre uma operação qualquer. A extensão de uma regra R está limitada à remoção de uma condição de R; em outras palavras, a extensão não implica alterações da regra.

Outro aspecto considerado essencial para que se possa oferecer uma compreensão mais abrangente das comparações inter-lingüísticas é o fato de que a mudança sintática não é aleatória, mas controlada por regras (Harris *op. cit.*: *ibidem*). Além disso, a motivação para a mudança num nível – e.g., morfológico ou sintático – estaria associada à operação de processos em outro desses níveis. Nesses termos, a reanálise implica uma motivação semântica e pode levar a fenômenos de extensão (Harris *op. cit.*: 536) como a difusão sintática.

Segundo a proposta de HC, portanto (215), a mudança sintática pode surgir graças a uma tendência a um alinhamento de traços; essa harmonia seria iniciada através da extensão e mantida por força do *status quo*. Por outro lado, a harmonia de ordenamento dos constituintes pode ser alterada pela reanálise de uma construção, que introduz uma ordem diádica que não é harmônica com outras de mesma natureza ou de tipos distintos (232). Isso leva à observação de que a ordem de palavras não resulta unicamente da operação de regras estruturais da oração, além do que cada ordem individual é dependente das demais em medidas que ainda cumpre determinar. De uma forma ou de outra, parece claro que a maioria das mudanças têm mais de uma causa. Um dos aspectos mais interessantes da causação múltipla é que as causas internas e externas freqüentemente interagem: por exemplo, um traço que parece ser induzido por contato pode também ter uma motivação interna; em outros casos, o desenvolvimento de um único caráter harmônico pode envolver tanto a reanálise quanto a extensão (215).

Ainda a respeito da harmonia, HC (215, 315 *et passim*) observam que as noções de ‘*targets*’, ‘conspiração’ ou ‘*deriva*’, anteriormente tratadas como possíveis causas da mudança em sintaxe e em outras partes da gramática, devem ser consideradas, mais corretamente, como caracterizações da maneira pela qual a mudança sintática freqüentemente procede. Para Lightfoot 1991 e HC *op. cit.*:231, é inegável a afirmação

empírica de que diversas línguas toleram sistemas não-harmônicos por longos intervalos de tempo; a flutuação das proporções das ordens minoritárias na determinação do caráter dominante da língua é abordada por Kaufmann 1993, Whaley 1997, Ruhlen 2000 e Briscoe 2000, que tomam não apenas a frequência das variantes, mas também os contextos comunicativos na caracterização tipológica. Apesar de a harmonia diádica não ser, portanto, um universal lingüístico mas uma tendência, ela é preferida por demandar uma única regra de colocação das palavras, ao passo que os sistemas não-harmônicos exigem várias; essa perspectiva, diga-se de passagem, faz eco à proposta de correlação das ordens de palavras à facilidade de processamento, encontrada em Kuno 1974 e em Hawkins 1990, ambos *apud* HC 238.

Na proposta teórica que adotei, é imprescindível considerar, por fim, que os estados transitórios, em que cada configuração restringe as possibilidades de variação e de mudança e lança as bases para os estados que lhe são posteriores, mostram que o sistema lingüístico se comporta de maneira complexa, determinada pela interação variável de diversos fatores. Nesse sentido, merecem nota as contribuições a) de Lightfoot 1991, que observa a patente estabilidade relativa de estados não-harmônicos por longos intervalos da história de diversas línguas, e b) de Whaley 1997:96ss, que mostra que a determinação da ordem dominante é uma função complexa não apenas de variáveis como a frequência, mas também (e até) do tipo textual e da porção do texto selecionada para análise. Em consonância com esse ponto-de-vista, a aferição das tendências de dominância e de harmonia sincrônicas e de estabilidade diacrônica podem ser representadas qualitativamente segundo modelos como os de Gleick 1984, retomados, para os sistemas lingüísticos, por Lightfoot 1991 e Briscoe 2000.

V - Metodologia

Foram selecionados textos em prosa escritos originalmente nas línguas em estudo (entendendo-se que os textos em questão sejam representativos da época em que foram escritos, como já colocado na seção II), 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. Esse *corpus* é tratado através de métodos e recursos computacionais, na medida do possível adequados à identificação dos contextos em que os dados ocorrem e de suas respectivas posições nos sintagmas - segundo os procedimentos utilizados em Totaro 1995, 1998, em que tais recursos foram utilizados para se buscar a agilização da construção da base de dados.

Ainda no que tange à construção das bases de dados, classificaram-se as ocorrências segundo critérios morfológicos, sintáticos e semânticos 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 lingüísticos amostrados. Os critérios morfológicos – 'natureza do sujeito', 'natureza do objeto', 'status do pronome' e 'tempo do verbo' – foram sugeridos a partir de trabalhos que contemplam a influência de traços morfológicos e do peso do sintagma na mudança (Wilmet 1986; Martins, 2002); os semânticos, a partir de Waugh 1977, Lehrer 1974 e de Beaugrande 1997. As hipóteses que orientam a identificação e a seleção dos fatores pertinentes são fruto da observação das ocorrências encontradas nos textos.

A identificação de combinações promissoras baseadas no critério *ambiente oracional* levou à determinação das proporções em que se distribuem os constituintes S-V-C para cada época nos textos tratados, verificando-se a estabilidade posicional ou a mudança de ordem através da quantificação da frequência das ocorrências por século. As tendências de comportamento desses sistemas lingüísticos ao longo do tempo foram lançadas em gráficos que sumarizam os resultados numéricos do tratamento estatístico das ocorrências codificadas. Vejamos, nas Seções seguintes, uma descrição detalhada desses procedimentos.

V.1 – Tratamento dos textos

As ocorrências textuais que compõem as bases de dados foram selecionadas e catalogadas a partir de arquivos de formato .DOC construídos a) depois de se passarem os textos eletrônicos por OCR (como se deu com os textos italianos e portugueses – q.v. Figs. PN.1 e PN.2 abaixo) ou b) depois de se passarem os textos por *scanner* e OCR (como foi o caso dos textos espanhóis e dos textos italianos e portugueses do séc. XX – q.v. Fig. PN.3 abaixo). No segundo caso, os textos foram digitalizados em resolução de 600dpi e gravados no formato .JPG, após o que foram submetidos às rotinas de reconhecimento de caracteres do OmniPage Pro; para os demais textos, utilizaram-se diretamente as imagens que constituem os arquivos .PDF das bibliotecas eletrônicas consultadas, e que vêm em resoluções de até 300dpi.

O OmniPage Pro permite a edição dos textos gerados e a atualização de um arquivo de parâmetros ajustáveis (*training file*) de reconhecimento de traços específicos, o qual pode ser utilizado em etapas posteriores que envolvam caracteres idênticos ou de traçado semelhante. Esse procedimento é particularmente interessante no tratamento de grandes massas de dados, pois diminui consideravelmente o trabalho de construção dos textos ou das planilhas que contêm as ocorrências: o programa, de fato, pode ser instruído para passar a reconhecer um *ð* ou *đ* como **d**, e não como **6**, **ò** nem **ô**; **ſ** ou **ſ** como **s**, e não como **l**, **j** nem **f**; capitulares ou maiúsculas floreadas, como **R** ou **P**, são reconhecidas como maiúsculas simples. Dependendo do caráter do texto a ser produzido e do grau de intervenção permitido, podem-se desdobrar as abreviaturas e transformar alguns elementos originais, ou, naturalmente, mantê-los como são: **&**, **ꝥ**, **C** e **p**, por exemplo, podem ser identificados automaticamente e transcritos como se desejar.

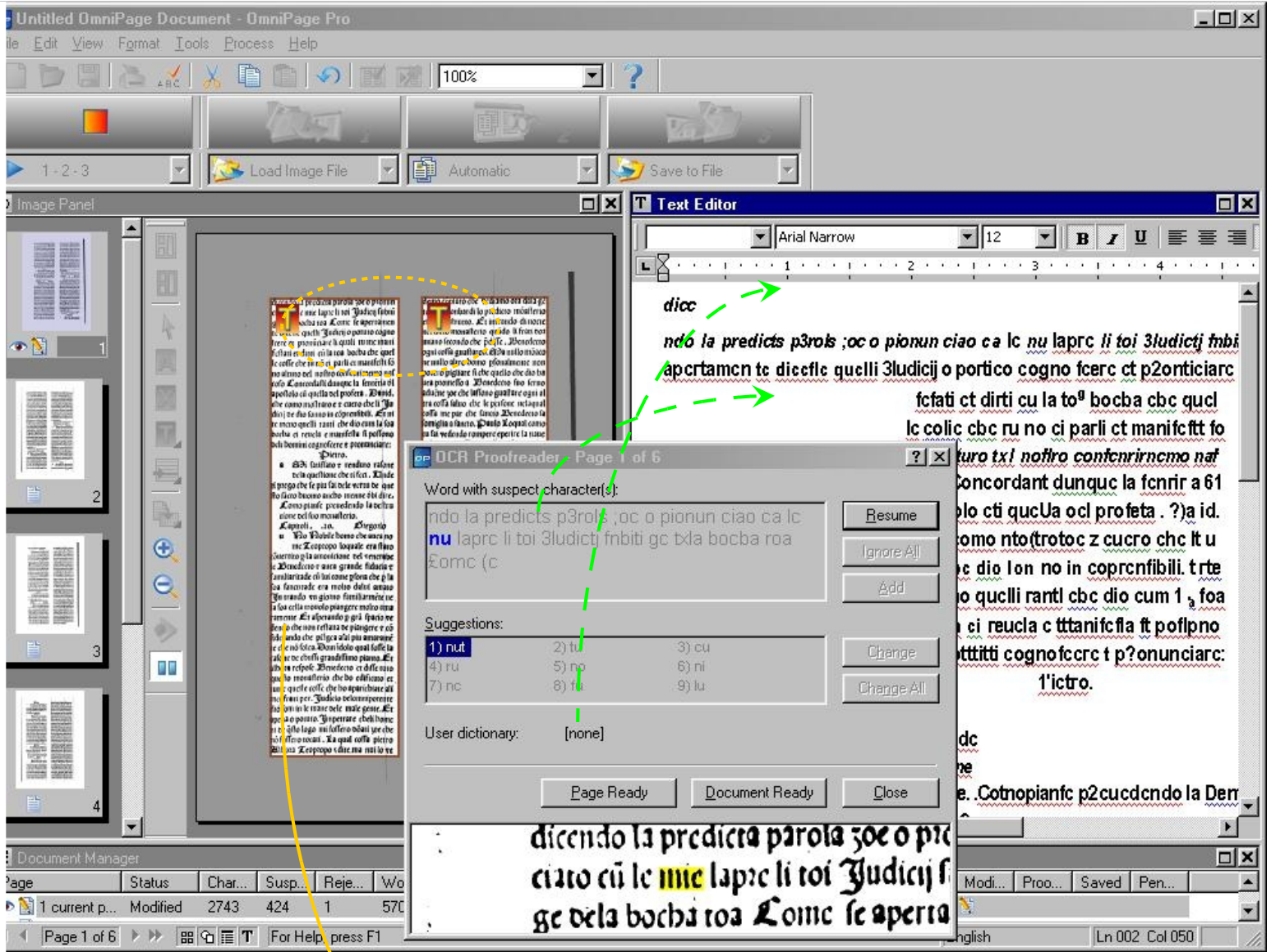


Fig. PN.1. Exemplo de um dos passos do tratamento dos textos a partir de um arquivo PDF (*Dialogo di Sancto Gregorio – séc. XV*). O revisor permite a gravação de um arquivo de informações que serão utilizadas pelo programa em sessões posteriores, aumentando a precisão do processo e diminuindo a necessidade de intervenções por parte do usuário. Observe-se ainda que, definidos os parâmetros da imagem na biblioteca de origem (col. Gallica), não é possível melhorar a qualidade da aquisição (*scanning*).

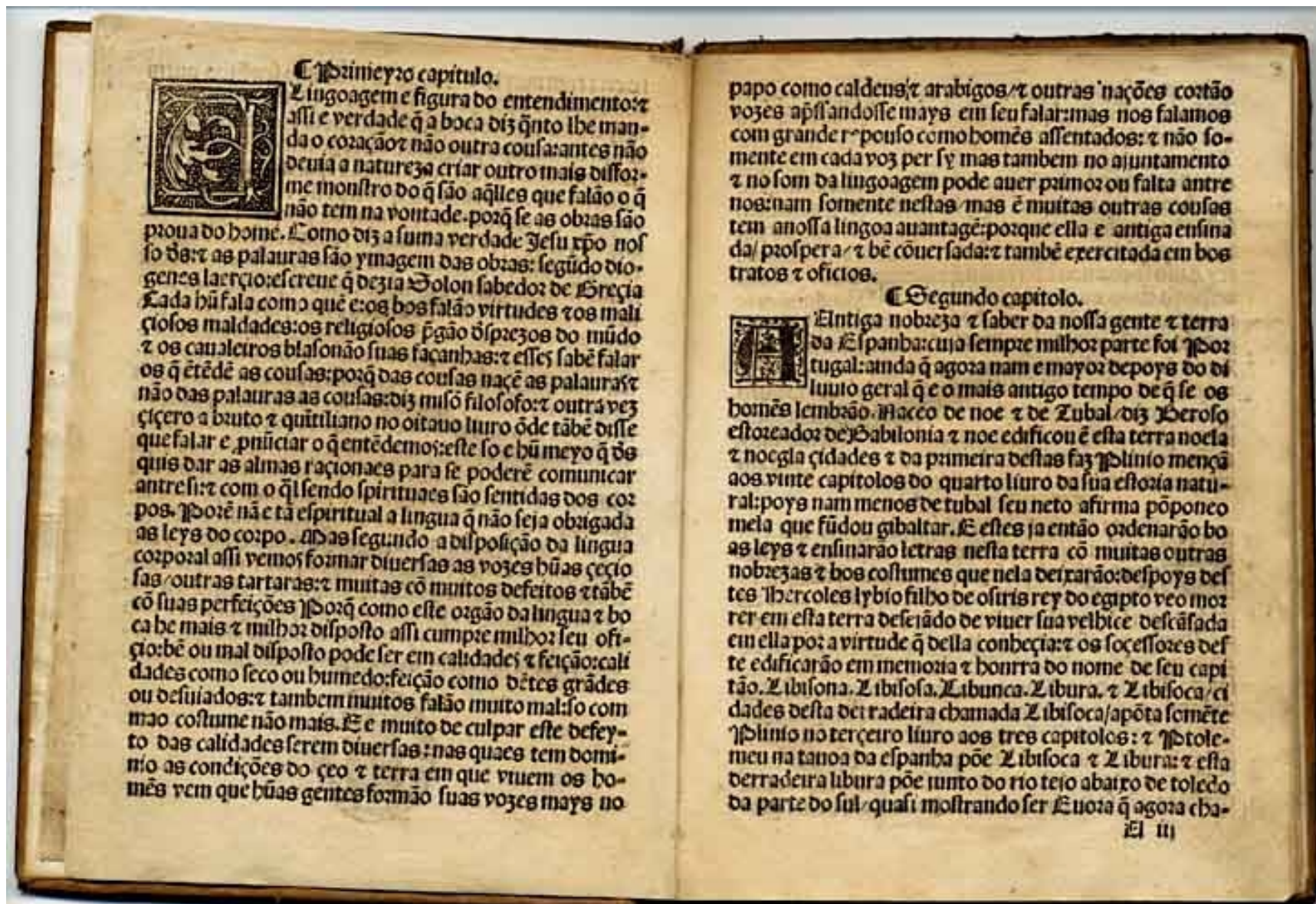


Fig. PN.2 – Duas das páginas da 'Grammatica' de Fernão de Oliveira (séc. XVI) utilizadas na amostragem.

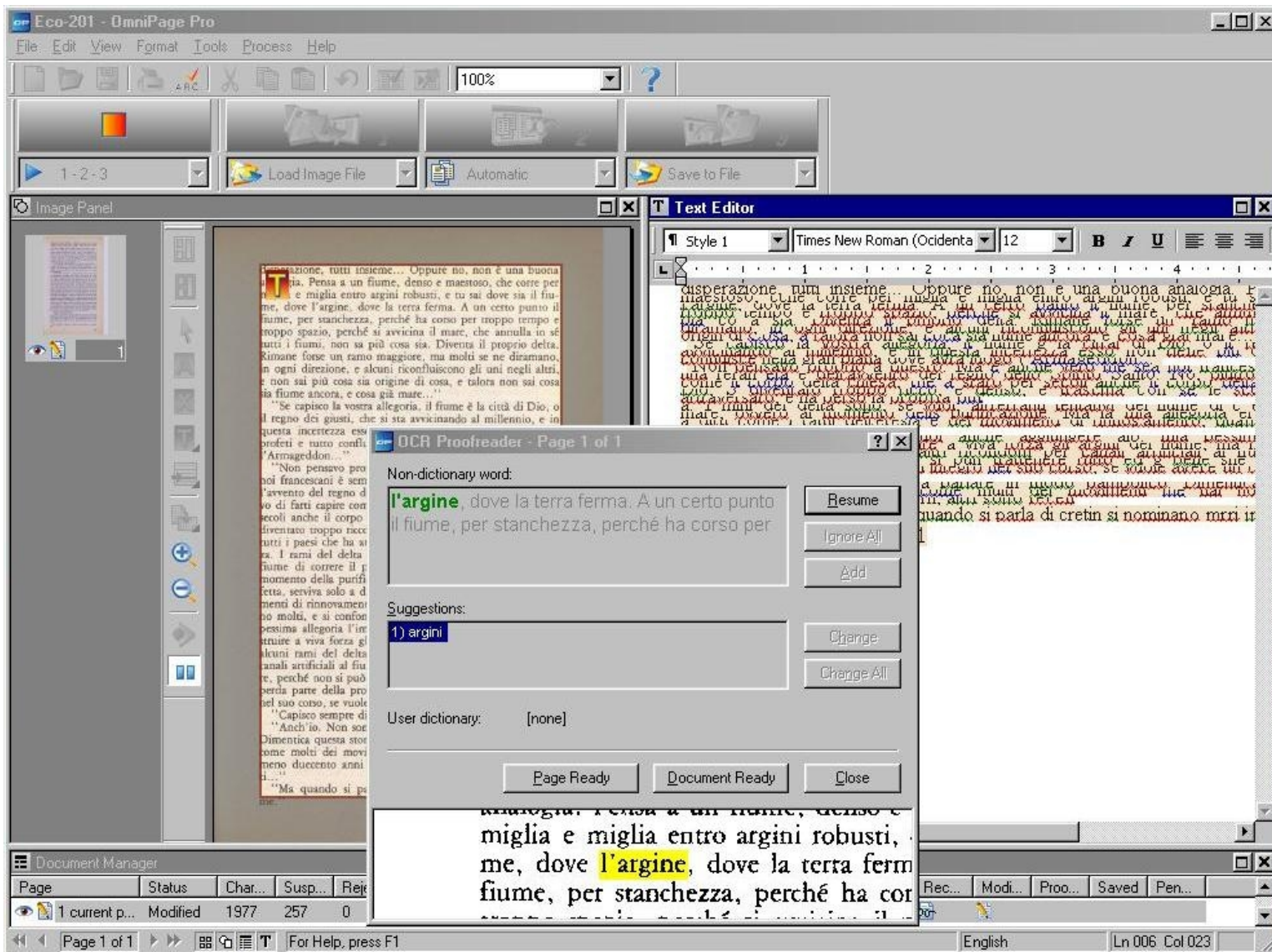


Fig. PN.3. Exemplo de um dos passos do tratamento dos textos a partir de um arquivo JPG (ref. pág. 201 de *Il nome della rosa* – séc. XX). Neste caso, a imagem foi gerada através de *scanner*, o que permitiu alterar-lhe as características a fim de se conseguir melhor desempenho do OCR. No caso dos arquivos-texto produzidos a partir de imagens de média resolução (~600dpi), quase não foi necessário introduzirem-se alterações – tanto em virtude do fato de se aprimorarem continuamente os arquivos de trabalho do OCR quanto graças ao ajuste dos parâmetros de luminosância e de contraste da imagem.

V.2 –Montagem das bases de dados

Constituídos os textos eletrônicos de suporte, as ocorrências e seus respectivos contextos foram copiados e transpostos (*copy / paste*) para planilhas do MicroSoft Excel nomeadas a partir do texto-fonte (Fig. PN.4). Os principais delimitadores utilizados para a definição do comprimento do contexto foram a) a inclusão dos elementos S, V e C expressos, b) a largura da coluna e c) qualquer elemento de pontuação. Estruturas intercaladas separando os constituintes S/V/C, como aposto ou orações subordinadas, foram substituídas por reticências e, sempre que necessário, passaram a constituir entrada separada na tabela⁶. Em seguida, utilizaram-se os critérios expostos adiante, nesta mesma Seção, para marcar as ocorrências segundo as características com que se apresentam; o resultado é uma tabela semelhante à da Fig. PN.5b, onde de encontram a ocorrência em seu contexto e os traços atribuídos à mesma. Esse procedimento foi utilizado em Totaro 1998 e em Lima 2003, e acompanha muitas das sugestões feitas por Leech *et al.* 1996 e Biber, Conrad & Reppen 1998 sobre a marcação sintática de *corpora*.

A coleta de ocorrências foi feita em textos espanhóis dos séculos XIII a XX (os mesmos textos utilizados para a constituição do *corpus* em Totaro 1998), portugueses dos séculos XV a XX (dois da Biblioteca da FaLe / os demais, a partir dos textos disponíveis na Biblioteca Nacional Digital - www.bnd.pt) e italianos dos séculos XV a XX (todos obtidos na coleção eletrônica Gallica da BNF – gallica.bnf.fr); consultaram-se as gramáticas históricas disponíveis em formato digital e na FaLe (v. referências bibliográficas abaixo) a fim de que se identificassem as estruturas sintáticas com a maior pertinência possível, 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.

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. Em

⁶ Esse procedimento é semelhante ao modelo híbrido *constituência-dependência* discutido por Teich 1998 para a identificação de sintagmas.

ambos os casos e especialmente com respeito às obras em que há grande número de abreviaturas, procedeu-se ao desdobramento manual das mesmas (cf. preceitos de Capelli 2001). Dessa maneira, a base de dados foi composta em formato .XLS, destacando-se as ocorrências oracionais do texto eletrônico e classificando-as segundo seus caracteres de estruturas subordinadas e/ou coordenadas. Para fazer essa classificação, utilizaram-se os princípios de tendência funcional para a avaliação da relação entre orações, como exposto em Pidal (1949) para a sintaxe histórica do espanhol, tendo-se estendido o mesmo procedimento ao português e ao italiano.

201

disperazione, tutti insieme... Oppure no, non è una buona analogia. Pensa a un fiume, denso e maestoso, che corre per miglia e miglia entro argini robusti, e tu sai dove sia il fiume, dove l'argine, dove la terra ferma. A un certo punto il fiume, per stanchezza, perché ha corso per troppo tempo e troppo spazio, perché si avvicina il mare, che annulla in se tutti i fiumi, non sa più cosa sia. Diventa il proprio delta. Rimane forse un ramo maggiore, ma molti se ne diramano, in ogni direzione, e alcuni riconfluiscono gli uni negli altri, e non sai più cosa sia origine di cosa, e talora non sai cosa sia fiume ancora, e cosa già mare..."

"Se capisco la vostra allegoria, il fiume è la città di Dio, o il regno dei giusti, che si sta avvicinando al millennio, e in questa incertezza esso non tiene più, nascono falsi e veri profeti e tutto confluisce nella gran piana dove avrà luogo l'Armageddon..."

"Non pensavo proprio a questo. Ma è anche vero che tra noi francescani è sempre viva l'idea di una terza età e dell'avvento del regno dello Spirito Santo. No, piuttosto cercavo di farti capire come il corpo della chiesa, che è stato per secoli anche il corpo della società tutta, il popolo di Dio, è diventato troppo ricco, e denso, e trascina con sé le scorie di tutti i paesi che ha attraversato, e ha perso la propria purezza. I rami del delta sono, se vuoi, altrettanti tentativi del fiume di correre il più presto possibile al mare, ovvero al momento della purificazione.

-
-
-

AN / NA

Oppure no, non è una buona analogia. Pensa a un fiume, denso
che corre per miglia e miglia entro argini robusti, e tu sai
sai dove sia il fiume, dove l'argine, dove la terra ferma. A un certo
il proprio delta. Rimane forse un ramo maggiore, ma molti se ne

-
-
-

S-V-C

e alcuni riconfluiscono gli uni negli altri, e non sai più cosa sia origine di cosa, e talora
e alcuni riconfluiscono gli uni negli altri, e non sai più cosa sia origine di cosa, e talora
e alcuni riconfluiscono gli uni negli altri, e non sai più cosa sia origine di cosa, e talora
e talora non sai cosa sia fiume ancora, e cosa già mare..."
"Se capisco la vostra allegoria, il fiume è la città di Dio, o il regno dei giusti, che
"Se capisco la vostra allegoria, il fiume è la città di Dio, o il regno dei giusti, che

-
-
-

Fig. PN.4 – Ilustração do procedimento de constituição das bases de dados .XLS de AN (italiano) e de S-V-C (para os três sistemas) a partir dos arquivos .DOC gerados no passo anterior pelo OCR.

V.3 - Métodos estatísticos

O número mínimo de ocorrências a ser colhido em cada texto foi determinado a partir da metodologia exposta em Woods, Fletcher & Hughes 1986 e em Těšitelová 1992, e leva a gráficos semelhantes a P5a; essa figura mostra ter-se atingido, próximo das 150 ocorrências, um patamar de estabilidade nas proporções relativas entre SVC (a ordem preferencial⁷) e as demais, considerando-se textos dos séculos XIII e XX – os extremos do intervalo de amostragem:

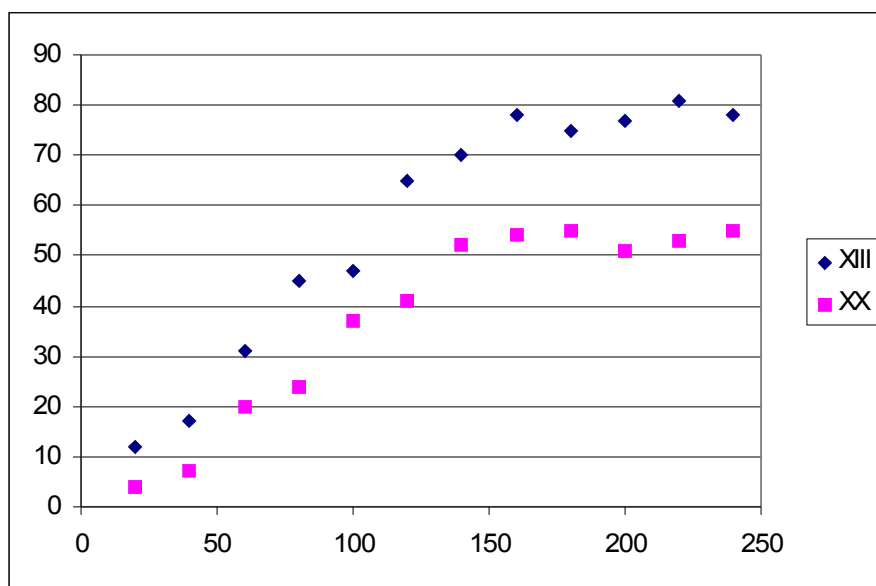


Fig. PN.5a – Comportamento da base de dados em relação a um único traço sintático (número X porcentagem de ocorrências).

A metodologia empregada na confecção do gráfico acima diminui as diferenças individuais existentes entre os textos de um mesmo século, fazendo com que a cada século corresponda a *média* dos valores de cada texto; por outro lado, uma curva ajustada (cf. Vieira 1975) aos pontos deixa claro que, com base nos dados coletados até o momento e para o estilo de texto utilizado na amostragem e constituição do *corpus*, há uma mudança na ordem de colocação de constituintes. O procedimento mostra ainda que esse decréscimo

⁷ Testes feitos para o par de constituintes A e N sugerem uma tendência à horizontalização da curva entre 100 e 125 ocorrências/texto/século (Oc/TS); para os constituintes S-V-C, o patamar da curva é atingido em torno de 150 Oc/TS. Na coleta de ocorrências, buscou-se observar esse número como limite inferior, a fim de que se pudesse trabalhar com uma margem de segurança.

ultrapassa os limites da simples variação individual ou dialetal presente nos textos, refletindo, outrossim, 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 parâmetro (Lightfoot 1991; Shen 1991; Niyogi & Berwick 1994 1995; Goertzel 1995a 1995b; Kirby 1997) decorrente da competição gramatical entre variantes (Kroch 1994). O mesmo perfil é apresentado por Batali (1998), em simulações feitas com redes neurais, corroborando a viabilidade e a propriedade do tratamento numérico como elemento de análise das ocorrências reais.

Os motivos que levaram à construção das bases de dados no MS Excel foram 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 seqüência 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 de ocorrências, o MS Excel oferece duas alternativas principais, cada qual com suas vantagens e desvantagens – uma totalmente automática, que envolve uma fórmula simples de contagem num intervalo definido, e que é sensível a quaisquer retificações ou alterações na base original; e uma absolutamente manual, que se presta à busca de valores / atributos de forma mais livre (por exemplo, se se deseja apenas verificar, de maneira prospectiva, o comportamento de um conjunto de dados). A primeira exige a definição de um intervalo de células dentro do qual os valores (numéricos ou não) devem ser buscados e contados, e retorna unicamente o valor numérico igual ao total de *tokens* encontrados. Dessa forma, as células das tabelas de tratamento numérico são interligadas, e uma alteração numa delas resulta no ajuste das demais, até que toda a cadeia seja recalculada e, em última instância, todos os dados sejam lançados em gráfico (cf. Figs. PN.6a e PN.6c). É necessário, obviamente, ater-se ao formalismo imposto pelo programa e cuidar para que todas as etapas estejam corretamente definidas em termos de parâmetros e

de referências; contudo, o tempo dispendido nessa fase é compensado pela confiança que se ganha no tratamento dos dados e nos resultados. Em outras palavras, a obtenção de um perfil como o da Fig. PN.5b (idêntica à Fig. II, para as proporções de AN em textos italianos), em que há um máximo local no século XVIII que 'quebra' a curva decrescente que vinha sendo construída desde o século XV, não é analisada como decorrência de erro de digitação, mas como o resultado de um comportamento excêntrico de um dos textos desse século em termos da relação AN/NA.

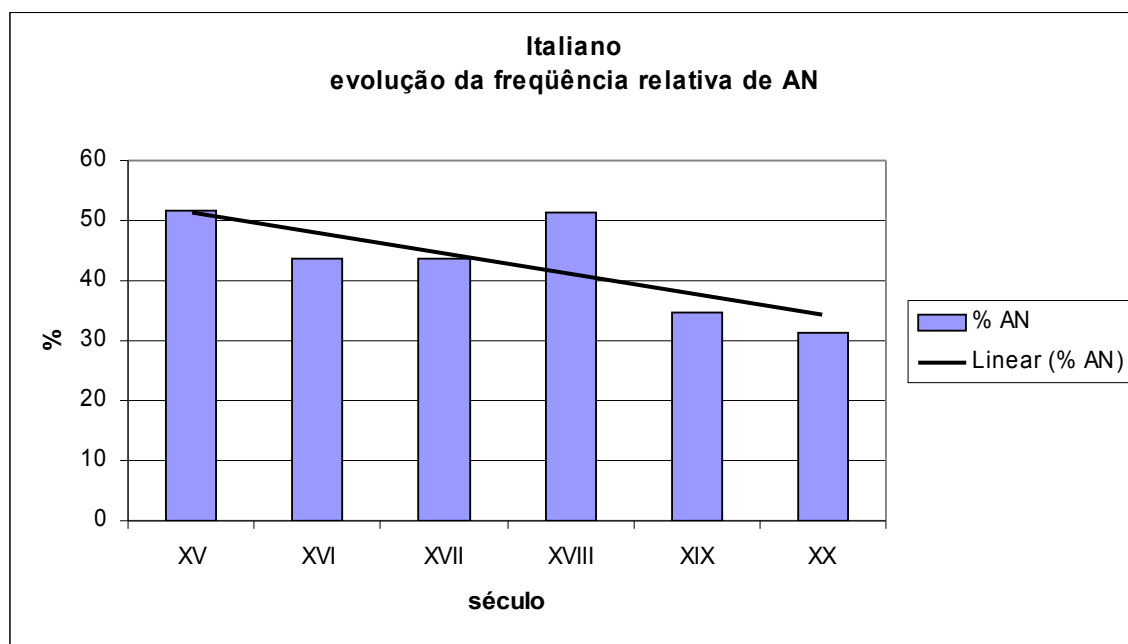


Fig. PN.5b – demonstração do resultado gráfico do tratamento automático de ocorrências: a dispersão de pontos diante de retas de regressão linear ou de curvas que indicam as tendências do conjunto é atribuída às peculiaridades de alguns textos.

A segunda alternativa implica a associação de um filtro (*Data>Filter>AutoFilter*) a um intervalo de células qualquer; esse filtro é acionado manualmente através da seleção de um dos valores compreendidos no intervalo de amostragem, e retorna o número de *tokens* que detêm a característica selecionada **além** de isolar as linhas que contém esses traços. Os filtros podem, por fim, ser concatenados, combinando-se os critérios de classificação das ocorrências (Fig. PN.6b).

fórmula que faz a contagem do número de ocorrências de AN na pasta referente ao texto *Sermoni del divo Bernardo*, do séc. XV, e retorna esse valor para a célula C4 da pasta corrente ('estatística geral').

Os valores de C4 (AN) e de C8 (NA) do mesmo texto são contados automaticamente. A soma desses valores dá o total de ocorrências colhidas no mesmo, que é utilizado no cálculo da porcentagem de AN encontrada no texto; esta, por sua vez, é usada para calcular a média dos valores de AN para cada século, a qual é imediatamente lançada no gráfico.

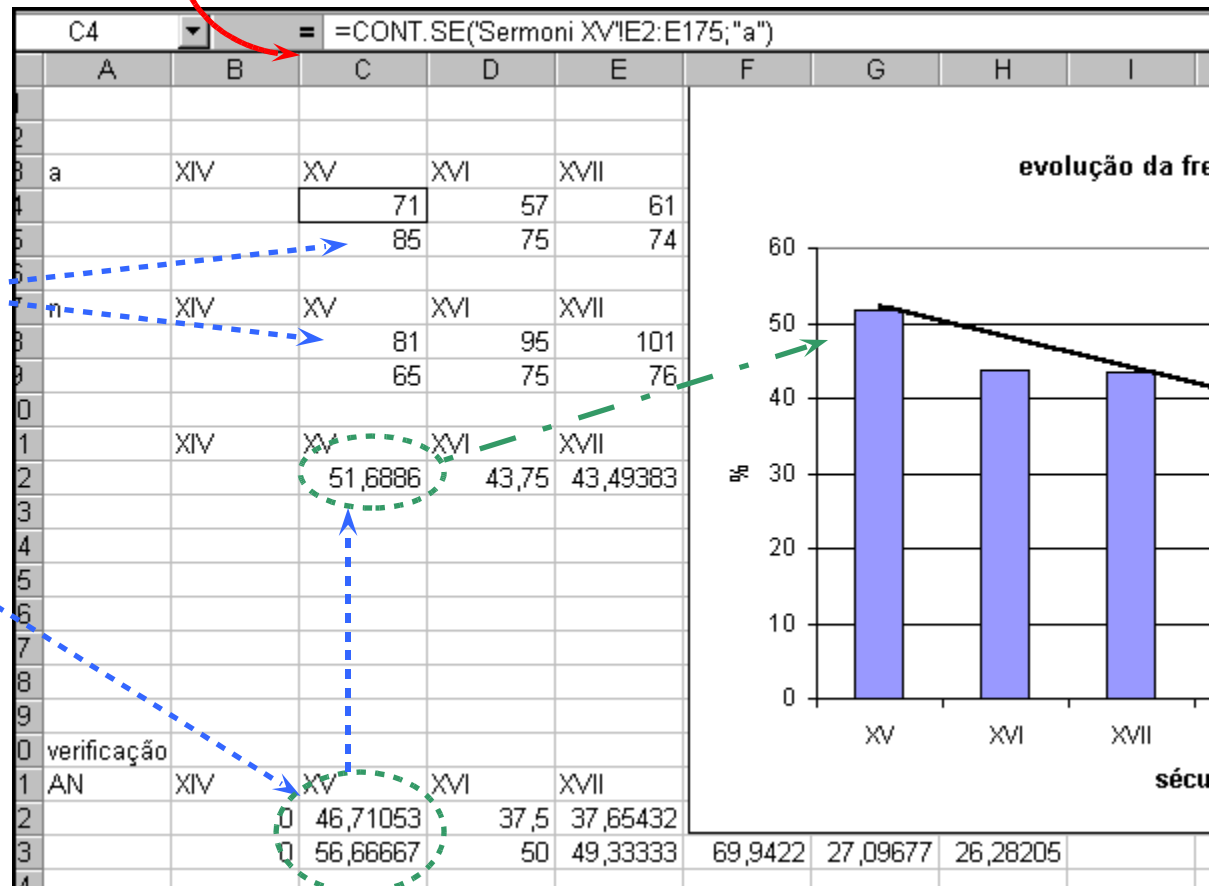


Fig. PN.6a – Descrição sumária do funcionamento do primeiro método de contagem e de cálculo de valores utilizados na confecção dos gráficos com o MicroSoft Excel. As fórmulas relacionam as células de várias planilhas, interligando-as e diminuindo a probabilidade de erro graças à automatização dos procedimentos de cálculo e de plotagem.

parâmetros de classificação

ocorrências

O35 = SEM														
	G	H	I	J	K	L	M	N	O	P	Q	R	S	U
1	F	N	N	S	A	R	A	T	CS	I	I	verbo	col	
26	i		n		3	w		α	DISP	h	n	fare	1	dice(n)do io priego idio che tifaccia ogni bene / et che tifaccia saluo
29	i		n		3	u		α	EMOT	h	h	amare	1	Et pero langelo che amaua la uergine Maria come contento dogni sua salute
30	i		n		3	u		α	COGN	h	n	avere	2	e / desideroso che lhauesse maggior gratia dadio che epsa non haueua prima
35	i		n		3	u		α	SEM	h	n	dire	3	Come se dicesse io sono contento che tu sia in gratia didio
41	i		p		3	u		α	SEM	h	n	dire	3	diciamo Aue cioe tu sia salua non ch(e) uogliamo dire che dio le dia salute che
49	i		p		3	u		α	SEM	h	n	dire	3	quasi sedicesimo Maria noi semo conte(n)ti della tua eterna gloria
53	i		p		3	u	z	α	COGN			sapere	2	Cosi come sappiamo che ladurera sâza dubio:et uogliamo ogni tuo bene
74	i		n		3	u		α	SEM	h	n	dire	3	che dice sancto Hieronymo onde lachiesa humilm(en)te confessa che
76	i		n		3	u		α	ATTR			avere	2	confessa che ha bisogno della sua sancta mano quando ladice Aue maria
78	i		p		3	u		α	SEM	h	n	dire	3	quasi dicessi tu sia sempre salua madonna mia et illuminatrice
85	i		n		3	x		α	ATTR			essere	2	et e Madre delRe del cielo et della terra lesu christo
87	i		n		3	x		α	IDENT			essere	2	ilquale e una substantia col padre:et e / tabernaculo dello spirito sancto
92	i		n		3	u		α	SEM	h	n	significare	1	ilquale significa quella Candida Vergine del cui purissimo sangue
99	i		n		3	u		α	DISP	h	h	illuminare	1	et purificata da quella luce che illumina tucti li huomini et ledonne
104	i		n		3	u		α	PERC	h	n	sentire	3	alla quale bisogna dirizare liochi quando sentiamo lafortuna / perche epsa
107	i		p		3	u		α	SEM	h	n	significare	1	Et pero questo nome e dolce / che significa quello che cidona
108	i		n		3	w		α	DISP	h	n	dare	1	quello che cidona mille dolce cōsolationi cioe / lastella del mare che
110	i		n		3	u		α	DISP	h	h	coniungere	2	laquale strectissimamente coniuuge la creatura rationale alsuo dulcissimo

linhas que correspondem à combinação dos fatores VC / 3 / α para as ocorrências do texto *Expositione del Reverendissimo*(séc. XV).

Fig. PN.6b – Descrição sumária do funcionamento do segundo método de contagem e de cálculo de valores utilizados na confecção dos gráficos com o MS Excel. A seleção dos parâmetros definidos em colunas (assinala-se aqui, em azul, o cruzamento dos traços [ordem: VC], [ambiente oracional: subordinadas] e [tempo do verbo: finito]) mostra as ocorrências a que esses traços foram atribuídos em conjunto. O próprio programa conta essas ocorrências e informa, no mesmo campo, o total de ocorrências do texto.

V.4 - Outros cálculos

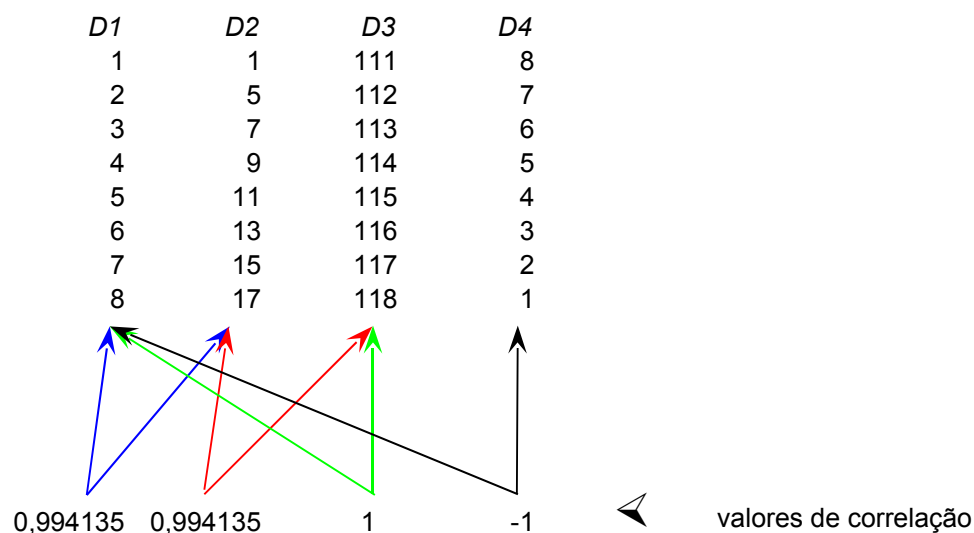
As funções mais extensamente usadas nas planilhas de tratamento estatístico foram as de cálculo de iRRL, de R^2 e de desvio-padrão, $dp = \sqrt{\text{var}}$. Sempre que necessário determinar uma característica de comportamento das ocorrências para um traço qualquer, os blocos de dados numéricos das pastas de tratamento estatístico são seguidos, à sua direita imediata, de um conjunto de células que participa do cálculo da inclinação das retas de regressão linear (iRRL) – valor definido por fórmula associada à célula colocada abaixo do bloco de dados original (Fig. PN.5c), e que define a inclinação da melhor reta para o subconjunto de dados. Os valores de iRRL foram usados para avaliar a tendência da amostra, e faz-se referência constante a eles no texto como instrumento de distinção entre os comportamentos dos *corpora*: uma iRRL positiva indica que a frequência de ocorrência do traço em questão aumenta com o tempo naquele sistema, e vice-versa.

Os cálculos do desvio-padrão, da distorção e de R^2 se prestam, no presente estudo, à definição de valores de referência para um critério de tratamento das ocorrências. O desvio-padrão mede a dispersão dos dados em relação à sua média: quanto maior o valor de dp , maior a variação dos dados individuais e, portanto, maior a assimetria da amostra – trata-se, portanto, de uma ferramenta de análise restrita a cada conjunto de dados, como o cálculo de distorção da amostra (que tem a vantagem de mostrar se a assimetria é positiva ou negativa). Já R^2 , chamado coeficiente de determinação, mede a confiabilidade de uma curva de tendência em relação aos dados – i.e., avalia o grau de ajuste de uma curva calculada para representar um conjunto de valores: quanto mais próximo da unidade, maior o grau de adequação de uma curva que se refere a um conjunto de dados.

Os cálculos de desvios-padrão, de distorção e de R^2 também são indicados por células em amarelo, situadas próximo de cada bloco de dados; contudo, esses cálculos compõem, de ordinário, apenas as planilhas .XLS, e destinam-se a verificar a consistência do tratamento e a auxiliar na identificação de erros na construção da base de dados. Dessa maneira, quando os resultados desses testes adicionais extrapolam os valores esperados para o conjunto de dados, as células que os contêm são assinaladas em vermelho; por exemplo, valores excêntricos de um fator qualquer levam a uma diminuição do valor de R^2 e a um aumento da dispersão da amostra. Nesses casos, procede-se, então, à verificação de

toda a cadeia de fórmulas que relaciona as células a fim de se eliminarem erros de digitação ou de ponteiros; os casos que permanecem em vermelho nas planilhas finais foram verificados, mas não foram atribuídos a erro e, sim, ao caráter diferenciado daquele texto diante de um critério de classificação dos dados⁸.

Nos casos em que se comparavam os resultados obtidos para os *corpora*, decidimos acrescentar uma ferramenta de análise externa, ou seja, que mede o grau de correlação entre dois conjuntos mais abrangentes de valores numéricos. Esse procedimento permite avaliar numericamente a similaridade dos comportamentos dos sistemas, tomados aos pares; na análise comparativa dos resultados individuais, apresentada adiante, os valores de correlação são utilizados para apontar a similaridade de tendências e/ou de distribuição das frequências entre os sistemas lingüísticos representados pelos textos da amostragem. Note-se que esse método não leva em conta a magnitude dos dados, mas a similaridade de seus comportamentos; veja-se o caso hipotético construído abaixo para efeito de ilustração:



Na tabela acima, os valores variam dentro de uma faixa estreita, mas os de D3 estão na ordem das centenas. Entretanto, a correlação entre D1 e D3 é de 100%, e a que se verifica entre D1 e D4 é de -100%, ou seja, é inversa.

Os gráficos que correspondem aos resultados dos cálculos de porcentagem, por fim,

⁸ Exemplo típico e de fácil acompanhamento, já que envolve apenas uma tabela, é o dos adjetivos em italiano (arquivo **i - AN3.xls**), já comentado acima. Em última análise, isso implica, mais uma vez, a necessidade de se avaliar esse dado específico com cautela e de se usar maior número de textos por século.

são elaborados pelo próprio Excel, que oferece diversos tipos diferentes de gráficos (dispersão, *pizza*, área / superfície, bolhas etc.) que podem ser selecionados a partir da marcação de uma porção da tabela e da seqüência *Insert>Graph**⁹. Optamos pelos gráficos em coluna porque esse tipo é bastante corrente, além de permitir a inclusão de linhas de tendência sem poluir visualmente o gráfico; apenas no caso dos gráficos de espanhol, em que há dois séculos adicionais compondo a amostra, optamos por apresentar os gráficos de barras em dois passos (sem as linhas de tendência / RRLs, e, em seguida, com essas linhas), a fim de não carregar demais a apresentação gráfica dos resultados num único passo expositivo. As figuras que representam a comparação entre as bases de dados das línguas individuais compreendem, por sua vez, apenas as linhas de tendência de cada conjunto – as quais são, obviamente, apresentadas em separado na discussão dos dados de cada sistema –, também com o propósito de aliviar visualmente o gráfico.

⁹ No caso de gráficos em barras horizontais ou verticais, o MS Excel sugere, baseado na assimetria da tabela de dados (número de linhas X número de colunas), a variável que estará nas abscissas; isso pode, naturalmente, ser alterado de acordo com as necessidades de representação.

V.3 – Critérios de codificação das ocorrências

Com respeito ao tratamento das ocorrências e 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 pela análise quantitativa encontram-se reproduzidos adiante na seção de tratamento numérico dos dados. Os mesmos gráficos 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 α a ι , e seus desdobramentos; adota-se, aqui, uma linha de trabalho semelhante à de Costa 1998, que situa, em primeiro lugar, a posição do elemento central em apreço, e, em seguida, os fatores que possivelmente contribuem para, ou restringem, (Vincent 2000; Kroch 2000) os fenômenos de mudança de ordem observados no sistema. Esse mesmo autor considera, ainda, vários graus e naturezas da interação entre os fatores - o que, em nosso caso, deverá ser quantitativamente apontado.

α . Posição dos constituintes - assinalaram-se as possibilidades greenbergianas SVC / SCV / VSC / VCS / CSV / CVS, além de se considerarem as ordens SV / VS / VC / CV para as ocorrências cujos argumentos não são explícitos (por elipse) ou não se aplicam (no caso de verbos intransitivos)¹⁰.

<u>SV</u>	por huma bahia asima onde <u>os turquos</u> <u>tem</u> <i>huma forsa bem fraqua de pedra</i>
<u>C</u>	[<i>Pe. Jerônimo Lobo</i>]
<u>SCV</u>	et p{er}o <u>langelo</u> <u>ladimando</u> piena digratia quasi mutandole ilnome p{er}
	[<i>Expositione del Reverendo</i>]
<u>VS</u>	<u>sabedes</u> <u>vos todos</u> <i>que se guarnesçen agora ellos quanto pueden contra nos</i>
<u>C</u>	[<i>Historia Troyana</i>]

¹⁰ De maneira a não menosprezar as diferenças entre os casos de elipse e os que envolvem a intransitividade do verbo, procurou-se observar se a estrutura de superfície poderia ser determinante de uma ou de outra ordem; ademais, distinguem-se os dois tipos de verbos por meio de critério próprio (*regência verbal*, explicitado adiante), e não através do uso de subcategorias que poderiam sobrecarregar indevidamente um dos fatores da análise pela sobreposição de traços sintáticos e semânticos.

VCS	aos quaaes nom deve <u>fazer perjuizo o que outras fezerom em contrairo</u> [<i>Leal Conselheiro</i>]
CSV	mas de que <u>lo vos vierdes</u> [<i>Libro del caballero y del escudero</i>]
CVS	<i>Molto auara</i> mi è <u>V.S.</u> delle sue lettere, da me [<i>Viaggi del pellegrino</i>]

SV	e <u>uma estatua de marmore</u> ... <u>ennegrecendo</u> a um canto na lenta humidade [<i>Os Maias</i>]
VS	All'opposto ben diversamente <u>corrano</u> <u>le cose</u> per la parola <i>Tamehu</i> [<i>Geografia etnologica e storica</i>]
VC	como al <u>trabar diálogo</u> con una persona por su nombre sólo [<i>Castro Urdiales</i>]
CV	de uma estratégia conjunta e complementar que <u>os tornasse</u> no fiel de balança [<i>Jangada de pedra</i>]

Distingue-se ainda o conjunto de ocorrências em que o complemento é anteposto ao verbo (assinalado, no texto, como [CV] e, nos gráficos, como CV) daquele em que é posposto ao mesmo (assinalado, no texto, como [VC] e, nos gráficos, como VC). Esse procedimento segue a proposta de Lehmann 1978, segundo a qual as construções baseadas nos arranjos de V e de C correspondem aos tipos lingüísticos primários, de maneira que as línguas são classificadas como VC ou como CV.

β. Natureza do sujeito – pretende-se investigar se a complexidade do sintagma nominal tem alguma contribuição para a mudança, como proposto, por exemplo, por Abeillé & Godard 1997 para o francês. Aplicam-se as mesmas classes à natureza do objeto: pleno, para aqueles constituídos apenas por [determinante + nome]; pronominal, no caso de pronomes retos ou oblíquos; e oracional, se representados por oração subordinada.

S pleno	De como <u>ElRey</u> partio de Lisboa, e do que passou [<i>Príncipe Dom João</i>]
S pronominal	ma niuna di esse (direi <u>io</u>) nè può mostrar maggiori, nè più fondate ragioni [<i>Christoforo Colombo</i>]
S oracional	es dudoso <u>que la judicatura se los hubiera reconocido</u> [<i>La ciudad de los prodigios</i>]

C pleno	dice{ <u>n</u> }do io priego idio che tifaccia <u>ogni bene</u> / et che tifaccia
----------------	---

	saluo [<i>Expositione del Reverendissimo</i>]
C pronominal	por que ella <u>o</u> faz pryncypalmente per sua virtude e bondade [<i>Leal Conselheiro</i>]
C oracional	Creían <u>estar en presencia de un santo</u> [<i>Ciudad de los prodigios</i>]

γ. Status do pronome – trata-se de uma particularização de β acima: se o caráter do sujeito ou do objeto for pronominal, deve-se observar se esses pronomes estão unidos ou não ao verbo; portanto, as classes logicamente possíveis são S ligado / S não-ligado / C ligado / C não-ligado. Com esse traço, pretende-se verificar a) a extensão cronológica da aplicação do pronome ligado ao verbo – especialmente para o espanhol e o italiano, em que o complemento pronominal do verbo ocorre, com frequência, unido a ele; e b) se é possível identificar particularidades em seu comportamento diante de sintagmas plenos e oracionais – que têm um peso muito maior que o de pronomes –, como previsto por Abeillé & Godard 1997.

S ligado	e dal non <u>hauer'egli</u> ... trouato mai notitia alcuna, che certa fosse; e pure [<i>Christoforo Colombo XVII</i>]
S não-ligado	Por dinero, respondía <u>él</u> [<i>Ciudad de los prodigios</i>]
C ligado	et p{er}o l'angelo <u>ladimando</u> piena digratia quasi mutandole il nome p{er} [<i>Expositione del Reverendissimo</i>]
C não-ligado	com cozimentos, e <u>o</u> aviventem na cor ou a quintura do sol o faz nos pedasos [<i>Pe. Jerônimo Lobo</i>]

δ. Ambiente pronominal - ainda com referência a pronomes explícitos, julga-se necessário investigar até que ponto a colocação dos mesmos é determinada por elementos vários, como partículas negativas ou relativos; as possibilidades cobertas por este item são Neg + Pro / qu- + Pro / outros + Pro

Neg + Pro	estão, porque <u>as não</u> ha de nenhuma, mas serras escavadas [<i>Pe. Jerônimo Lobo</i>] ma se uogliamo parlare di quello Re <u>non lo</u> nomineremo signore in Ferrara [<i>Expositione del Reverendissimo</i>]
qu- + Pro	sy nos de aquesta vez estamos <u>que nos</u> non vengemos de Ia desonrra [<i>Historia Troyana</i>]

	et quello <u>che</u> <u>lha</u> guardisi dalli ladri <u>che</u> <u>lanongli</u> sia furata perche epso [<i>Expositione del Reverendissimo</i>]
outros + Pro	que vio correr huma vea de agoa vermelha <u>e</u> <u>o</u> atreuhia a algum [<i>Pe. Jerônimo Lobo</i>] <u>pero</u> <u>la</u> aplicaban a su modo, expeditivamente [<i>Ciudad de los prodigios</i>]

- ε. Ambiente oracional - pretende-se investigar a independência dos ambientes hipotáticos ou paratáticos em relação às orações principais; com esse objetivo, as subclasses associadas a este fator são apenas principal / coordenada / subordinada / coordenada-subordinada, e não são contemplados, por exemplo, os subtipos oracionais de estruturas hipotáticas. Reporta-se a sugestão desse fator à consideração do exposto por Bybee 2001.

principal (Pr)	Faz êle uma doação ao mosteiro de Grijó a três de Outubro [<i>Conde Dom Henrique</i>]
coordenada (Co)	y el paisaje fue tomando serios acordes de grandeza salvaje [<i>Monasterio de Silos</i>]
subordinada (Sb)	Similmente <u>quando</u> noi incominciamo afare oratione alle <u>primariamente</u> [<i>Expositione del Reverendissimo</i>]
coordenada-subordinada (CS)	de rogar para que ele ... nelas entrasse <u>e</u> <u>derramasse</u> as suas penúltimas seivas, se [<i>Jangada de pedra</i>]

- ζ. Regência e tempo do verbo - partindo-se do pressuposto de que a mudança sintática possa ter origem em outros níveis de organização do sistema, e visto que a delimitação prospectiva do conjunto de condicionadores envolve a investigação da natureza da complementação do verbo (Bybee *et al.* 1994), propõe-se uma observação sobre a expressão da regência do verbo no VP (direta / indireta / direta-indireta / ligação / intransitivo) e sobre a estrutura configuracional que o mesmo assume (finito / reduzido / locução / composto).

direta	<u>Sigue</u> la carretera su cinta silenciosa llena de claridad cegadora [<i>Monasterio de Silos</i>]
---------------	---

indireta	Parésceme que te <u>plaze</u> saber, muy caro amigo [<i>Virtuossas mugeres</i>]
bitransitivo	e non <u>giouando</u> a Muhammèd Bascià ... tutte le diligenze, che hà fatte [<i>Viaggi del pellegrino</i>]
ligação	omnipote{n}te <u>essendo</u> Iesu christo figliuolo uero di ambi due [<i>Expositione del Reverendissimo</i>]
intransitivo	E la mezquina ... <u>lloraba</u> muy fuertemente [<i>Libro de los enxemplos</i>]

finito	no seculo seguinte já nenhuns vestigios delle <u>restavam</u> [<i>Castelo de Faria</i>]
reduzido	ch'è vno stupore, il vedere vn castello tãto ristretto si sito; <u>dare</u> tãta somma [<i>Compendio delle regole</i>]
locução	<u>Deixaram</u> por isso ... <u>sumir</u> nas paredes de um claustro pedras que foram [<i>Castelo de Faria</i>]
composto	con el afecto que le <u>había ofrecido</u> los otros hijos [<i>Vida de Sor M. de Ágreda</i>]

η. Classe semântica do verbo - da mesma forma como foi feito em Cohen 1986/1990 e Totaro 1998 para os adjetivos, aventa-se, aqui, a possibilidade de o condicionamento da mudança estar relacionado às características semânticas dos verbos; utilizaram-se, para a classificação dos verbos, as definições apresentadas por de Beaugrande 1997 (cap. IV)

ATTRibutive	Aprazível é o sitio, sombreado de velhas arvores [<i>Castelo de Faria</i>]
CIRCumstantia	se <u>encuentra</u> el peregrino en un quartito de limpio y modesto adorno l [<i>Castro Urdiales</i>]
COGNitive	Che significhino, non sò : ma m' <u>imagino</u> le insegne della Giustitia, e del gouerno [<i>Viaggi del pellegrino</i>]
DEVELOpment	<u>heláronse</u> muchos ríos tan vecinos al mar, que formaba margen el hielo [<i>Marqués de San Felipe</i>]
DISPositive	& de' proprij cõtadini che <u>lauorano</u> detti luoghi [<i>Compendio delle regole</i>]
EMOTional	e se Affonso n'aquella edade <u>amava</u> o socego de Santa Olavia [<i>Os Maias</i>]
ENACtive	de haber llegado <u>llegó</u> también el Príncipe Cardenal Alberto [<i>Alonso de Contreras</i>]

EVALuative	por alguu, <u>achando</u> nom boo meu conselho, me nom prasmarr [Leal Conselheiro]
EXISTential	il Quatrefages e l'Ilamy ... opinarono che ... <u>esista</u> un certo numero di diversi tipi [Geografia etnologica e storica]
IDENtity	Los buenos <u>son</u> los que dirigen las almas á la salvación [Torres Villaroel]
PERCeptual	E que quiser <u>veer</u> a pronunçiação propria de .am. & quam diferente he [Duarte Nunes do Lião]
POSSessive	até os limites que <u>possuem</u> os cristãos para a parte do Meio-dia [Conde Dom Henrique]
PRODuct	perche questa soldatesca di Grecia <u>fà</u> vna gran somma [Viaggi del pellegrino]
SEMantic	e qual Colomba altresì... <u>annuciar</u> buona nuoua dell'Euangelica pace [Christoforo Colombo]
VOLitional	<u>quiero</u> yo que seas tu señor de aqueste rreyno [Historia Troyana]

- θ. Agente e Paciente - outro par de traços que pretende detectar a possível associação entre a mudança de ordem e a semântica; os valores permitidos são humano / não-humano

agente	humano	quasi che vogliamo esprimere il nostro desiderio dice{n}do [Expositione del Reverendissimo]
	não-humano	O Estado da fortuna pode cõceder ou tirar fauor aos estudos liberaes [Fernão de Oliveira]
pacient e	humano	La infinita bondad y hermosura de Diós <u>la</u> cautivó [Vida de Sor María de Ágreda]
	não-humano	de guerra, q continuamente faziaõ estragos por mar no Reyno de Hespanha [Príncipe Dom João]

- ι. Item lexical – o comportamento da mudança, liderado por elementos específicos, deve também ser coberto por este fator; com relação ao verbo, os itens lexicais são divididos em grupos segundo a frequência em que ocorrem nos *corpora* e de acordo com sua conjugação. Não se tecem, aqui, considerações sobre as diferentes conjugações de um verbo de língua para língua – como se observa entre recibir (E) e receber (P), aggrandire (I) e engrandecer(P).

<i>freqüência</i>	<i>intervalo</i>	<i>exemplos portugueses</i>
alta	$F > 66\%$	ser / fazer
média	$33\% < F < 66\%$	dizer / ter
baixa	$F < 33\%$	aplicar / enobrecer

<i>conjugação</i>	<i>exemplos</i>
I	Onde noi <u>salutiamo</u> coloro liquali amiamo [<i>Expositione del Reverendissimo</i>]
II	<u>mantendo</u> ... a sua posição com relação aos pontos cardeais por que [<i>Jangada de pedra</i>]
III	la traycion que <u>urdio</u> el conde do Yllan et ell arçobispo [<i>Crónica General de España</i>]

Na base de dados, a codificação dos traços atribuídos às ocorrências obedece à tabela de correspondências apresentada abaixo:

PC	posição dos constituintes
a	SVC
b	SCV
c	VSC
d	VCS
e	CSV
f	CVS
g	SV
h	VS
i	VC
j	CV

NS	natureza do sujeito
k	pleno
l	pronominal
m	oracional

NO	natureza do objeto
n	pleno
o	pronominal
p	oracional

SP	status do pronome
-----------	--------------------------

q	S ligado
r	S não-ligado
s	C ligado
t	C não-ligado

AO	ambiente oracional
1	principal
2	coordenada
3	subordinada
4	coordenada em subordinação

RV	regência do verbo
u	direto
v	indireto
w	dir./indireto
x	ligação
xx	intransitivo

AP	ambiente pronominal
y	Neg+Pro
z	qu- + Pro
zz	outros + Pro

TV	tempo do verbo
α	finito
β	reduzido
γ	locução
δ	composto

agent	agentividade	
A	h	agente humano (Ah)
	n	agente não-humano (An)
P	h	paciente humano (Ph)
	n	paciente não-humano (Pn)

CSV	classe semântica do verbo
ATTR	atributivo
CIRC	circunstancial
COGN	cognitivo
DEVEL	evolutivo
DISP	dispositivo
EMOT	emocional
ENAC	deslocamento
EVAL	avaliativo
EXIS	existencial

IDENT	identificador
PERC	perceptual
POSS	possessivo
PROD	resultante
SEM	semiológico
VOLIT	volitivo

CONJ	conjugação
1	
2	
3	

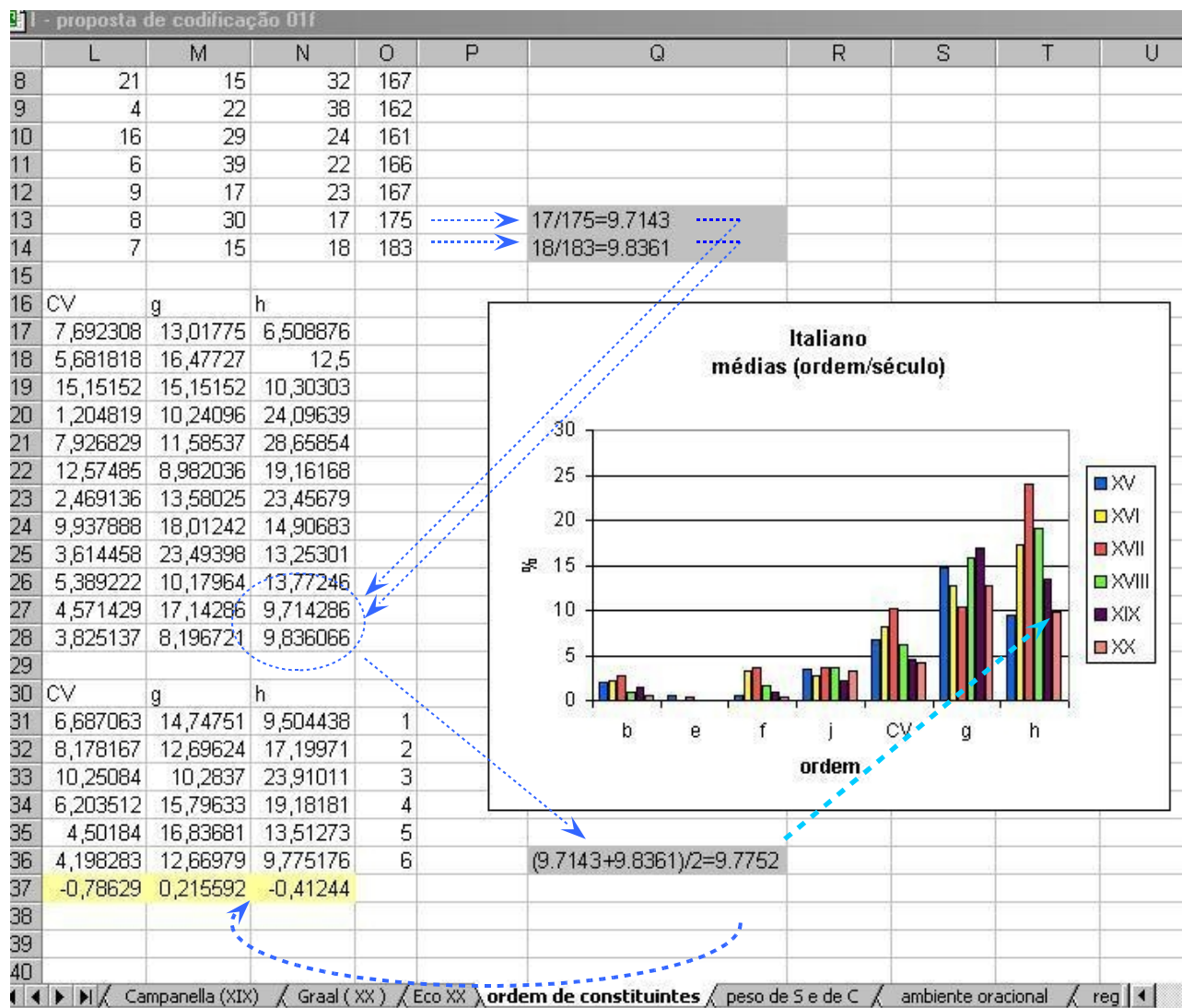
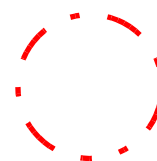
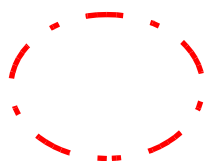


Fig. PN.6c – Ilustração dos cálculos básicos que levaram à construção dos gráficos de tratamento estatístico dos dados. As células em cinza não fazem parte dos arquivos .XLS

originais, e representam, na figura, apenas o cômputo feito entre os blocos numéricos. Assim, as células Q12 e Q13 mostram o cálculo das porcentagens simples de ocorrência da classe h nos textos italianos do século XX, e que é realizado pelas fórmulas que constam das células N27 e N28. Em seguida, procede-se ao cálculo das médias de ocorrência por século, que é tornado visível na célula Q36 e é efetivamente realizado na célula N36; esse é o valor que corresponde à altura da barra h para o século XX, apontado pela seta em azul-claro. As células amarelas contêm uma fórmula que calcula a inclinação do conjunto de valores numéricos – no caso da classe h , o bloco N31 a N36 – a partir das células de referência O31 a O36 (que servem para o cálculo da iRRL para todas as classes do gráfico).

Quando foi preciso ressaltar a amplitude das curvas menos expressivas – por exemplo, nos casos em que se buscava perceber se os máximos locais teriam valores coincidentes para as abscissas –, procedeu-se à normalização das frequências para 100%. Esse procedimento aproxima as curvas equiparando suas frequências a uma base individual, e não à mesma base. Em outras palavras, cada curva é (re)calculada convertendo-se o número total de ocorrências que contêm um traço em 100% (abaixo, em α , para *ambiente oracional*), em vez de se computarem os resultados individuais sobre a soma de todas as ocorrências (em β).

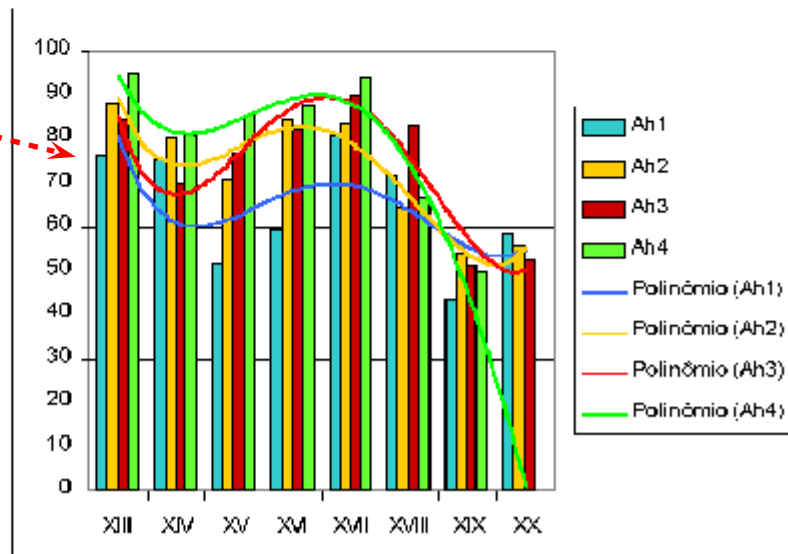
α



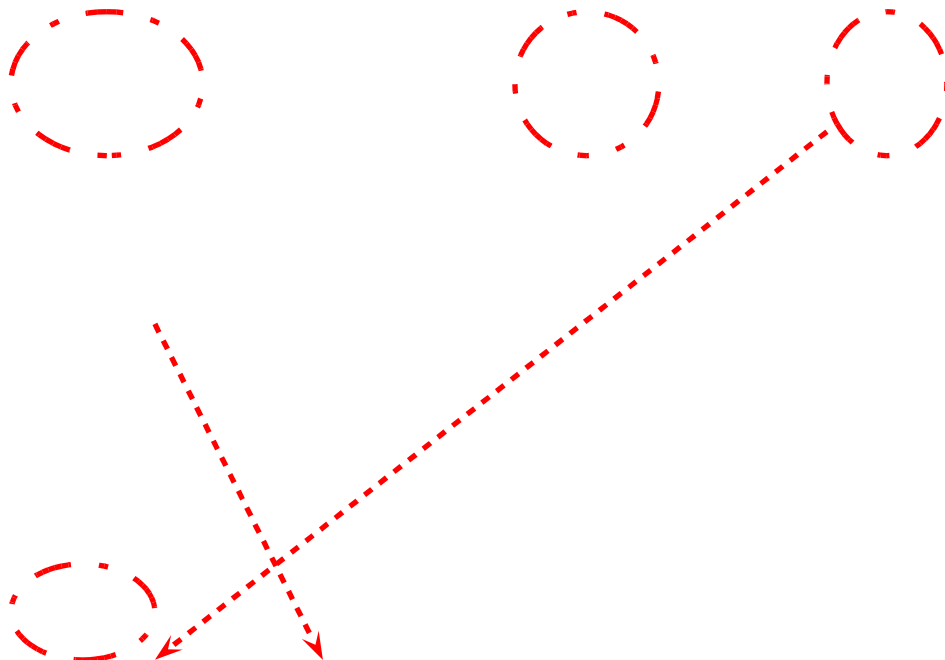
	1	Ah	An	Ph	Pn	σ
XIII		12	5	0	11	17
		9	2	5	3	11
XIV		19	2	3	14	24
		10	3	2	8	14
XV		14	1	3	4	16
		3	15	1	7	19
XVI		22	10	0	21	32
		1	1	0	1	2
XVII		35	12	15	19	49
		41	3	4	23	46
XVIII		34	20	10	30	55
		45	10	7	33	55
XIX		25	14	0	22	52
		20	32	7	37	52
XX		76	15	11	54	91
		30	52	5	38	88

	1	% Ah	% An
XIII		70,58	29,41
		81,81	18,18
XIV		79,16	8,333
		71,42	21,42
XV		87,50	6,25
		15,78	78,94
XVI		68,75	31,25
		50,00	50,00
XVII		71,42	24,48
		89,13	6,52
XVIII		61,81	36,36
		81,81	18,18
XIX		48,07	26,92
		38,46	61,53
XX		83,51	16,48
		34,09	59,09

médias1	Ah	An
XIII	76,20	23,79
XIV	75,29	14,88
XV	51,64	42,59
XVI	59,37	40,62
XVII	80,27	15,50
XVIII	71,81	27,27
XIX	43,26	44,23
XX	58,80	37,78



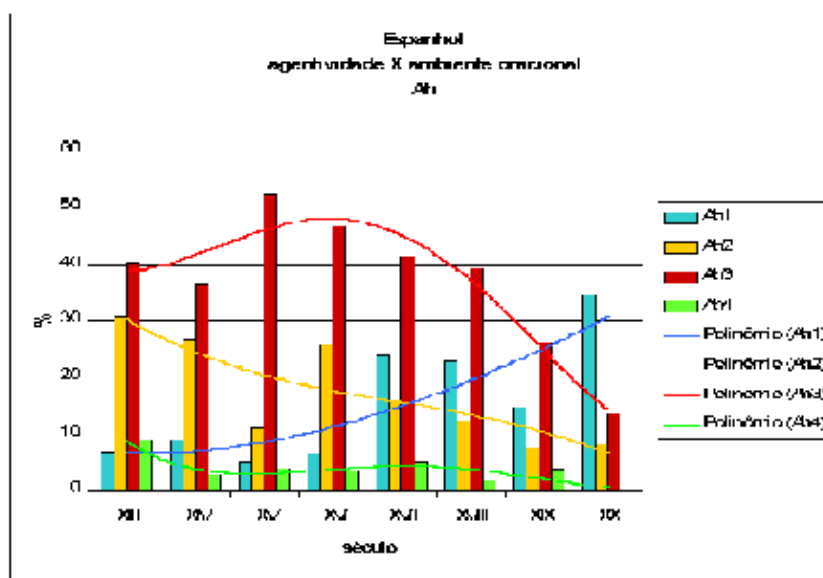
β



	1	Ah	An	Ph	Pn	σ
XIII		12	5	0	11	157
		9	2	5	3	155
XIV		19	2	3	14	162
		10	3	2	8	165
XV		14	1	3	4	168
		3	15	1	7	169
XVI		22	10	0	21	178
		1	1	0	1	162
XVII		35	12	15	19	154
		41	3	4	23	167
XVIII		34	20	10	30	172
		45	10	7	33	176
XIX		25	14	0	22	156
		20	32	7	37	153
XX		76	15	11	54	155
		30	52	5	38	151

	1	% Ah	% An
XIII		7,64	3,18
		5,80	1,29
XIV		11,72	1,23
		6,06	1,81
XV		8,33	0,59
		1,77	8,87
XVI		12,35	5,61
		0,61	0,61
XVII		22,72	7,79
		24,55	1,79
XVIII		19,76	11,62
		25,56	5,68
XIX		16,02	8,97
		13,07	20,91
XX		49,03	9,67
		19,86	34,43

médias1	Ah	An
XIII	6,72	2,24
XIV	8,89	1,53
XV	5,05	4,74
XVI	6,49	3,12
XVII	23,64	4,79
XVIII	22,67	8,65
XIX	14,55	14,94
XX	34,45	22,06



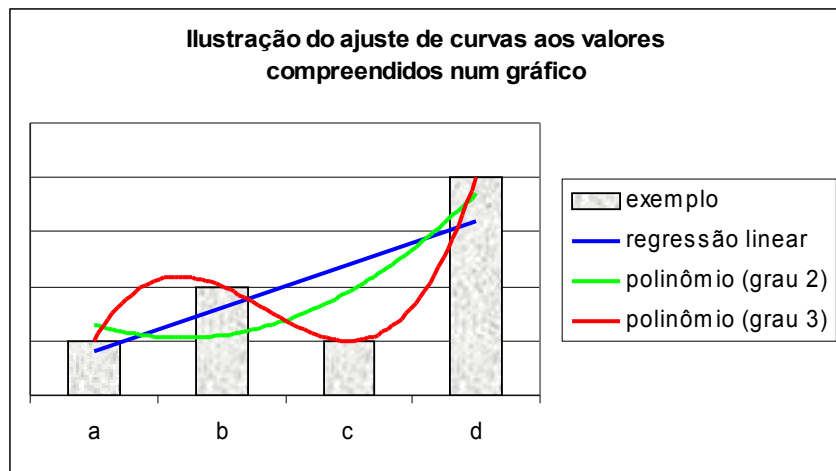
Observe-se que as matrizes em α e em β foram transpostas com o objetivo de permitir o agrupamento dos resultados afins e a construção dos gráficos comparativos ; efeito idêntico é obtido, no Excel, pela seleção de uma forma diferente (linha / coluna) de se combinarem as informações de uma tabela.

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 de uma curva polinomial¹¹, mas é natural que haja um ou outro dado que aumente a dispersão do grupo. Esse efeito individual pode ser reduzido, como visto, graças à ampliação longitudinal (para alcançar outros séculos) ou transversal (utilizando-se outros textos como 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.

A fim de demonstrar o efeito produzido pela inserção de pontos adicionais nos extremos do intervalo de amostragem, pode-se conduzir um teste que envolve o cálculo sucessivo de RRLs para amostras de tamanho crescente. Na figura abaixo, utilizam-se os dados de frequência de orações subordinadas em espanhol, sistema que tem o intervalo de amostragem mais longo neste estudo. A primeira RRL (em magenta) foi traçada considerando-se apenas os dois primeiros séculos; a segunda (em azul), os três primeiros, e assim por diante, até a sétima (em vermelho), que considera os dados de todos os oito séculos da amostra. Notam-se claramente a) a progressão das inclinações das retas, conforme indicação feita pelas setas em cinza, de maneira a se aproximarem as RRLs daquela que reúne todos os valores de frequência, e b) o salto provocado pelo ponto referente ao século XV, que distoa dos demais – mas não impede o ajuste gradual das iRRLs. É óbvio, portanto, que os intervalos de amostragem mais longos são mais

¹¹ Retas de regressão linear e curvas que se ajustam aos pontos dos gráficos são traçadas automaticamente pelo MS Excel, que permite adequar o grau de refinamento das curvas de tendência. Nota-se, na figura abaixo, que 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.



adequados, uma vez que as tendências de conjunto representadas pelas RRLs são definidas com mais precisão e propriedade; além disso, já que essas retas são calculadas a partir da consideração de todos os dados do intervalo, um ponto excêntrico causa maior distúrbio em amostras reduzidas (simbolizadas pela reta 3.2, em azul, que utiliza apenas três pontos), mas esse efeito é gradualmente minimizado pelos pontos adicionais, que levam a uma diminuição paulatina das inclinações a partir dessa mesma reta 3.2. Por fim, percebe-se que, a menos que os pontos sucessivos distoem das tendências mantidas pela língua – o que equivaleria a uma mudança catastrófica em Lightfoot 1991 –, as RRLs não se sucedem em saltos contrários à propensão do sistema (simbolizados pela maior inclinação positiva da reta azul em relação à da reta magenta, em vez de se respeitar a tendência decrescente do conjunto).

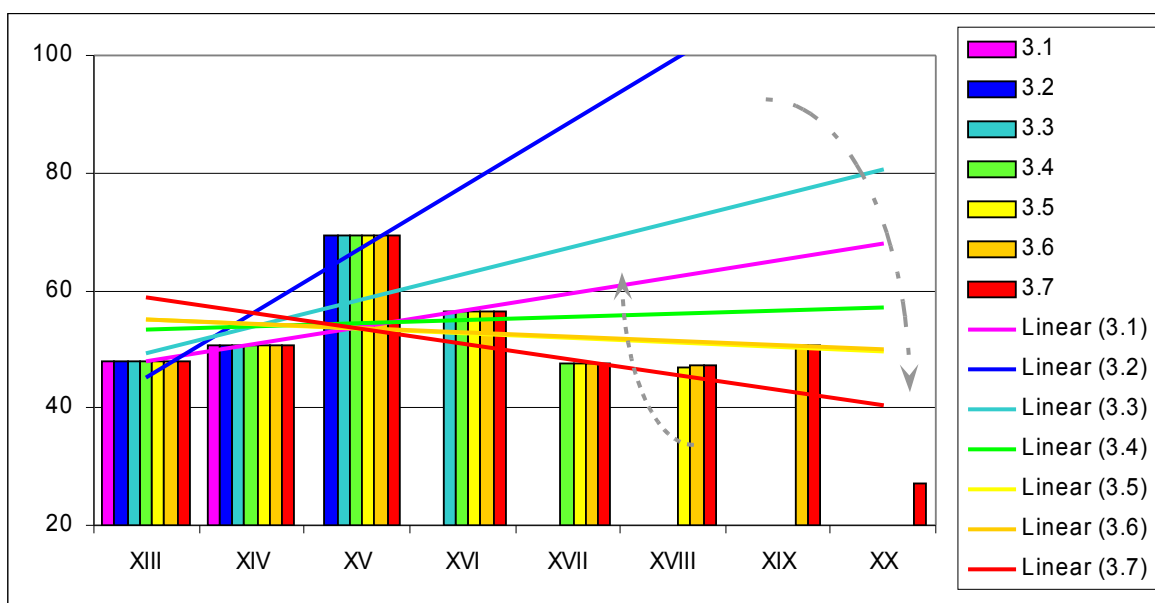


Fig. PN.6d – Demonstração da mudança de iRRL diante da inclusão de pontos adicionais no extremo do intervalo de amostragem.

Em segundo lugar, as curvas polinomiais têm traçados que são definidos por equações construídas automaticamente sobre a amostragem. Necessita-se, nesse procedimento, de um mínimo de três pontos para se lançar a expressão gráfica de um polinômio; infere-se daí que a precisão do traçado de uma curva é maior nas regiões internas do intervalo compreendido pela amostra, e deve ser avaliada com mais cautela à medida que se aproxima das margens. Veja-se, em PN.7, o efeito provocado pela adição de

um ponto ao final do intervalo: no primeiro gráfico, a ausência de pontos nos séculos XXI e XXII leva a uma projeção sempre crescente para a curva azul, enquanto as demais tendem a diminuir; se se acrescenta um ponto hipotético (equivalente à média dos valores dos séculos XIII a XX) no século XXI, a projeção se reverte para a maioria das curvas no segundo gráfico. Observação idêntica pode ser feita diante da possibilidade de expansão da amostra de maneira a contemplar os séculos anteriores ao XIII.

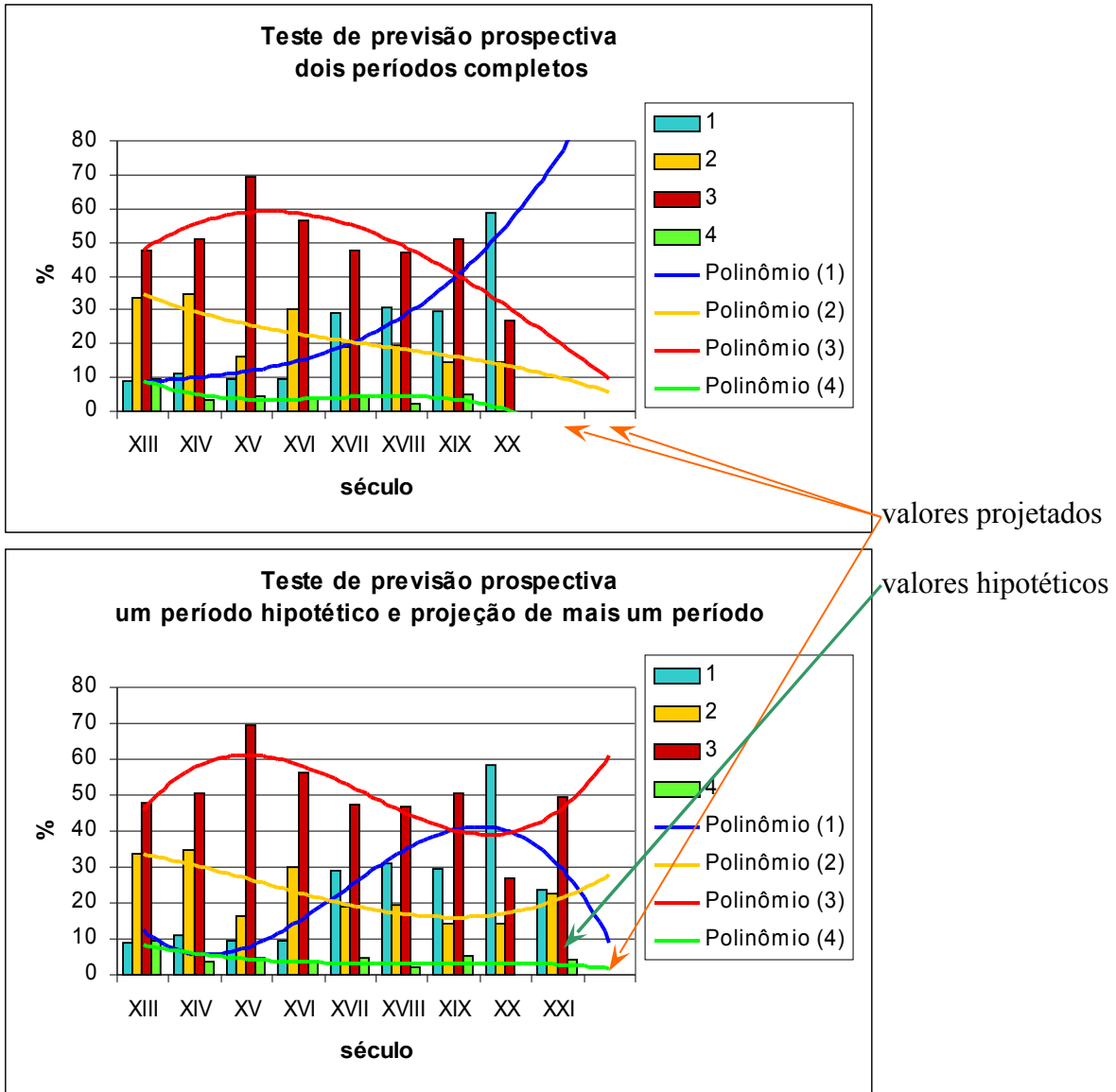


Fig. PN.7 – demonstração da sensibilidade das regiões periféricas da amostragem no traçado de curvas de tendência. No segundo gráfico, os valores plotados no século XXI são hipotéticos e equivalentes à média dos valores dos séculos anteriores, e provocam alteração nos perfis das curvas de tendência no final do intervalo.

Tratamento numérico das ocorrências: o *corpus* português

As ocorrências do *corpus* português se distribuem, segundo a ordem dos constituintes, em torno de quatro categorias sintáticas principais. As freqüências absolutas se concentram preferencialmente em torno das ordens **a** (SVC), **i** (VC), **g** (SV) e **h** (VS), e o *corpus* português apresenta, por vezes, comportamentos diferentes para esses subconjuntos de ocorrências: por exemplo, a categoria **a** (SVC) tem iRRL positiva em italiano e em espanhol (q.v. adiante), apesar de ser negativa (-1,91) em português; já a classe **i** (VC) tem iRRL negativa para os três *corpora*, mas é mais alta em português (-3,00) do que nos outros dois sistemas, e contribui para a inclinação elevada da classe [VC]. As categorias **g** (SV) e **h** (VS) têm, respectivamente, iRRLs iguais a 1,50 e -,74.

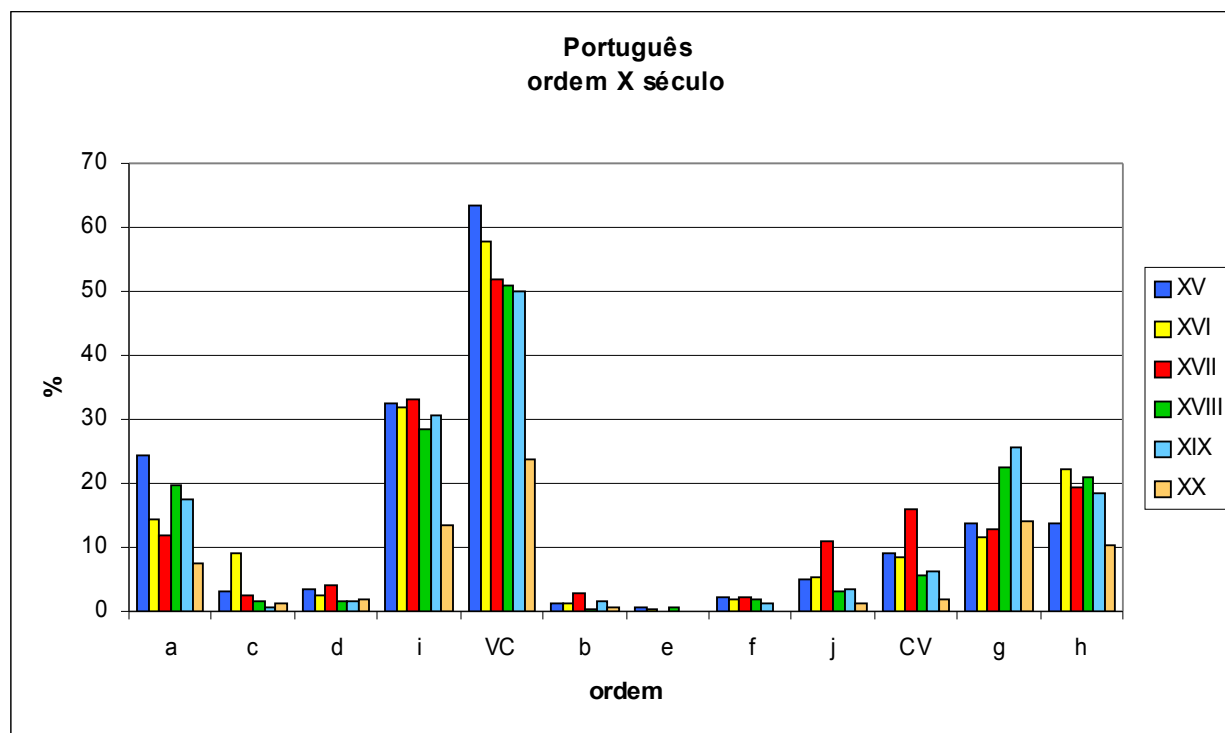


Gráfico P1 – distribuição geral das ocorrências segundo a ordem dos constituintes.

É interessante observar, nesse gráfico geral, as proporções absolutas de ocorrência das classes do fator utilizado na análise tipológica (cf. Greenberg 1966): a ordem **a** (SVC) corresponde, na verdade, ao grupo mais freqüente dentre os que apresentam os três constituintes

S-V-C, mesmo atingindo valores abaixo dos 30% de frequência absoluta. As ordens **i** (VC), **g** (SV) e **h** (VS), que implicam a ocorrência de apenas um dos argumentos do verbo, têm, comumente, frequências mais elevadas que as de **a** (SVC) no *corpus* analisado. Em outras palavras, se a definição de perfis nítidos de decréscimo ou de aumento de frequências no *corpus* é considerada indicativo a) de representatividade da língua pelos textos e b) de fenômenos de mudança nesse mesmo *corpus*, o Gráfico P1 mostra que o português é, de fato, caracterizado por uma ordem **a** (SVC) decrescente; entretanto, a despeito disso, sempre foi **i** (VC) por excelência e tem-se transformado num sistema predominantemente **g** (SV) + **i** (VC), em que há preferência pela expressão de apenas um dos argumentos do verbo. Em linhas mais gerais, isso indica claramente a necessidade de se trabalhar com frequências normalizadas (q.v. seção que trata da Metodologia) em estudos que focalizem a mudança tipológica, uma vez que as ordens de **a** (SVC) a **f** (CVS), que correspondem às combinações dos três constituintes S-V-C, não representam todas as possibilidades de estruturação de orações que os envolvem.

O cruzamento desse critério de classificação das ocorrências com os demais fatores produz indicações interessantes sobre o alcance da mudança em português. Procedemos, inicialmente, à verificação da interseção de ordem e de *agentividade*, que forneceu os perfis sumarizados no gráfico P2 abaixo. Destaca-se, no mesmo, a predominância das associações Ah e Pn para os grupos [VC] e [CV], paralelamente à baixa expressividade das combinações An e Ph (sistematicamente abaixo dos 10%). Os comportamentos dos grupos mais comuns na amostra – e que também são as classes prototípicas para agente e para paciente – são resolvidos por regressão linear, observando-se facilmente que as frequências de ambos são idênticas: $iRRL(Ah) = -2,18$, $iRRL(Pn) = -2,1$. Dito de outro modo, as construções que envolvem verbos em que não se identificam agente nem paciente, como as apresentadas abaixo, são cada vez mais comuns no *corpus*.

"no seculo seguinte já nenhuns vestigios d'elle restavam" [Castelo de Faria, séc. XIX]

"D'ahi partia um amplo corredor, ornado com as peças ricas de Bemfica" [Maias, séc. XIX]

"a população subia até os telhados, e pendia, como cariátides, das telhas" [Serões, séc. XX]

"a pretexto de que a capital espanhola se encontra ... no centro geométrico" [Jangada de Pedra, séc. XX]

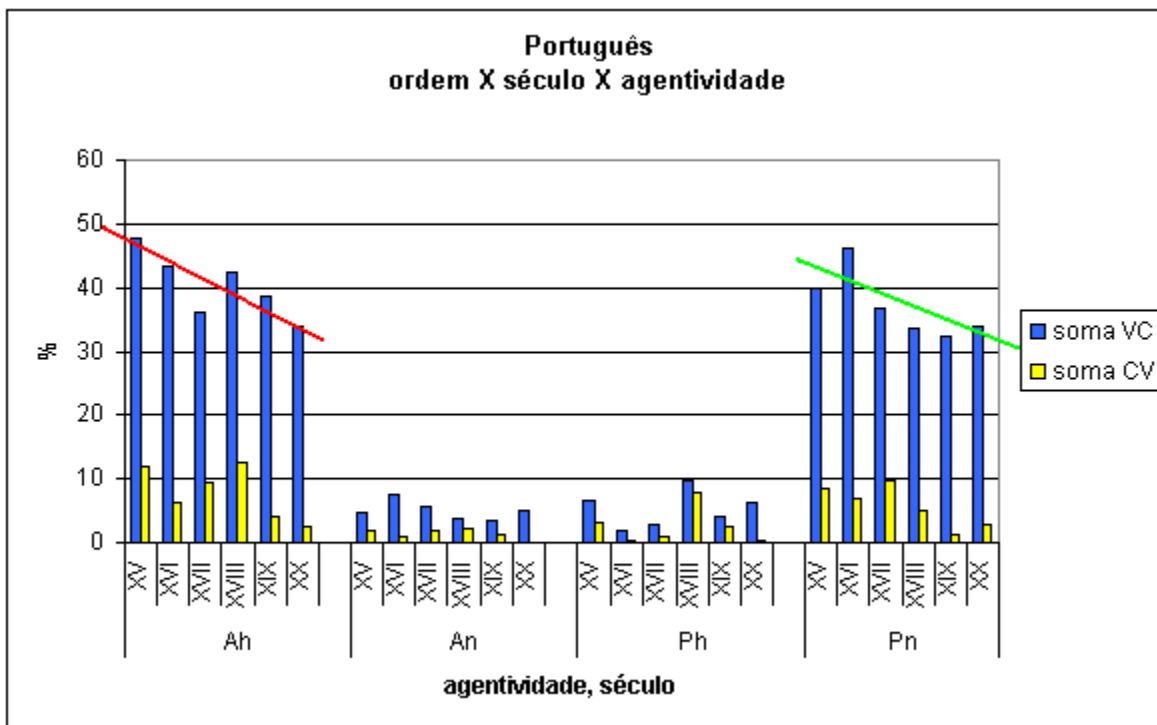


Gráfico P2 – distribuição das freqüências relativas gerais em função da agentividade.

Passamos, em seguida, à observação do comportamento das subclasses sintáticas mais freqüentes no que toca à *agentividade*, cujo perfil é apresentado nos gráficos P3.

Apesar de o perfil geral de *a* (SVC) se refletir na predominância de Ah e de Pn, nota-se que esses dois subconjuntos mais freqüentes têm taxas de decréscimo ($iRRL(Ah) = -0,29$; $iRRL(Pn) = -0,48$) inferiores às apresentadas no gráfico P2 acima. Isso sugere que a subclasse *a* (SVC) acompanha a evolução do conjunto, mas não é a principal responsável pela diminuição sensível das freqüências das combinações Ah e Pn.

São exemplos de ocorrências associadas a essa classe:

com Castella. eu quero ordenar huuas festas rreaes que durem todo hum anno [Ceuta XV]
e o mesmo Paio Guterres restaurou o mosteiro de Tibães e jaz sepultado [CondeDHenr XVII]
os reis e os vassallos deixam voluntariamente os seus reinos e as suas pátrias [IlhAmor XVIII]

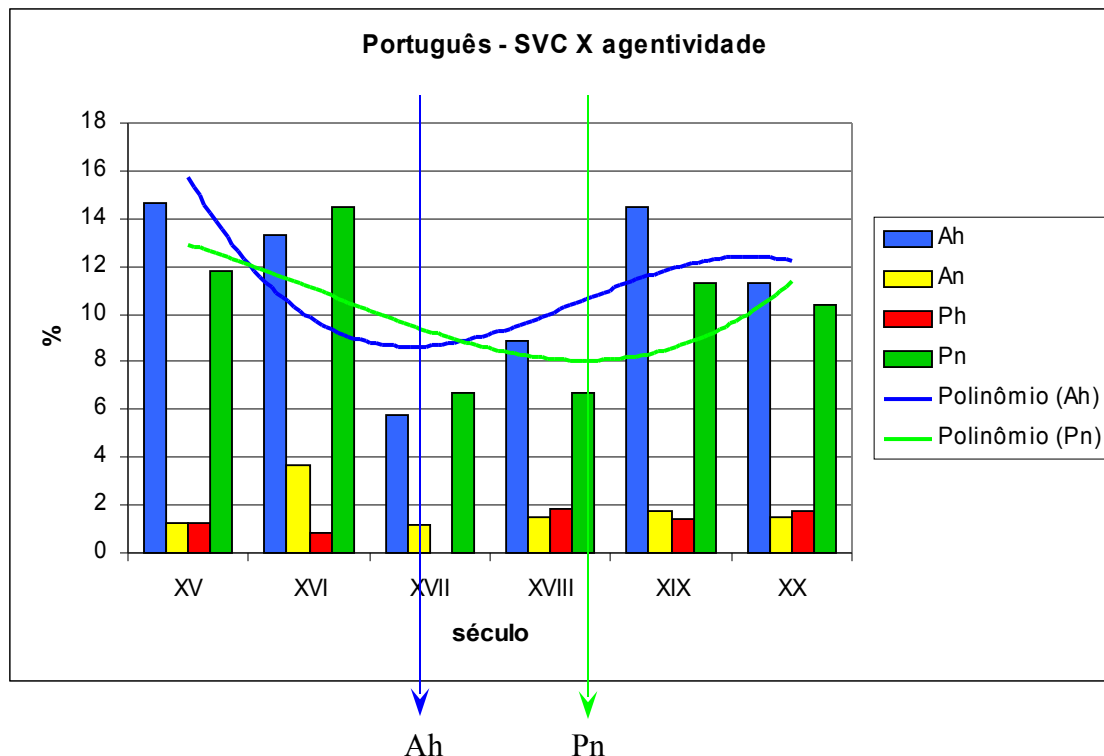


Gráfico P3a – distribuição das freqüências relativas de SVC em função da agentividade.

Outra das categorias mais expressivas é *i* (VC¹²), cujo comportamento na amostra é traduzido no gráfico P3b abaixo. É imediatamente perceptível o decréscimo das classes mais proeminentes, a exemplo do que se disse para *a* (SVC) acima; entretanto, pode-se observar que as inclinações das retas que resolvem a evolução dessas mesmas categorias são maiores do que a média para o conjunto. Isso indica que, em termos do fator *agentividade*, a subclasse *i* (VC) é uma das principais determinantes do comportamento de [VC], o grupo que lhe é homônimo. Além disso, a freqüência relativa dessa subclasse é da ordem do dobro da observada para *a* (SVC), o que confere também maior peso ao perfil do gráfico P3b. Veja-se ainda que os perfis das curvas para as classes mais comuns, Ah e Pn, não são coincidentes, e que a curva de Ah parece ter um período menor que o de Pn.

Exemplos de ocorrências associadas ao traço *i* (VC) são:

quando com ella suprimos o .m. com que formamos alguus diphthôgos [DNLião XVI]
o qual arancava do fundo molhos de erva que era a que fazia a dita aparencia [PeJerLobo XVII]
arripiando-se, rindo, ao metter o pésinho n'agoa [Maias XIX]

¹² Conforme observado à pág. 36, trata-se, aqui, da classe *i* (VC), e não de [VC], a somatória de todas as ocorrências que verificam a anteposição de V a C.

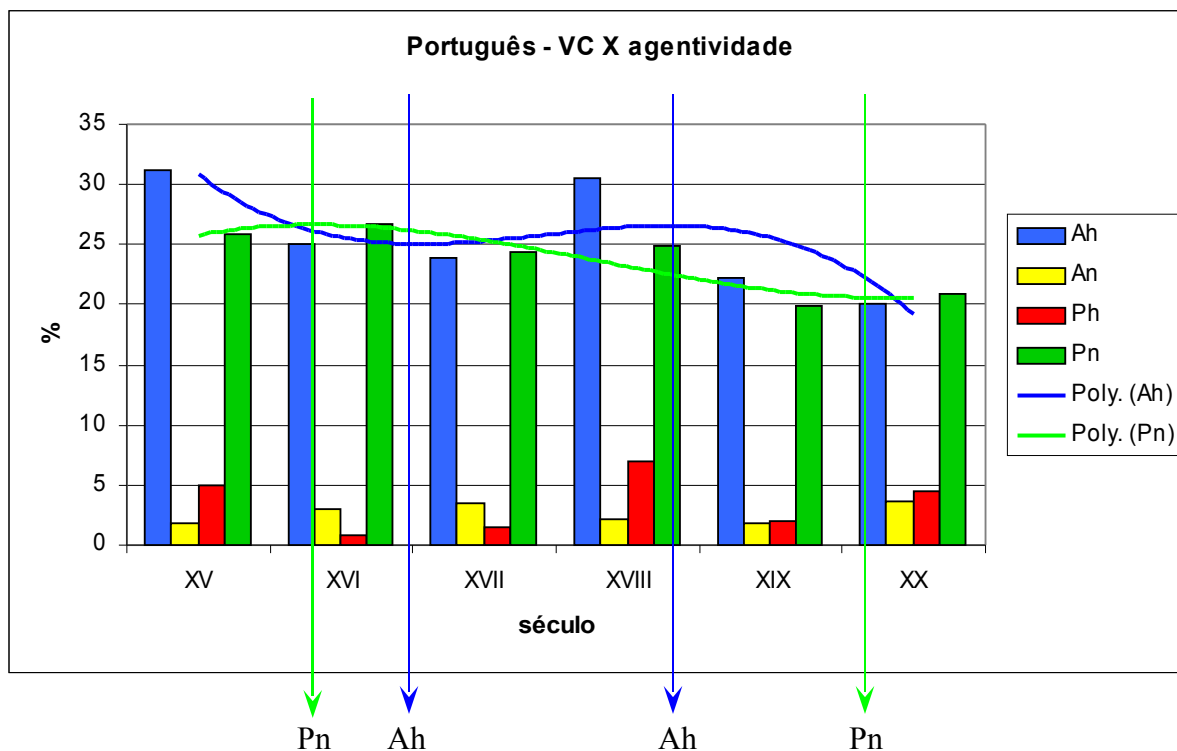


Gráfico P3b – distribuição geral das ocorrências VC segundo o fator *agentividade*.

O próximo grupo de dados a merecer considerações em maior detalhe é a classe **j** (CV), cuja disposição de constituintes é especular em relação à anterior, mas apresenta, naturalmente, a mesma tendência geral de predominância de Ah e de Pn observada no gráfico P2. Portanto, as tendências são decrescentes para Ah e para Pn, como se pode ver pela inclinação negativa das retas de regressão linear apresentadas em P3c. Observe-se que esta classe está no limiar das categorias minoritárias, e que sua expressividade é baixa, não ultrapassando os 10% de frequência absoluta.

São exemplos dessa categoria os listados abaixo:

a graça do senhor deos por bem e sagesmente o bem fallar praticarmos [LealCons XV]
porq também agora a possamos vsar na nossa antiga & nobre lingua [FernOliveira XVI]
que nella recebe, a destruiã de todo, e matãraõ a ferro [PrincDJoão XVIII]

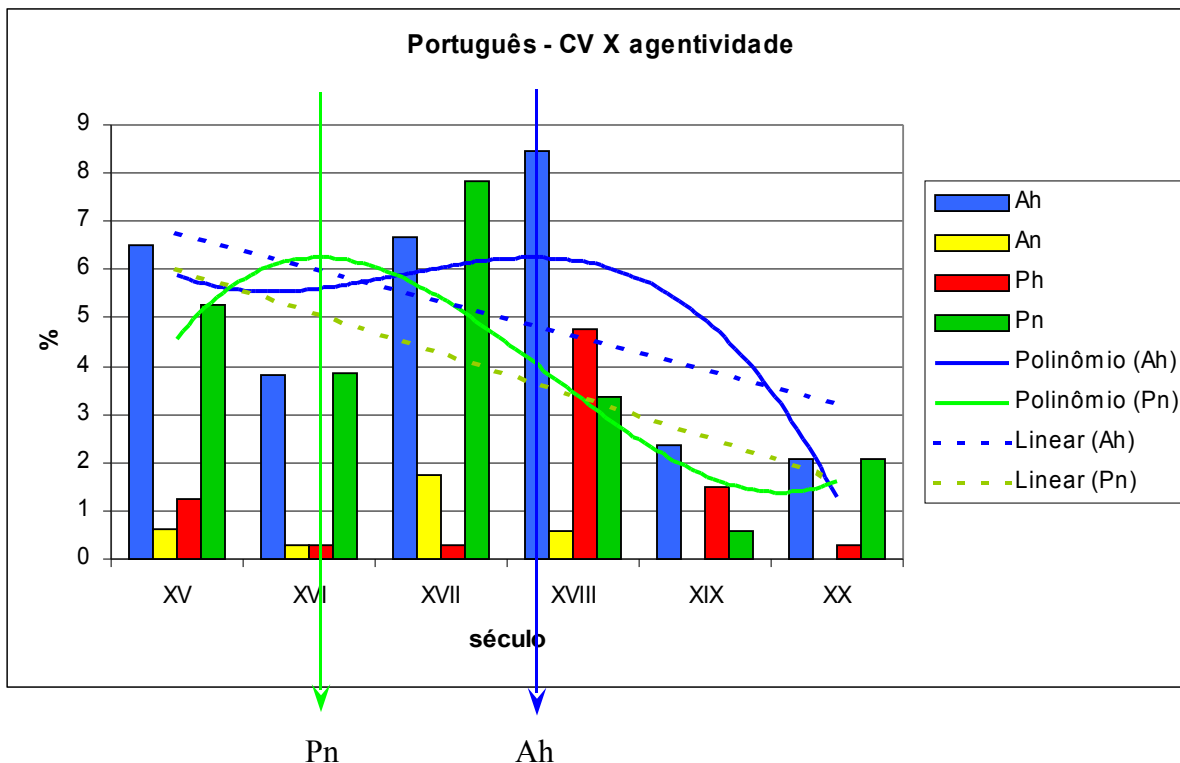


Gráfico P3c – distribuição das frequências relativas de CV em função da agentividade.

As categorias cujo complemento verbal não é expresso – apresentadas nos gráficos P3d e P3e abaixo, e correspondendo às ordens mais freqüentes, cf. Gráf. P1 acima – mostram tendências mais distintas em função da agentividade, sugerindo a relevância desse fator para a definição do comportamento diacrônico dos ambientes sintáticos.

A ordem *g* (SV) mostra um perfil mais simétrico para Pn do que para Ah, que tem, no século XIX, um máximo absoluto da ordem do dobro da média dos demais textos; em decorrência do fato de estar no final do intervalo de amostragem, esse ponto força a reta de regressão linear para cima, de maneira que ela tenha inclinação positiva – contrária, portanto, ao perfil global observado no *corpus*. Mesmo o conjunto Pn apresenta um perfil de inclinação bastante modesta, sugerindo tratar-se antes de uma tendência para essa fatia das amostras, e não de uma distorção do padrão de comportamento dos dados face à seleção de textos excêntricos.

São exemplos de ocorrências desse traço:

podemos dizer q e aquella das çinco çidades ja ditas a que elles chamarão Lisibona - FernOliv XVI

A agoa lhe vem por mar todos os dias em odres - PeJerLobo XVII

inesperadamente, Carlos apareceu em Lisboa com um architecto-decorador - Maias XIX

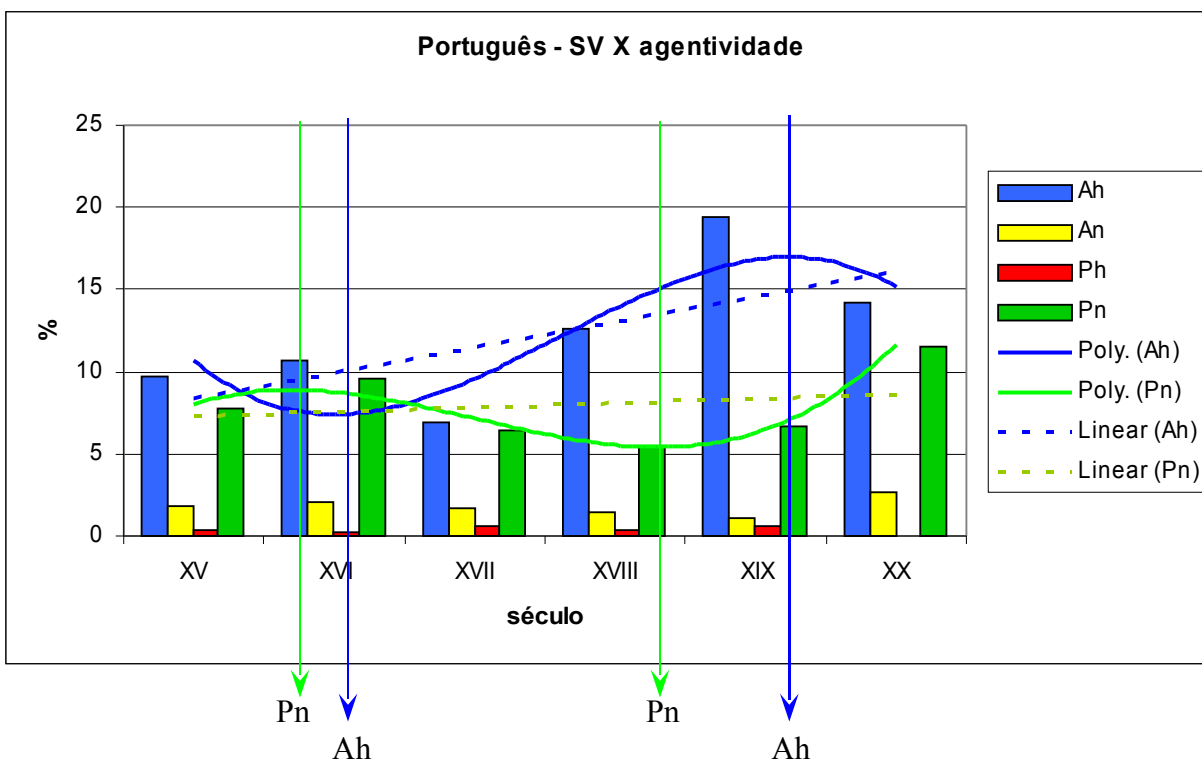


Gráfico P3d – distribuição das frequências relativas de SV em função da agentividade.

As ocorrências da ordem VS associadas a Ah e a Pn também têm uma distribuição decrescente ao longo do tempo, verificando-se um perfil igualmente distinto daqueles dos gráficos P3a a P3c: as séries Ah e Pn correm paralelas em quase todo o *corpus*, mas os valores iniciais são sensivelmente mais baixos que os intermediários e finais. Isso faz com que ambas apresentem tendências que sugerem a manutenção das médias absolutas de ocorrência, a despeito de sua flutuação moderada.

São exemplos desse caráter no *corpus* português:

quando entre duas vogaes vem hua consoante - DNLião XVI

Apareceu o sol sobre o horizonte - IlhaAmor XVIII

formidáveis parangonas, Nasceu a Nova Atlântida, No Xadrez Mundial - JangPedra XX

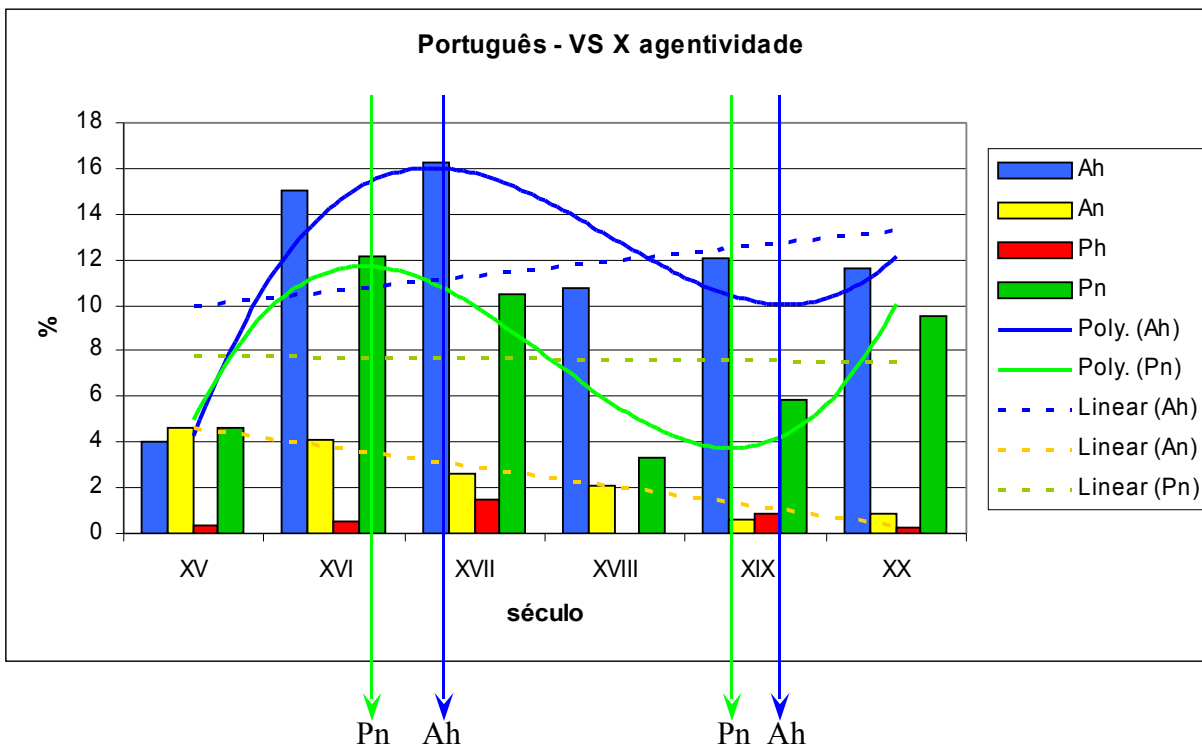


Gráfico P3e – distribuição das freqüências relativas de VS em função da agentividade.

A regressão linear das subclasses mais comuns no *corpus* mostra ainda que os dois blocos de ocorrências não apresentam a mesma taxa de diminuição observada para todo o conjunto (cf. gráfico P2 acima); aliás, a reta que resolve Ah é subhorizontal, indicando uma tendência distinta da global, que é crescente.

Em linhas mais gerais, a combinação dos traços *ordem de constituintes e agentividade* leva a tendências bem definidas, como mostra a tabela abaixo; apesar de haver flutuações nas freqüências absolutas de cada subclasse, pode-se identificar uma queda nas proporções em que as classes [VC] ocorrem (em consonância com o exposto nos Gráfs. P1 e P2), e aumento ou estabilidade das ordens que não têm o complemento verbal expresso. Para a ordem mais comum dentre as que possuem os dois argumentos verbais expressos, **a** (SVC), há um mínimo da curva polinomial Ah que antecede tanto a mudança de AN para NA (que teria ocorrido no século XVIII, cf. Cohen 1986/1990) quanto o mínimo da curva Pn; entretanto, a variação nas freqüências dessa classe **a** (SVC) no *corpus* português é demasiado restrita para permitir advogar uma relação clara entre os fatos internos ao SN e esses pontos de mínimo do Gráfico P3a (como proposto por Canale 1992).

ordem	Ah	Pn
a (SVC)	↘	↘
i (VC)	↘	↘
j (CV)	↘	↘
g (SV)	↗	→
h (VS)	↗	→

Tabela P1 – tendências gerais das classes dominantes de *agentividade* segundo o *ambiente oracional*.

Uma observação mais refinada é oferecida pelo cruzamento dos fatores *agentividade* e *ambiente oracional*, especificando-se subclasses mais propensas à alteração das frequências absolutas. Temos, então, o desdobramento dos subconjuntos Ah e Pn em três grupos principais, cujos comportamentos são distintos: o das orações subordinadas – o mais freqüente – apresenta um perfil decrescente, enquanto o de orações principais e o de coordenadas mostram perfis subhorizontais ou crescentes. As séries de gráficos P4 e P5 sumarizam todas as informações pertinentes à combinação desses dois critérios.

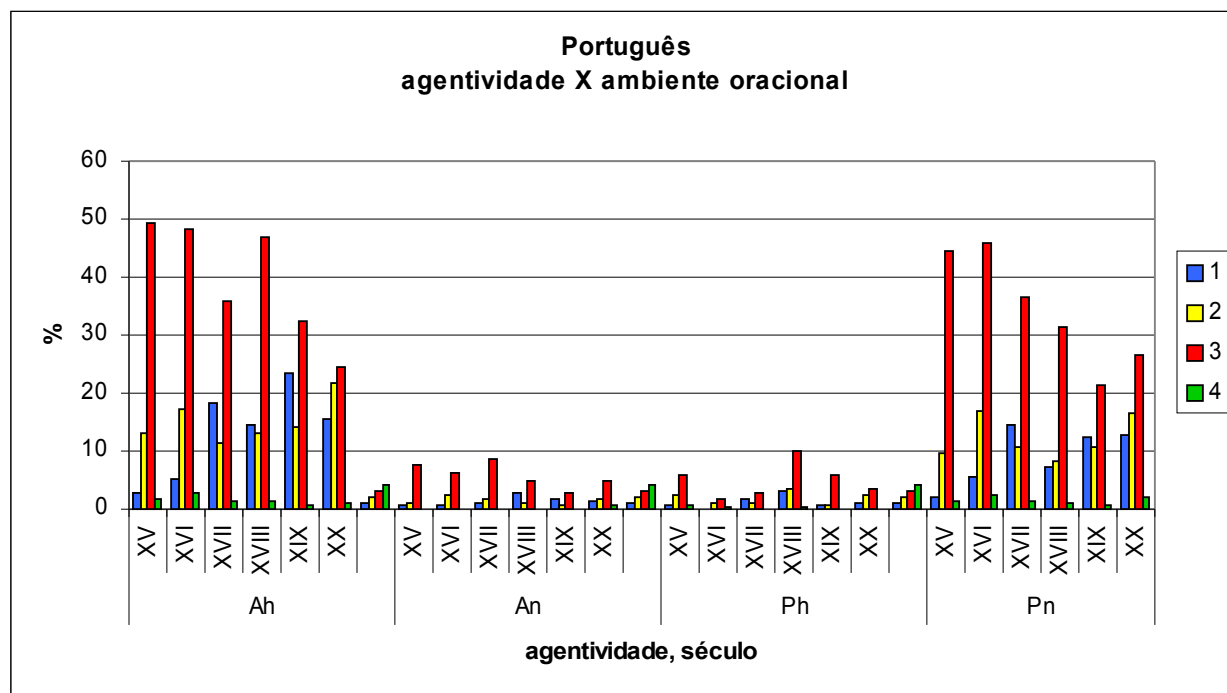


Gráfico P4a – distribuição geral das ocorrências para o cruzamento dos fatores *agentividade* e *ambiente oracional*.

Já foi observado em P4a que o perfil das orações subordinadas é diferente daquele relativo às demais. De fato, para a classe Ah, as subordinadas apresentam iRRL negativa (-4.59), enquanto as principais e as coordenadas têm inclinações positivas (iRRL(1) = 3.26; iRRL(2) = 1.01). É interessante observar que as médias de frequência de orações principais ultrapassam as de orações coordenadas no meio do intervalo de amostragem, mas estas voltam a ganhar espaço novamente no século XX. Dados os valores das iRRLs, as proporções iniciais e a extensão temporal da amostragem, as três curvas convergem no final do período considerado, indicando um estágio transitório de quase-equilíbrio entre os subconjuntos de ocorrências. Os exemplos apresentados abaixo foram colhidos no *corpus*:

- 1 - Apontey isto para que desta nossa propria & natural nobreza - FernOliv XVI
- 2 - e em presença dos condes ... firmou a doação que tenho - CondeDHenr XVII
- 3 - inquietação em que sem dúvida vos tem posto a incerteza - IlhaAmor XVIII
- 4 - e representando um grande ramo de girasoes atado por uma fita - Maias XIX

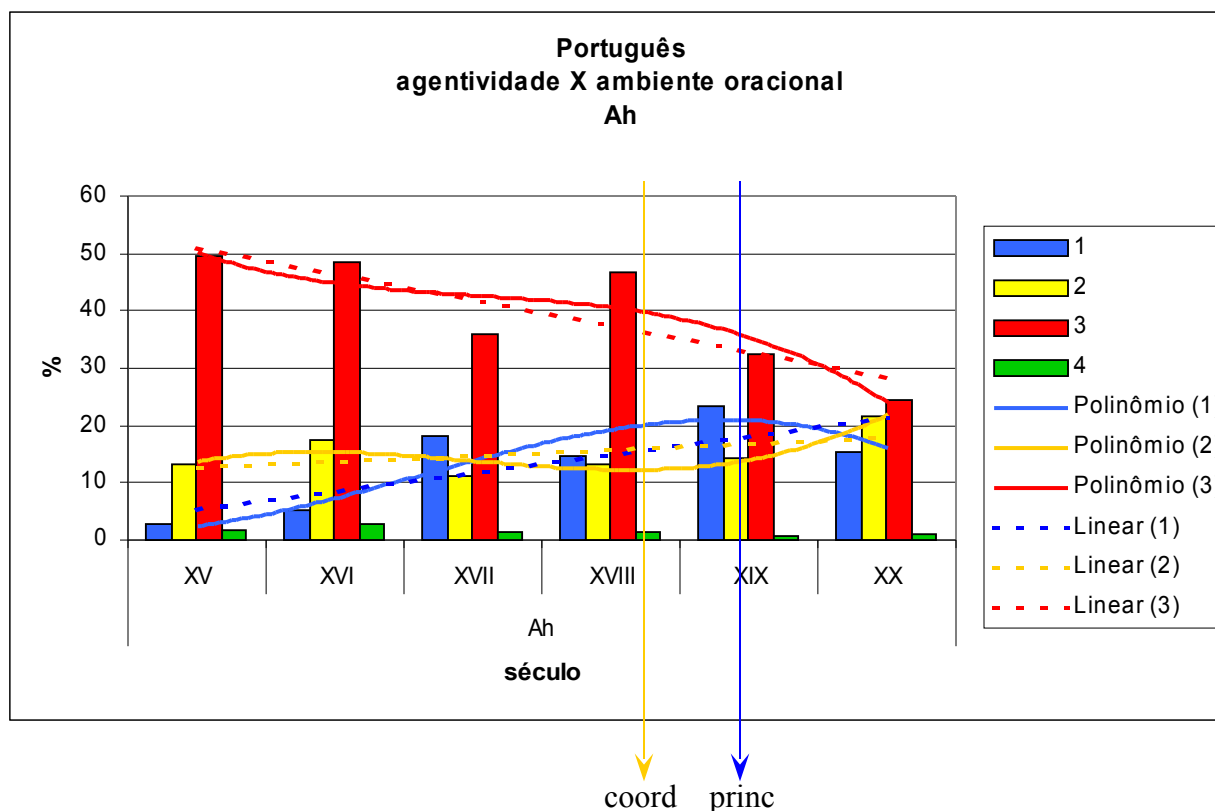


Gráfico P4b – distribuição das ocorrências para a classe Ah segundo o ambiente oracional.

Apesar de as conformações das curvas para Pn mudarem ligeiramente, mantêm-se as tendências já assinaladas no Gráfico P4b: há prevalência das orações subordinadas (no gráfico, correspondentes à classe 3) em todo o *corpus*, mas as médias de frequência absoluta dessas orações tendem a decrescer para o traço Pn. Em contrapartida, as médias de orações principais (no gráfico, correspondentes à classe 1) e coordenadas (no gráfico, correspondentes à classe 2) tendem a aumentar, fazendo, mais uma vez, com que a desproporção entre esses subconjuntos de ocorrências diminua. Nota-se ainda que o período da curva de coordenadas é menor que o de subordinadas – i.e., as mudanças se processam mais rapidamente naquelas do que nestas.

Tomem-se os exemplos seguintes:

- 1 - Esto digo segundo mynha tençom - LealCons XV
- 2 - pronunciamolo de maneira, que fica com a vogal precedente - DNLião XVI
- 3 - asima onde os turquos tem huma forsa bem fraqua de pedra - PeJerLobo XVII
- 4 - até que ElRey... mandou publicar, e apregoar guerra géral - PrincDJoão XVIII

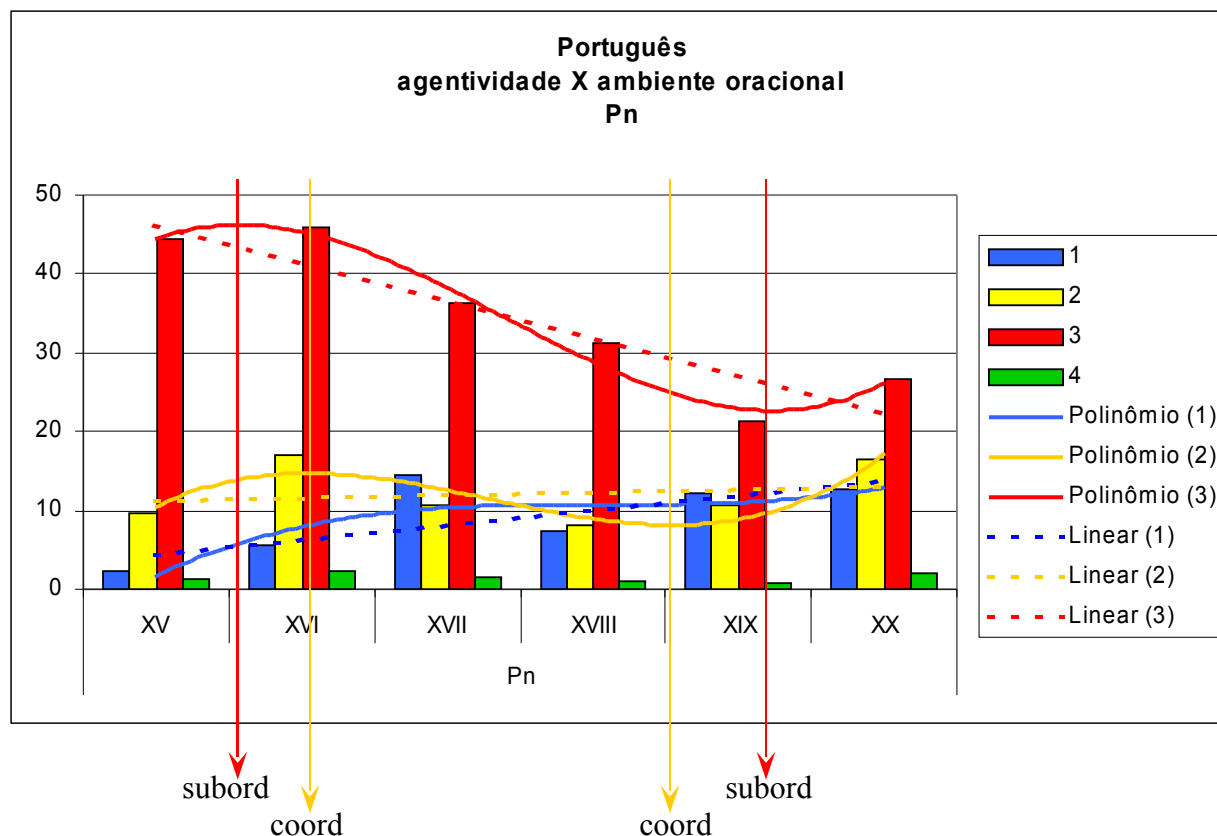


Gráfico P4c – distribuição das ocorrências de Pn segundo o ambiente oracional.

Em resumo, as tendências decrescentes para Ah e Pn (cf. Gráf. P2) são determinadas pelas iRRRLs das orações subordinadas, cuja perda em termos de frequência não é equilibrada pelo aumento das orações principais e coordenadas. Esse é um indicativo de que o ambiente oracional desempenha um papel importante na distribuição das subclasses de *agentividade* – o que, de fato, transparece na Tabela P2 abaixo: a análise dirigida pelo *ambiente oracional*, que corresponde à apreciação segundo as linhas da tabela, mostra um padrão – especialmente para orações principais (cujas frequências aumentam) e subordinadas (cujas proporções absolutas diminuem) – que será explorado nos Gráficos P5. Por outro lado, a análise dirigida pela *agentividade*, i.e., a leitura vertical da tabela, não mostra um padrão igualmente ordenado. Em qualquer dos casos, não há indícios claros de que a mudança no SN ocorrida no século XVIII (Cohen 1986/1990), de AN para NA, foi antecedida de uma alteração clara nas proporções das classes mais frequentes de *agentividade*; por vezes (a exemplo das orações principais em P4b e o de subordinadas em P4c), os pontos críticos das curvas são posteriores ao século XVIII, mas as relações causais e temporais entre esses fenômenos merecem estudos posteriores. Em outras palavras, os dados do *corpus* português indicam, até o momento, que a ordem dos constituintes A e N é independente, nesse intervalo, do caráter [\pm humano] de agente e de paciente. Isso não refuta, naturalmente, a observação de que fenômenos de ordens diversas possam se processar concomitantemente.

ambiente oracional	Ah	Pn
1 (or. principais)	↗	↗
2 (or. coordenadas)	↗	→
3 (or. subordinadas)	↘	↘

Tabela P2 – tendências de distribuição das ocorrências segundo os fatores *agentividade* e *ambiente oracional*.

A inversão dos fatores combinados nos Gráficos P4 acima mostra, portanto, o efeito do *ambiente oracional* na distribuição das classes de *agentividade*. No Gráfico P5a, tem-se uma visão geral dessa sobreposição de critérios de classificação das ocorrências, e percebe-se facilmente que as frequências absolutas das classes prototípicas Ah e Pn são praticamente equilibradas, século a século, em todo o intervalo em estudo; nesse ponto, não há alteração do padrão da língua em qualquer ambiente oracional, já que se continua atribuindo preferencialmente o traço +humano ao agente, e –humano ao paciente. Também transparece nesse recorte a propensão decrescente das classes associadas a orações subordinadas diante dos perfis crescentes de orações principais e coordenadas, ou seja, há uma tendência à simplificação das estruturas oracionais através da diminuição da hipotaxe e concomitante elevação das proporções de ocorrência de estruturas paratáticas.

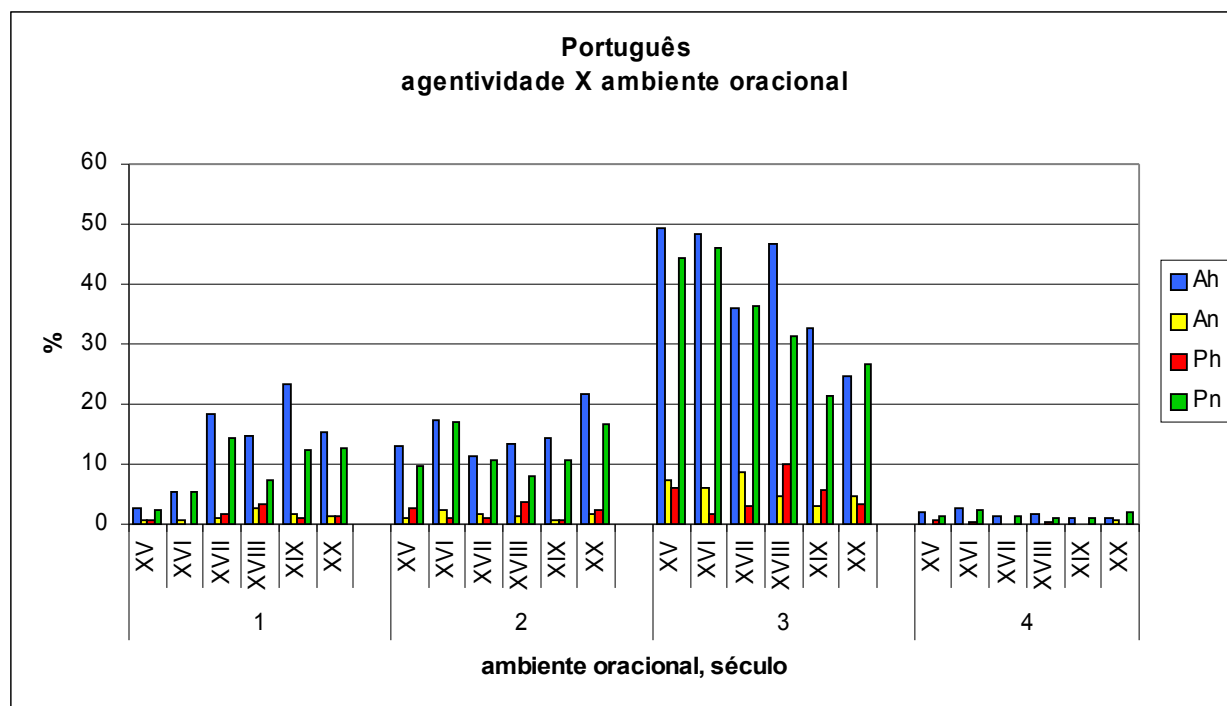


Gráfico P5a – panorama geral da distribuição das ocorrências segundo a combinação dos fatores *agentividade* e *ambiente oracional*.

Com respeito ao *ambiente oracional*, observemos, em primeiro lugar, como se comportam as classes de *agentividade* nas orações principais do *corpus* português. As duas categorias mais comuns, Ah e Pn, têm máximos absolutos de freqüência no século XIX; os dois subconjuntos de ocorrências têm iRRLs positivas ($iRRL(Ah) = 3,26$; $iRRL(Pn) = 1,89$), ressaltando-se que as freqüências de Ah aumentam mais rapidamente que as de Pn nesse ambiente oracional. Assinalei acima o fato de que as relações entre os dois fatores são mais nítidas em orações principais e subordinadas, e, com efeito, percebe-se com clareza o caráter ascendente das curvas Ah e Pn – que têm, inclusive, máximos locais indubitavelmente lançados no século XIX (posteriores, portanto, à mudança de AN para NA identificada por Cohen 1986/1990). As ocorrências apresentadas abaixo são exemplos desse traço:

- Ah - com Castella. eu quero ordenar huuas festas rreaes que duresm todo - Ceuta XV
 An - O Estado da fortuna pode cõceder ou tirar fauor aos estudos - FernOliv - XVI
 Ph - (pois não consta que tivesse aquêle fidalgo segunda mulher) - CondeDHenr XVII
 Pn - Sente-se alli o murmurar das aguas e a bafagem suave do vento - CastFaria XIX

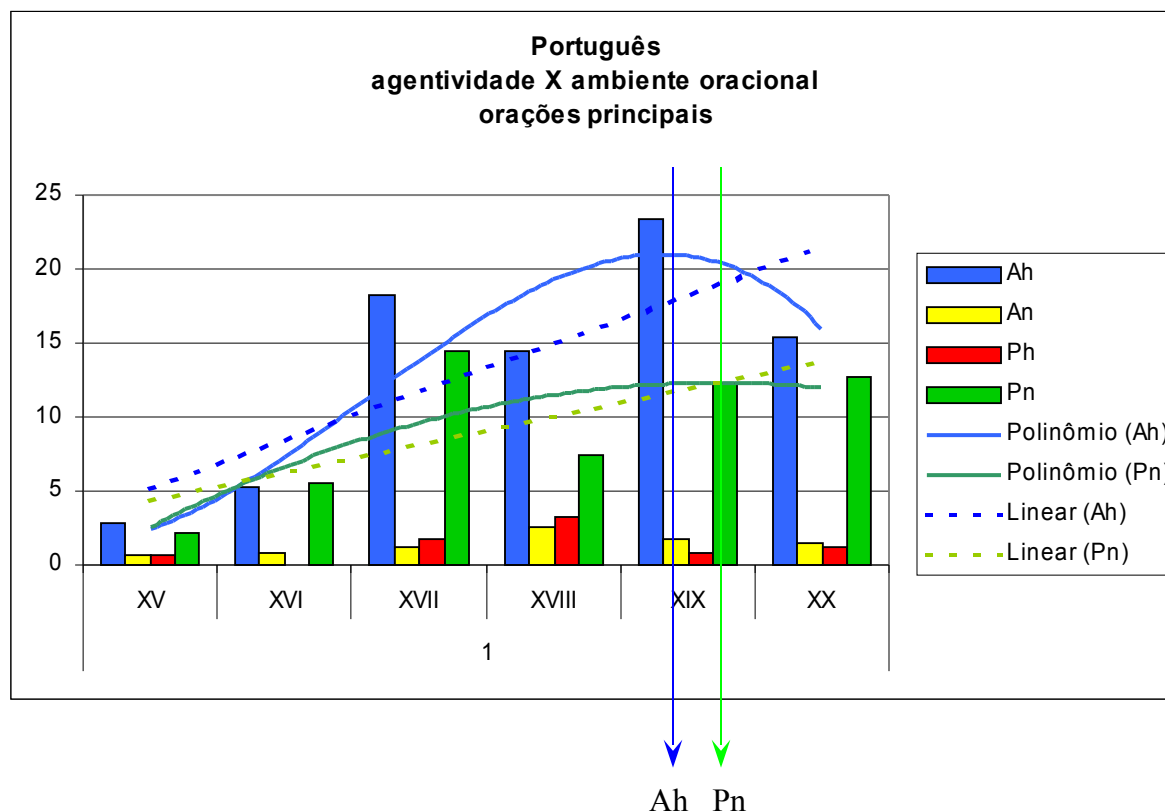


Gráfico P5b – distribuição das classes de agentividade para orações principais.

Para as orações coordenadas, obtém-se um perfil diferente do anterior, apesar de as iRRLs também indicarem aumento discreto das médias gerais das classes mais frequentes: iRRL(Ah = 1,01; iRRL(Pn) = 0,37). As curvas para esses subconjuntos de ocorrências mostram máximos locais praticamente coincidentes no século XVI, mas os mínimos locais apresentam uma defasagem mais nítida.

- Ah - E nos pelo contrario, nas dictas dições sentimos hu som muito - DNLião XVI
- An - mas de contino lhes tinha banhado as raizes a agoa salgada até meio - PeJerLobo XVII
- Ph - *Porém, alborotados os sentidos, a não podem ver nem ouvir* - IlhaAmor XVIII
- Pn - e as povoações e os rios ... apresentam-lhe o panorama variadissimo - Castfaria XIX

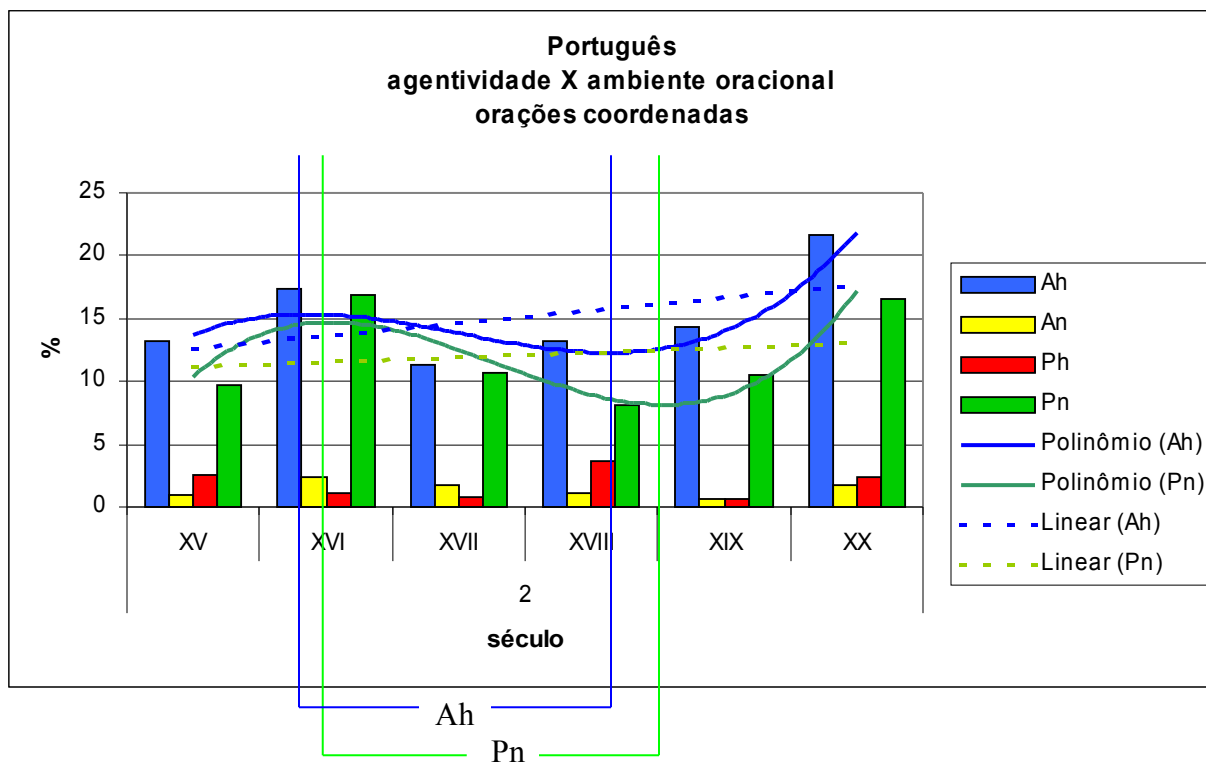


Gráfico P5c – distribuição das classes de agentividade para orações coordenadas.

Já foi mencionado acima (q.v. Gráficos P4a e P5a) que as proporções de orações subordinadas diminuem no sentido positivo do eixo das abscissas; entretanto, é interessante notar que, dentre essas orações, os perfis de Ah e de Pn são diferentes: a de Ah é sempre decrescente – apesar de as taxas de decréscimo variarem no intervalo considerado –, e Pn tem um perfil senóide cujo mínimo absoluto ocorre no mesmo momento em que as classes Ah e Pn têm orações principais (cf. Gráf. P5a) atingem seus valores máximos: o século XIX. As iRRLs têm praticamente o mesmo valor, e são, como observado anteriormente, negativas: $iRRL(Ah) = -4.59$; $iRRL(Pn) = -4,79$.

- Ah - Como elRey tinha vontade de fazer grandes festas em Lixboa pera fazer - Ceuta XV
- An - porque nesta paragem desem das serras ribeiras que se escondem - PeJerLobo XVII
- Ph - que ElRey ... quizera mandar aquella Armada toda contra - PrincDJoão XVIII
- Pn - fotografia que mostrava a península, se definitivamente não - JangPedra XX

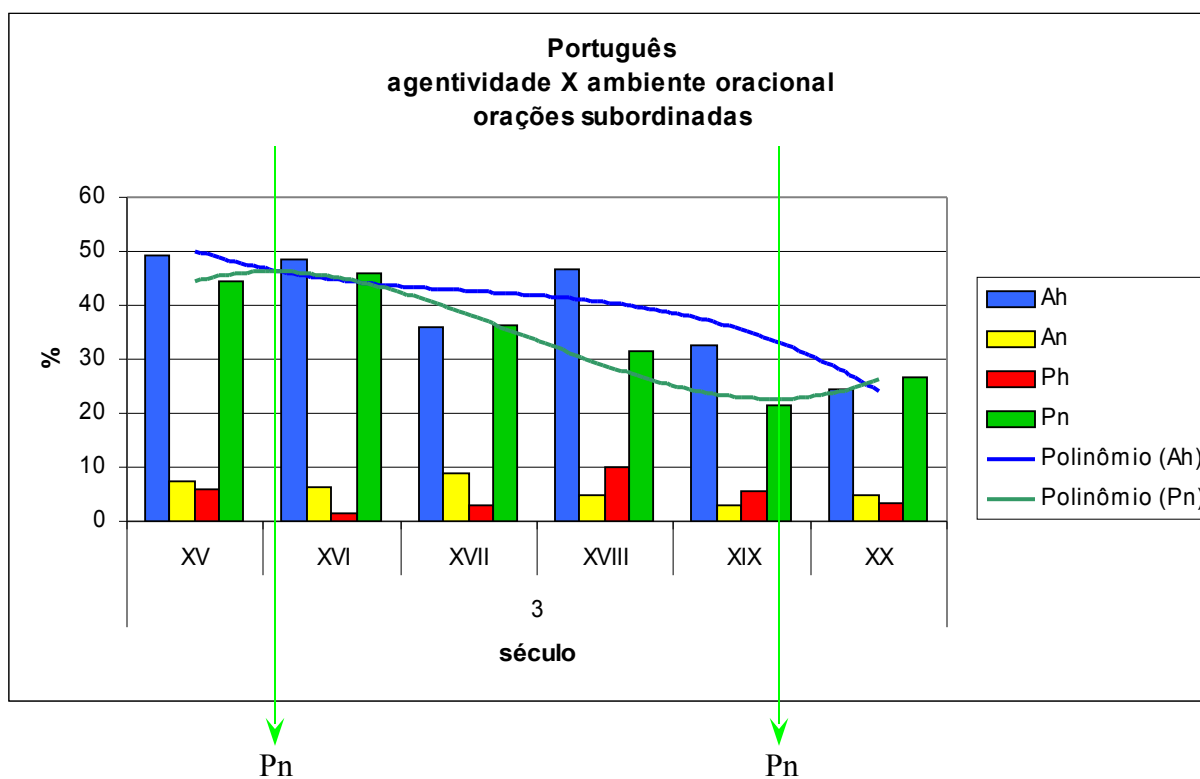


Gráfico P5d – distribuição das classes de agentividade para orações subordinadas.

O fator *ambiente oracional* aponta uma mudança na configuração geral da língua portuguesa, no sentido de um equilíbrio entre as proporções de orações principais (no gráfico, correspondentes à classe 1), coordenadas (no gráfico, correspondentes à classe 2) e subordinadas (no gráfico, correspondentes à classe 3). Em termos mais abrangentes, verificam-se, no *corpus* português, a) aumento nas médias de frequência absoluta de orações principais (iRRL = 5.35), b) flutuação nas médias de frequência absoluta de orações coordenadas, com leve aumento dessas proporções (iRRL = 1.67), e c) diminuição generalizada das médias de frequência absoluta de orações subordinadas (iRRL = -6.83) ao longo de todo o intervalo de amostragem. Essa mesma tendência geral de decréscimo da frequência de estruturas hipotáticas é observada para os outros *corpora* tratados neste trabalho.

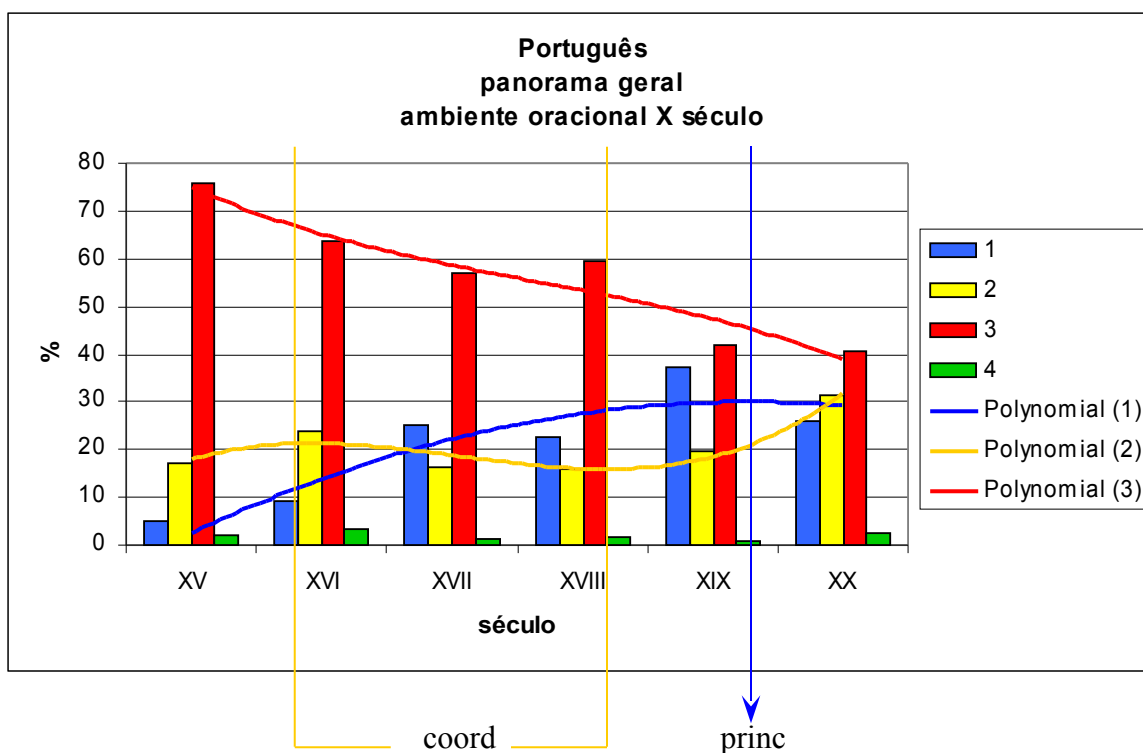


Gráfico P6 – distribuição geral das ocorrências segundo o fator *ambiente oracional*.

O cruzamento dos fatores *ambiente oracional* e *ordem* leva a perfis como os apresentados nos gráficos P6a a P6c, nos quais é possível identificar uma tendência à especialização de algumas ordens em certos ambientes. A partir do exame de P6a, por exemplo, percebe-se que, em orações principais, a ordem [VC] é cada vez mais freqüente (iRRL = 1.78, assinalada no gráfico) – e, dentro desse grupo, sobressai a subclasse SVC (1.76); as combinações sintáticas em que o complemento do verbo não é expresso, **g** (SV) e **h** (VS), também ganham espaço em orações principais (iRRL(**g**) = 1.95; iRRL(**h**) = 1.05). Algumas ordens que ocorrem em orações principais são exemplificadas abaixo:

- a (SVC) - O coração pelos v sentidos filha principalmente amor - LealCons XV
- c (VSC) - Conta diodoro siculo no seisto liuro da sua bliblioteca q foy - FernOliv XVI
- d (VCS) - com sua espada e poder real, favorecendo-o Deus, tomou - CondeDHenr XVII
- i (VC) - *Ides buscar, aventureiros, o verdadeiro gosto que Amor inspira* - IlhaAmor XVIII

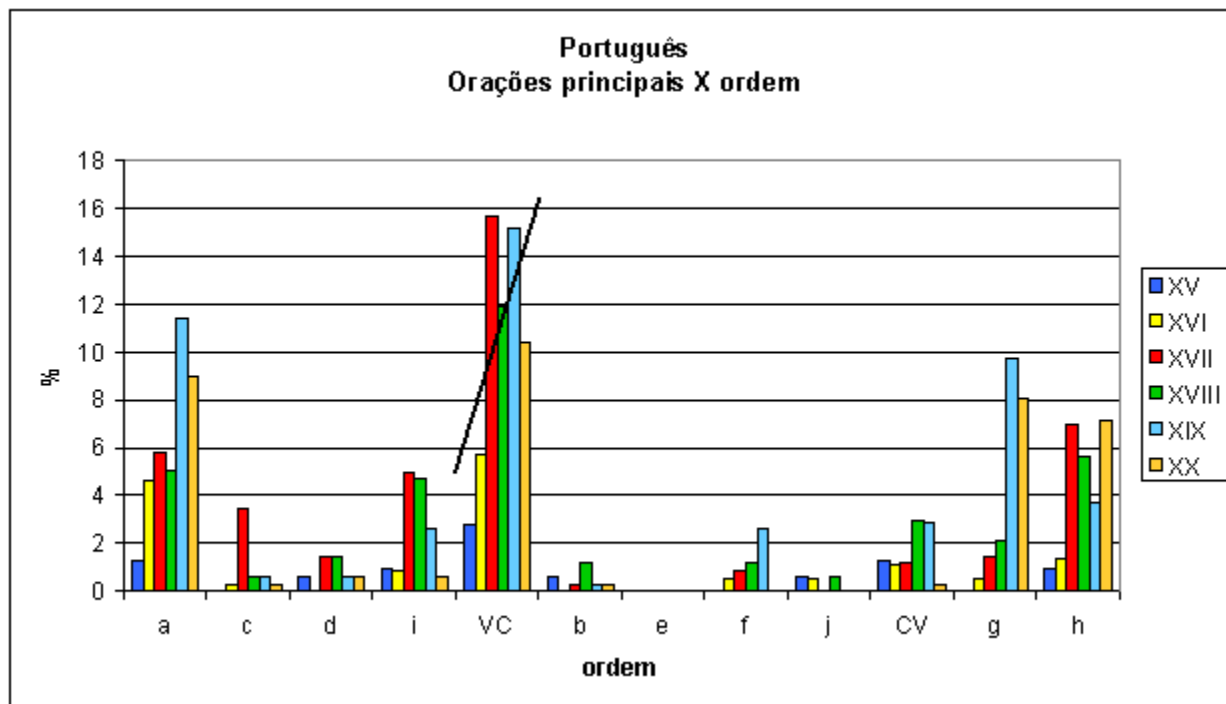


Gráfico P6a – distribuição das freqüências relativas da ordem de constituintes em orações principais.

Em P6b, temos o mesmo aspecto geral apresentado no gráfico anterior: a classe geral [VC] predomina e tem inclinação levemente positiva ($iRRL = 0.14$), o que também se passa com respeito às classes **g** e **h** ($iRRL(SV) = 1.61$; $iRRL(VS) = .22$); a escala de ordenadas tem ainda, para os dois gráficos, praticamente a mesma extensão, indicando que os valores máximos alcançados em ambos os gráficos se aproximam muito um do outro. Entretanto, a subclasse mais expressiva no grupo [VC] em orações coordenadas é **i** (VC), e não SVC, como acontece em principais. Esse traço é ilustrado graças aos exemplos apresentados abaixo:

- a (SVC) - e a mal ordenada sparge e cria muytos desacordos e pellejas - LealCons XV
- j (CV) - Êle... povoou esta cidade e a sustentou com grande valor - CondeDHenr XVII
- g (SV) - mas elle desejava habitar sob tectos tradicionalmente seus - Maias XIX
- h (VS) - e não tivessem sido ainda decididas as estações e os tempos - JangPedra XX

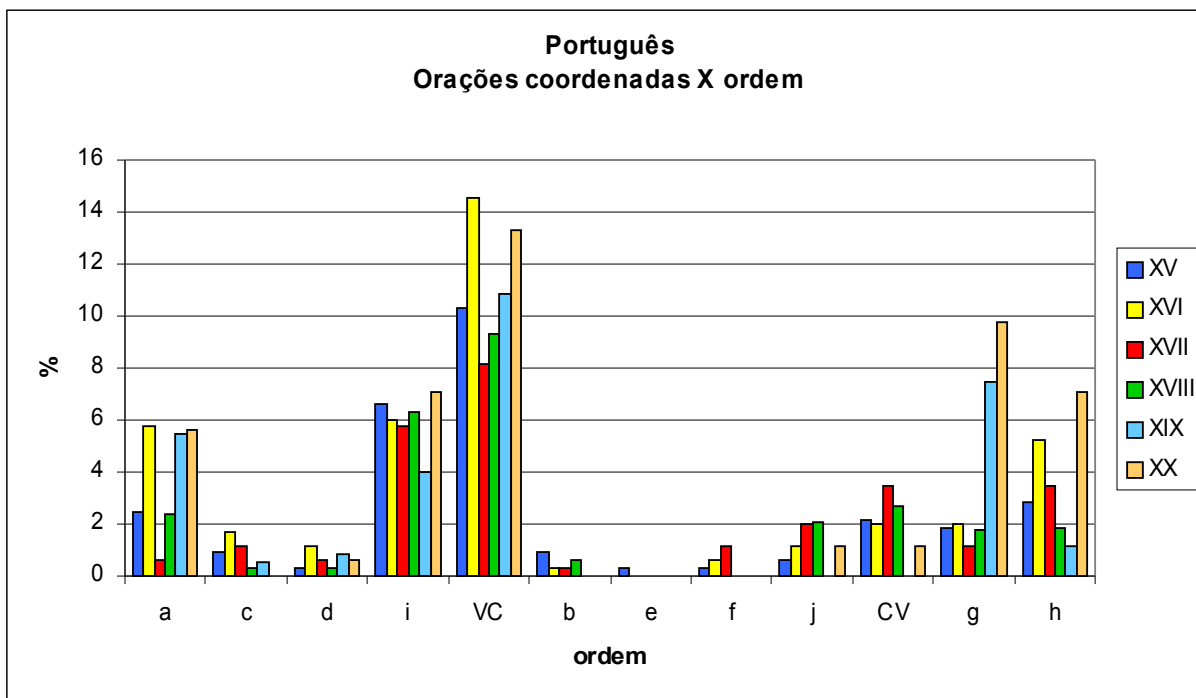


Gráfico P6b – distribuição das frequências relativas da ordem de constituintes em orações coordenadas.

Configuração distinta é apresentada em P6c, segundo o qual a preferência por estruturas subordinadas tem diminuído sistemática e sensivelmente ao longo do intervalo de amostragem. Nota-se, em primeiro lugar, que as construções hipotáticas são mais freqüentes que principais e paratáticas, mas os perfis são inversos: enquanto as orações subordinadas ocorrem em proporções cada vez mais baixas no *corpus*, as principais e as coordenadas aumentam em freqüência relativa, como visto em P6a e em P6b acima. O decréscimo da preferência por subordinadas é observado em todas as ordens de constituintes aqui estudadas. Por fim, observe-se que a ordem **i** (VC) é, a exemplo do que foi visto para as coordenadas, a mais freqüente no grupo [VC], e tem comportamento alinhado ao desse grupo. Apresentamos os exemplos:

- a (SVC) - antre sy que semelhante maneira de cauallaria nam era honrrosa - Ceuta XV
i (VC) - lingoa Portuguesa, que sempre põe .m. no fim das dições - DNLião XVI
g (SV) - que eu tinha visto e que com desejo ardente - IlhaAmor XVIII
h (VS) - Era esta lagea a mesa em que costumava comer Salat-ibn-Salat - CastFaria XIX

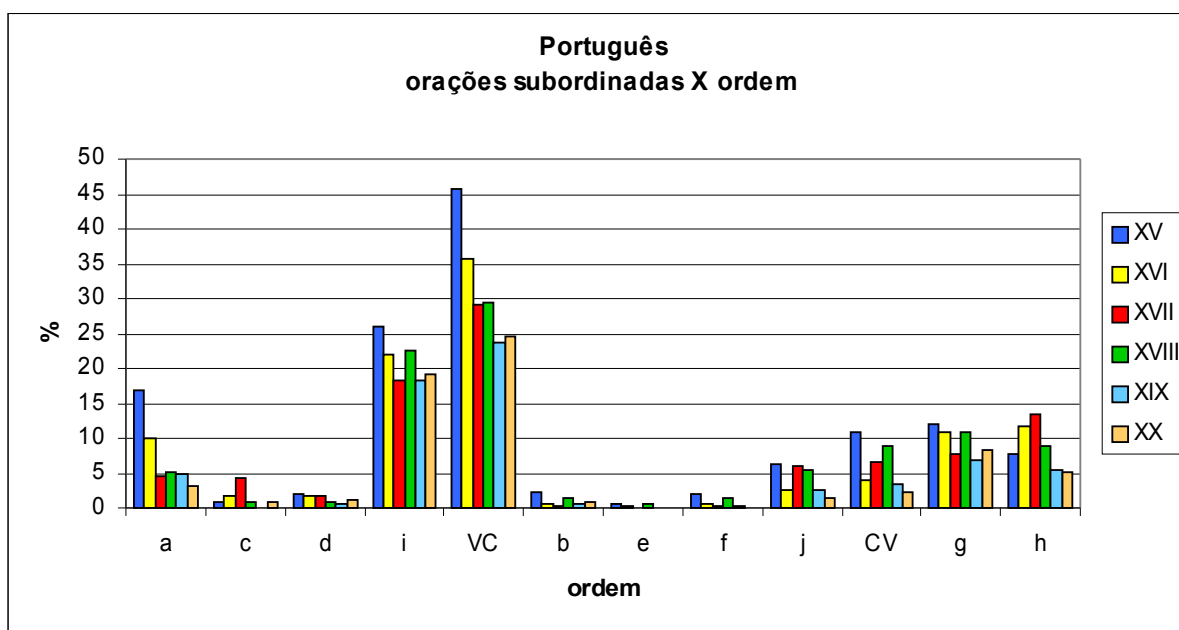


Gráfico P6c – distribuição das freqüências relativas da ordem de constituintes em orações subordinadas.

Para as orações coordenadas em subordinação, as freqüências são demasiado reduzidas (pico de **i** (VC) = 2.6%) para permitir conclusões sobre o comportamento do sistema; entretanto, pode-se perceber que, para os agrupamentos mais comuns, como o conjunto de ordens [VC], esboça-se uma tendência ao decréscimo nas proporções absolutas. Isso leva à hipótese de que

essas orações teriam comportamento idêntico ao de coordenadas, apesar de estarem em subordinação; a verificação da validade dessa hipótese demanda, contudo, um número muito maior de ocorrências em virtude da baixa frequência dessa classe.

As tendências identificadas nos Gráficos P6 acima podem ser resumidas na Tabela P3 abaixo; nota-se claramente a definição de um padrão de comportamento segundo o *ambiente oracional*, como já apontado acima. As classes mais recorrentes mostram a mesma propensão assinalada para os traços mais comuns de *agentividade* (cf. Gráf. P2 acima): aumento generalizado em orações principais e queda na frequência de orações subordinadas.

<i>ordem</i>	<i>1 (or. princ.)</i>	<i>2 (or. coord.)</i>	<i>3 (or. subord.)</i>
<i>a (SVC)</i>	↗	↗	↘
<i>i (VC)</i>	→	→	↘
<i>[VC]</i>	↗	→	↘
<i>g (SV)</i>	↗	↗	↘
<i>h (VS)</i>	↗	→	↘

Tabela P3 – tendências gerais das classes de *ordem de constituintes* segundo o *ambiente oracional*.

Os gráficos P7 correspondem a todas as possíveis associações entre a natureza de S e de C e a ordem de constituintes, segundo as quais também se identificam perfis distintos.

Em P7a, o traço utilizado para estudar a ordem dos elementos na oração foi **k**, sujeito pleno; os padrões de ordenamento de constituintes mais favoráveis à ocorrência desse tipo de sujeito são **a** (SVC), **g** (SV) e **h** (VS), que dominam definitivamente todo o intervalo coberto pelo *corpus*. No entanto, nota-se que essas ordens apresentam tendências semelhantes ao longo do tempo: a frequência de **a** (SVC), por exemplo, eleva-se levemente (iRRL = 1.13), ao passo que a regressão linear de **g** (SV) e de **h** (VS) fornece valores mais altos, mas também crescentes (iRRL(**g**) = 3.88; iRRL(**h**) = 2.02).

Tomem-se os casos oferecidos abaixo como exemplos desse traço:

- a (SVC) - a outros alguus parecese grande a elles pareçeo muy pequeno - Ceuta XV
- c (VSC) - e sujeitando outros os templos e mosteiros célebres - CondeDHenr XVII
- d (VCS) - rumor de passos morria: e ornavam-n'a divans cobertos de tapetes - Maias XIX
- g (VC) - quando a península, sem dar aviso, se moveu outra vez - JangPedra XX

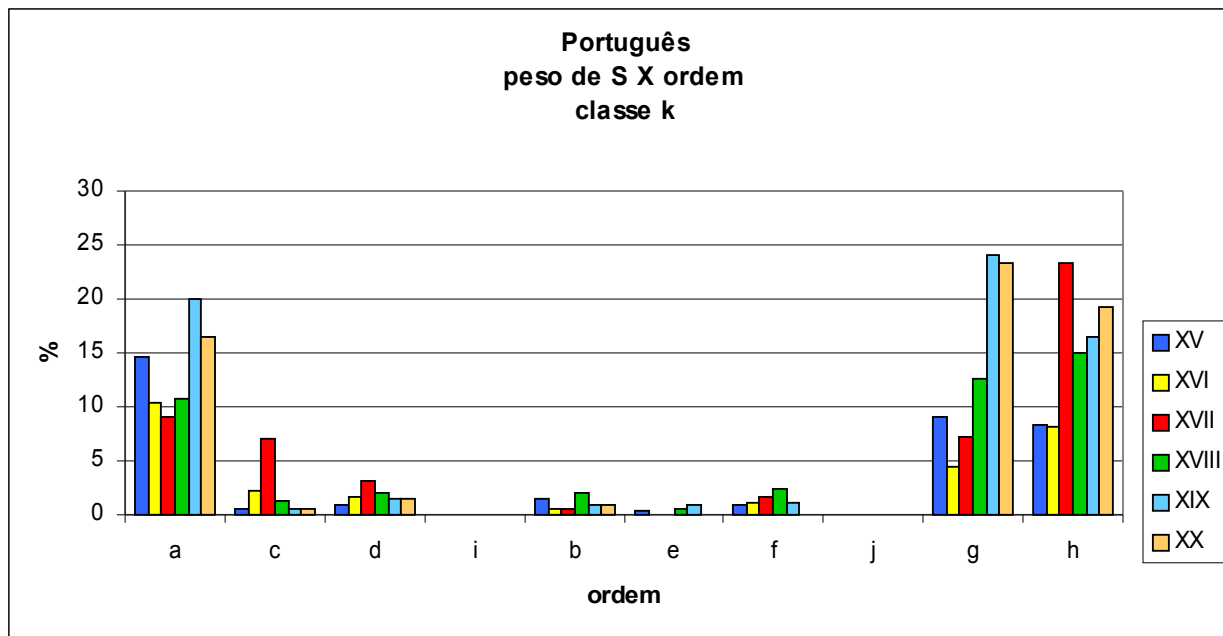


Gráfico P7a – distribuição das frequências relativas da ordem de constituintes diante de S pleno.

No caso de S pronominal (Gráfico P7b), as frequências são muito mais baixas, mas é importante observar que os perfis se invertem em relação ao gráfico anterior: tanto para **a** (SVC) quanto para **g** (SV), temos inclinações negativas para a reta que define a tendência geral dos blocos de dados (respectivamente, -0.56 e -0.53). Em outras palavras, o sujeito pronominal tem sofrido uma diminuição de uso generalizada, o que se reflete nas categorias principais **a** (SVC) e **(g)** SV. Como a escala de ordenadas tem um alcance menor do que a do gráfico anterior, a expressividade da mudança aqui considerada é, conseqüentemente, reduzida.

Os exemplos apresentados abaixo ilustram esse caráter:

- a (SVC) - porende eu das boas screvo esta maneira que com ellas - LealCons XV
c (VSC) – dizem, villano, villanos, diremos nos, villãos, cidadãos, aldeãos - DNLIão XVI
g (SV) - mas tambem como alguns querem ha outro muito arvoredado - PeJerLobo XVII
h (VS) - Saíra este com alguns soldados para soccorrer o conde de Cêa - Castfaria XIX

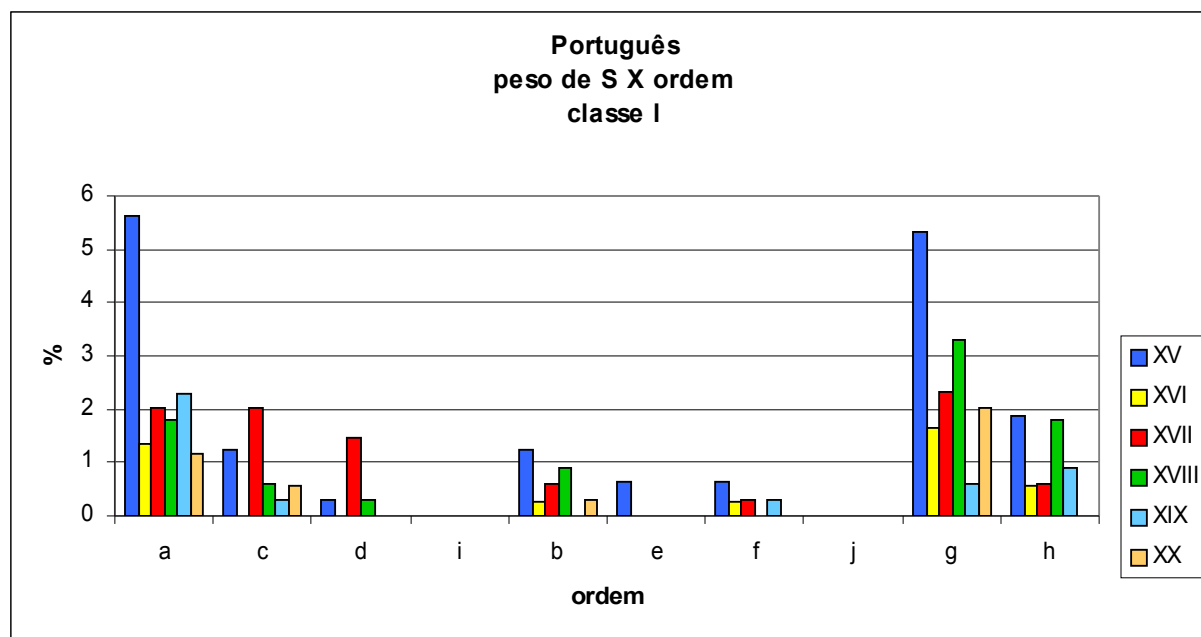


Gráfico P7b – distribuição das freqüências relativas da ordem de constituintes diante de S pronominal.

O sujeito oracional ocorre em proporções extremamente baixas, cujos máximos absolutos não passam dos 2%. O tratamento das ocorrências permite afirmar que essa classe de sujeito ocorre principalmente nas ordens **d** (VCS), **g** (SV) e **h** (VS) – todas decrescentes, o que indicaria um comportamento semelhante ao de S pronominal, ou mesmo a possibilidade de se considerarem dois grandes conjuntos de S: de referência direta, como o S pleno (**k**), e de referência indireta, como o S pronominal (**l**) e o S oracional (**m**).

Para o complemento verbal pleno, ao contrário do que se passa com S pleno, verificam-se perfis cujas retas de regressão têm inclinação negativa, apontando uma diminuição de moderada (para a ordem **a** (SVC): $iRRL = -.95$) a discreta (para a ordem **i** (VC): $iRRL = -.45$) no que respeita às classes mais frequentes. Observe-se a predominância absoluta dos padrões de ordenamento em que o verbo é seguido de seu complemento, diante dos valores muito baixos de ordens de tipo CV (**b** (SCV), **e** (CSV), **f** (CVS) e **j** (CV)).

No *corpus* português, colheram-se as ocorrências:

- a (SVC) - em que os escollares apuram todallas partes das outras çienças - Ceuta XV
 c (VSC) - destas faz Plinio mença aos vinte capitulos do quarto liuro - FernOliv XVI
 d (VCS) - He contudo necessario ajuntar aqui hum parafago da rezão - PeJerLobo XVII
 i (VC) - com idéa d'installar lá a Nunciatura, seduzido pela gravidade clerical - Maias XIX

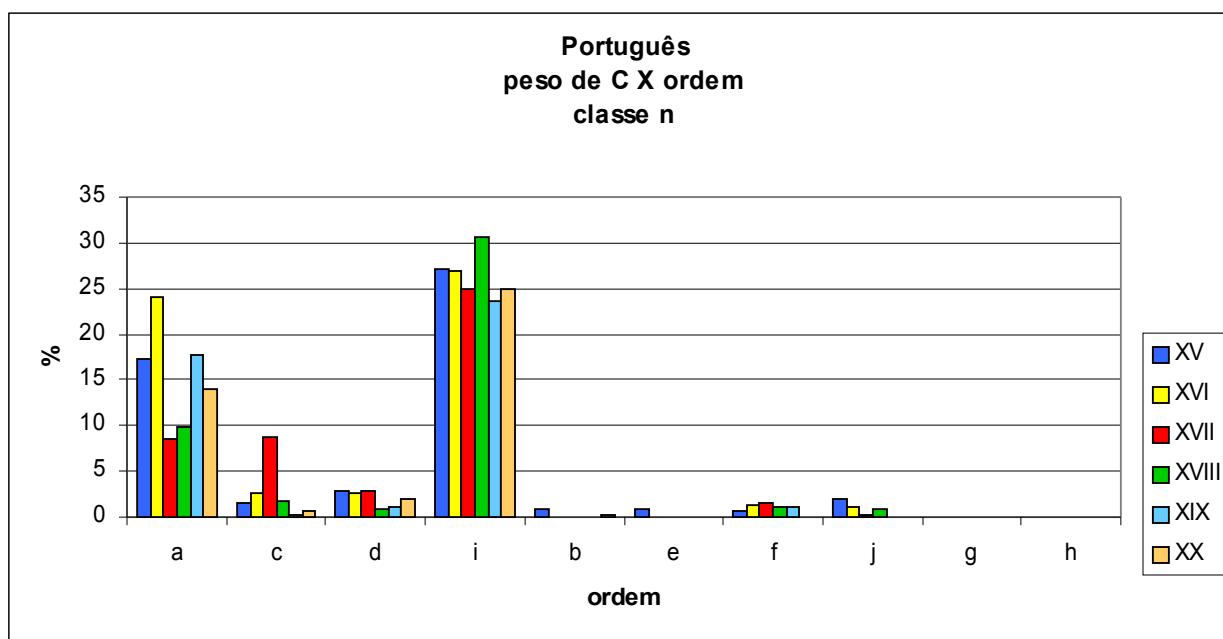


Gráfico P7c – distribuição das frequências relativas da ordem de constituintes diante de C pleno.

A frequência com que o objeto pronominal (classe **o**) ocorre também diminui proporcionalmente para diversos padrões de ordenamento, como se pode observar no Gráfico P7d: as classes mais comuns – **j** (CV), **i** (VC) e **b** (SCV) – sofrem decréscimo leve do século XV ao XX; a que mostra perda mais sensível é precisamente **j** (CV). Deve-se observar que, apesar de as frequências das categorias pronominais serem muito mais modestas que as de sintagmas nominais plenos, os subgrupos **l** (sujeito pronominal, Gráf. P7b) e **o** (complemento pronominal) apresentam tendências decrescentes na amostra, ao contrário de **k** (sujeito pleno, Gráf. P7a) e **n** (complemento pleno, Gráf. P7c). Passemos aos exemplos e ao gráfico de distribuição:

b (SCV) - por que ella o faz pryncypalmente per sua virtude e bondade - LealCons XV
e (CSV) - que fiquem scriptas, como as nos pronuciamos - DNLião XVI
f (CVS) - Assim o conta a *História dos Gôdos*, com estas palavras - CondeDHenr XVII
j (VC) - Logo que a encontrei, vi com ela um homem que me atemorizava - IlhaAmor XVIII

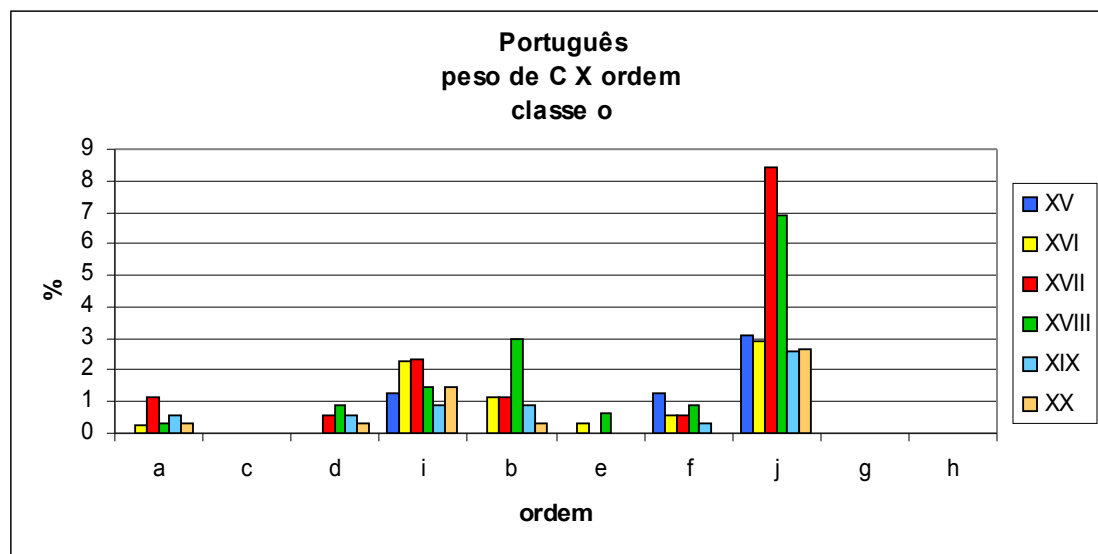


Gráfico P7d – distribuição das frequências relativas da ordem de constituintes diante de C pronominal.

A última das categorias de peso, complemento oracional (classe **p**), mostra comportamentos diferentes para duas de suas ordens principais, VC (**i**) e SVC (**a**). Uma vez mais, diante da reduzida escala dentro da qual os dados se inserem, é preciso cautela para fazer afirmações sobre o comportamento dessa classe: veja-se que, se se considerar como excêntrico o ponto relativo à média de valores dos textos do século XV, os grupos **a** e **i** apresentarão retas de inclinação positiva; por outro lado, se se exclui o ponto relativo ao século XX (como sugerido acima para normalizar o comportamento da classe **m**), as retas terão inclinação negativa. No primeiro caso, o complemento oracional posposto ao verbo seria cada vez mais freqüentemente encontrado nos textos; no segundo, poder-se-ia falar novamente numa superclasse englobando os objetos pronominais e oracionais.

Exemplos da classe **p** são:

a (SVC) - Porem a rrazom bem demostra que, onde os ha, nom he aquella - LealCons XV
c (VSC) - ainda porẽ q diz mersilo q de Hetruria tem a Ytalia as letras - FernOliv XVI
i (VC) – meu pae : dizeí-lhe que eu o espero - CastFaria XIX

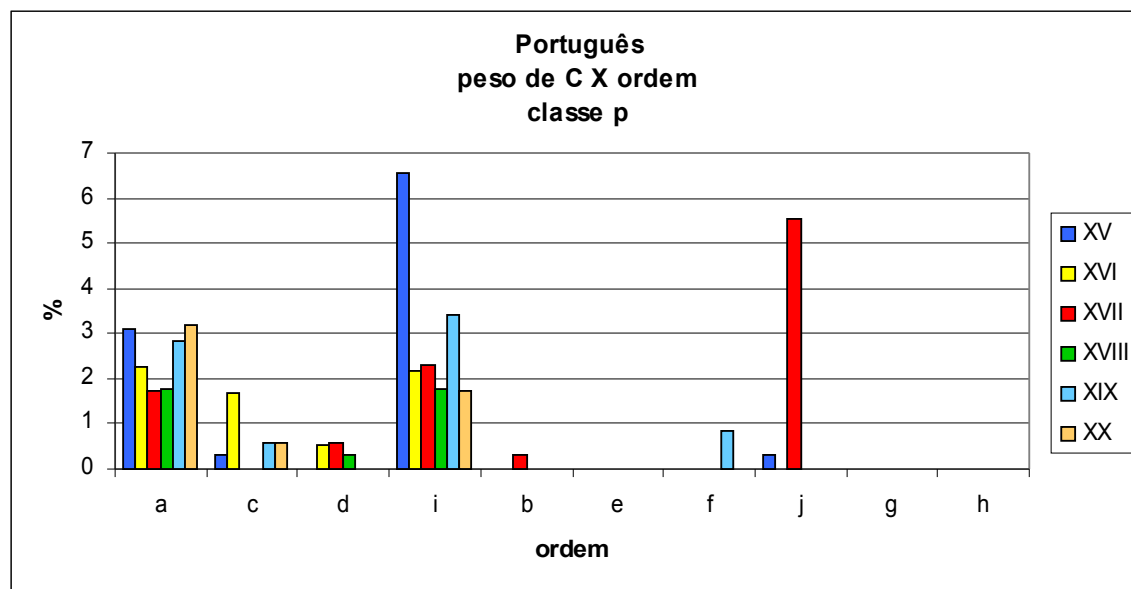


Gráfico P7e – distribuição das freqüências absolutas da ordem de constituintes diante de C oracional.

Em resumo, o único bloco de dados cujo comportamento parece contrapor-se a todos os demais é **k** (sujeito pleno), cujas ordens mais freqüentes apresentam retas de regressão de inclinação positiva; isso quer dizer que essa classe tende a aumentar em freqüência absoluta

diante das demais classes do mesmo fator, e nas ordens **a** (SVC), que tem os dois argumentos verbais expressos, e **g** (SV) / **h** (VS), que não trazem o complemento verbal. Os comportamentos dos subconjuntos mais freqüentes de ocorrências segundo o fator *peso de S e de C* estão representados na Tabela P4 abaixo, em que se destacam as tendências crescentes para as ordens associadas à classe **k** (sujeito pleno) e decrescentes para S pronominal (classe **l**) e C pleno (classe **n**). Nota-se, na Tabela, que o comportamento das classes é melhor definido pelo fator *peso de S e de C*, o que transparece no surgimento de padrões segundo uma leitura vertical da mesma:

	<i>k (S pleno)</i>	<i>l (S pronominal)</i>	<i>n (C pleno)</i>
<i>a (SVC)</i>	↗	↘	↘
<i>i (VC)</i>			↘
<i>g (SV)</i>	↗	↘	
<i>h (VS)</i>	↗	↘	

Tabela P4 – tendências gerais de comportamento das classes de *ordem de constituintes* segundo o *peso de S e de C*.

A sobreposição dos fatores *peso de S e ambiente oracional*, ilustrada nos gráficos P8, mostra, inicialmente, como as ocorrências de sujeito pleno (classe **k**) se distribuem no *corpus*

português. Surgem, também aqui, as curvas apresentadas no Gráfico P6, que apontam o crescimento franco das proporções de orações principais, o aumento mais suave das freqüências de coordenadas e a diminuição das proporções de subordinadas. No caso da classe **k**, observa-se que as médias de freqüência absoluta de orações principais ultrapassam as de subordinadas no século XVIII (o mesmo século em que há mudança da orientação AN para NA no português, cf. Cohen 1986/1990); ainda assim, as três curvas tendem a uma distribuição mais equilibrada ao final do intervalo de amostragem.

Ocorrências como as apresentadas abaixo levam à construção do gráfico P8a:

- 1 – Antre os boos amygos ... estas cousas muy necessariamente se requerem - LealCons XV
- 2 – E estes diphthongos se formão em cada lingua de diferentes maneiras - DNLião XVI
- 3 – escandalizados os portugueses do têrmo com que el-rei os tratava - CondeDHenr XVII
- 4 – parangonas, Nasceu a Nova Atlântida, No Xadrez Mundial Moveu-se - JangPedra XX

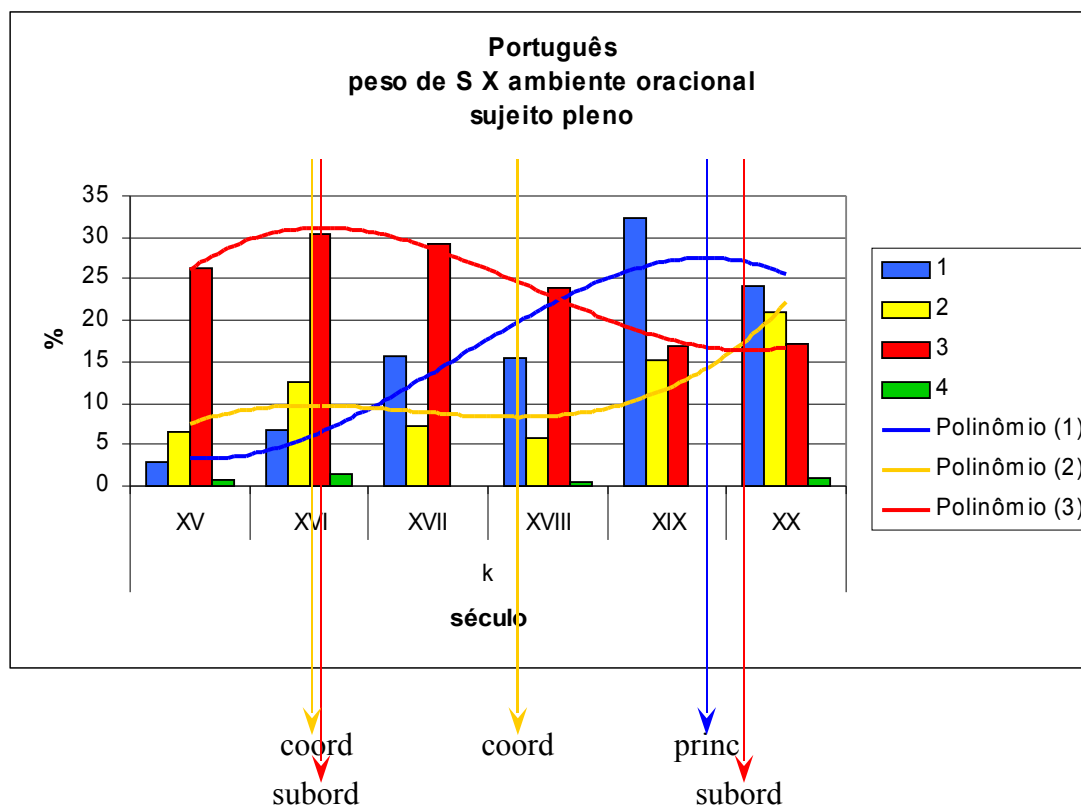


Gráfico P8a – distribuição das ocorrências de *sujeito pleno* segundo o *ambiente oracional*.

Para o sujeito pronominal, **I**, as freqüências absolutas são, como já assinalado acima, mais reduzidas que as de sujeito pleno; entretanto, é possível discernir claramente o decréscimo das

proporções absolutas de orações subordinadas – o ambiente mais comum para o sujeito pronominal –, que apresentam iRRL moderada: -1,98. Tomem-se os exemplos:

- 1 – com Castella. eu quero ordenar huuas festas rreaes que durem - Ceuta XV
- 2 – e por este respeito saem muitos por toda a costa de Melinde - PeJerLobo XVII
- 3 – e fazendo os mesmos gestos e as mesmas açções que ele fazia - IlhaAmor XVIII
- 4 – hua mesma dição, sem formar diphthongo, & fazer cada hua syllaba per si - DNLião XVI

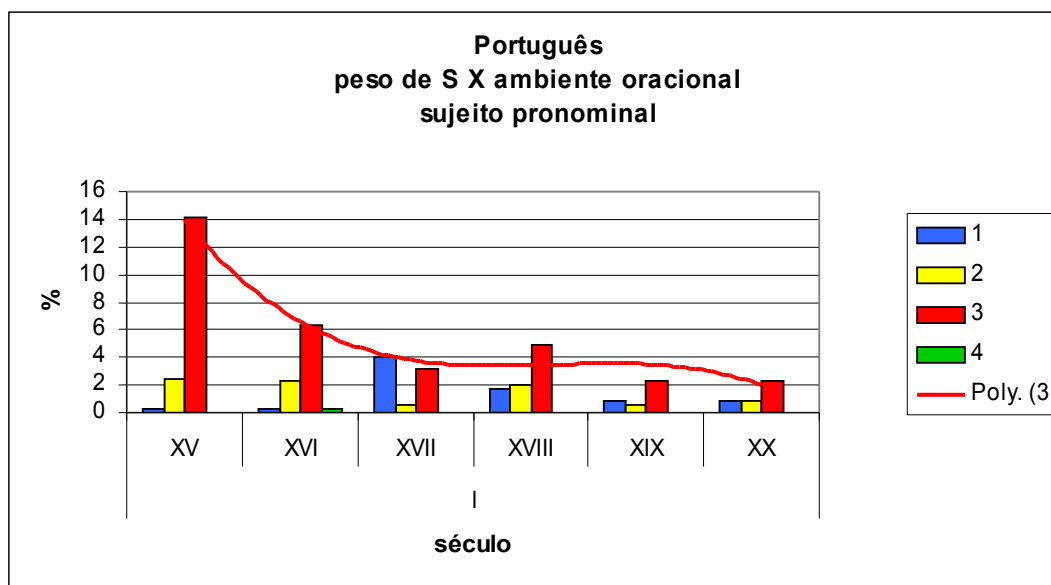


Gráfico P8b – distribuição das ocorrências de *sujeito pronominal* segundo o *ambiente oracional*.

Como já observado acima, as baixas proporções segundo as quais o sujeito oracional ocorre no *corpus* português não autorizam conclusões seguras sobre o comportamento desse subconjunto de dados, razão pela qual não serão feitos comentários a esse respeito.

No caso do complemento verbal pleno, classe **n**, temos, mais uma vez, o perfil crescente das freqüências associadas a orações principais (iRRL = 1.27); contudo, a flutuação das freqüências de **n** em orações coordenadas corresponde a uma situação de estabilidade (iRRL = 0,16), enquanto a freqüência desse tipo de complemento verbal em orações subordinadas diminui mais dramaticamente (iRRL = -3.80). Este último ambiente oracional, a despeito das proporções decrescentes em que ocorre, é o mais comum em todo o *corpus* português. Observa-se, mais uma vez, que algumas classes têm pontos críticos compreendidos no século XVIII, como **n1** (compl. verbal pleno em orações principais) – máximo absoluto – e **n2** (compl. verbal pleno em orações coordenadas) – mínimo absoluto. Exemplos desse traço são:

- 1 – nom sey pera as semelhantes ... mais proveitosa regra - LealCons XV
- 2 – E e manifesto que as linguas Grega & Latina primeiro forão grosseiras - FernOliv XVI
- 3 – a qual em seu tempo começou a levantar cabeça das ruínas - CondeDHenr XVII
- 4 – que o sol fazia a trajectória do costume nem podia fazer outra - JangPedra XX

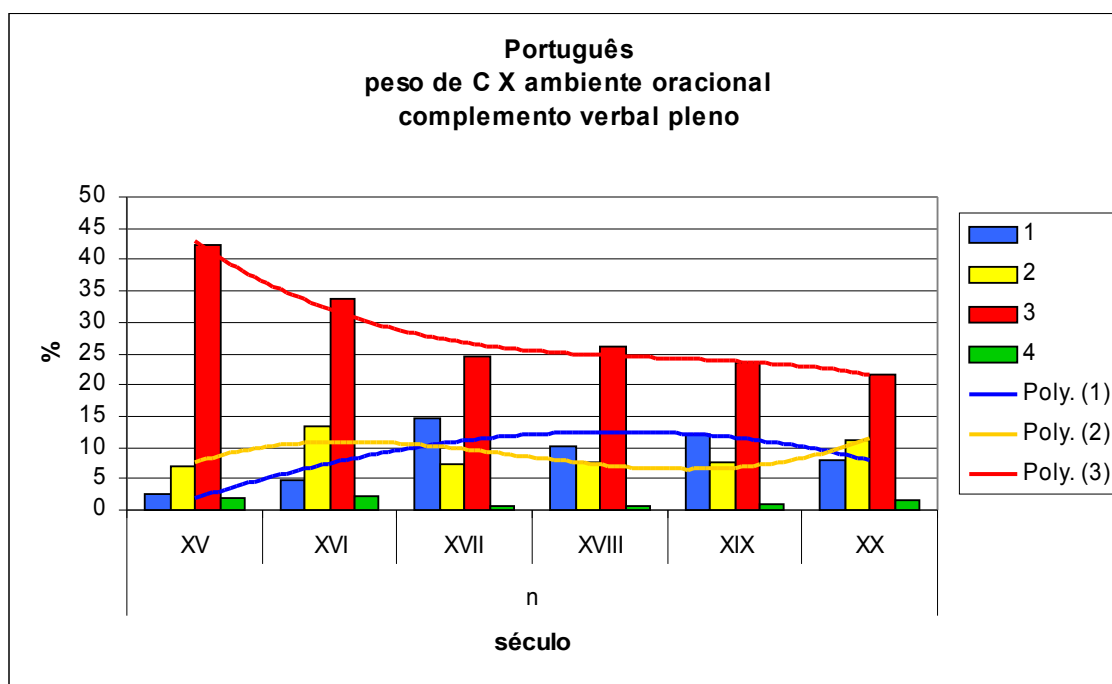


Gráfico P8c – distribuição das ocorrências de *complemento verbal pleno* segundo o *ambiente oracional*.

Situação similar se observa com respeito às ocorrências do complemento verbal pronominal, classe **o**: o ambiente oracional **3** (orações subordinadas) é o mais freqüente na base de dados portugueses, e as orações coordenadas mostram uma leve flutuação. Além disso, as médias absolutas de freqüência de orações subordinadas sofrem decréscimo ao longo do intervalo de amostragem – apesar de esse perfil de diminuição ser diferente daquele obtido para o complemento verbal pleno (classe **n**). Chama ainda a atenção o fato de o máximo absoluto da curva que resolve o perfil das ocorrências de **o3** abaixo coincidir com o século XVIII, a exemplo de outros pontos críticos para outras classes acima e da mudança de $AN > NA$ (Cohen 1986/1990). Exemplos desse caráter são apresentados abaixo:

- 1 – O caso nada tinha de assombroso, pode-se mesmo dizer que - JangPedra XX
- 2 – o rumor de passos morria: e ornavam-n'a divans cobertos de tapetes - Maias XIX
- 3 – O mesmo homem que nos tinha instruído do nome da ilha me disse - IlhaAmor XVIII
- 4 – que vio correr huma vea de agoa vermelha e o atreuhia a algum - PeJerLobo XVII

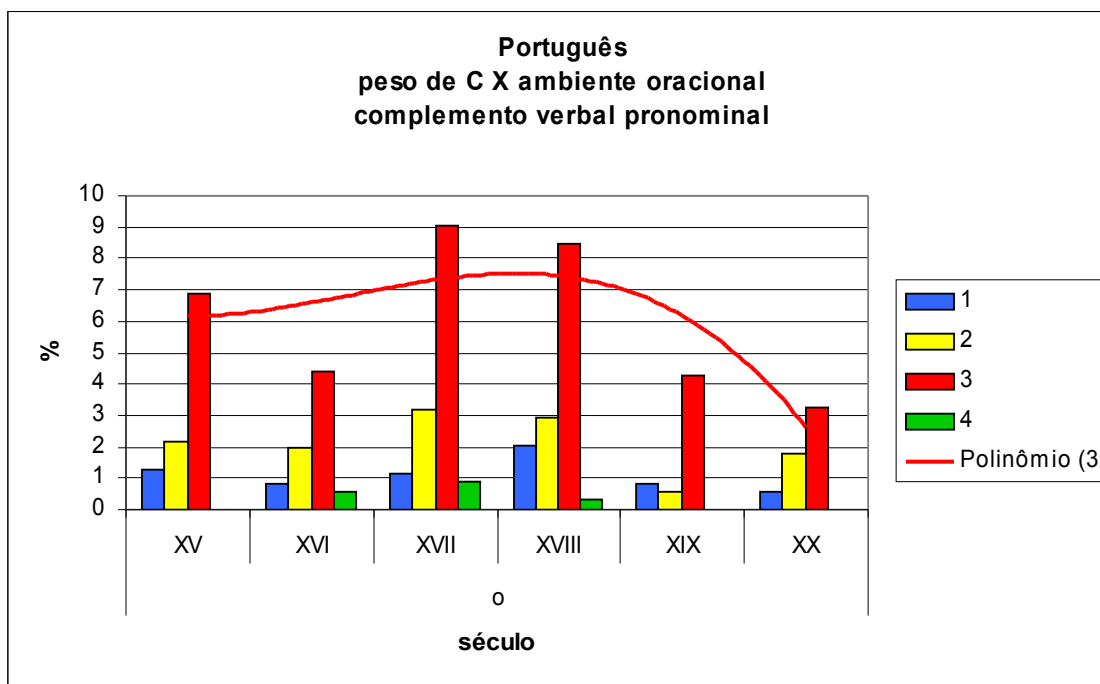


Gráfico P8d – distribuição das ocorrências de *complemento verbal pronominal* segundo o *ambiente oracional*.

Na Tabela P5 abaixo, em que se resumem as informações apresentadas nos Gráficos P8, percebe-se claramente a definição de padrões crescentes para as orações principais (1) e

decrecentes para as subordinadas (3); nessa tabela, portanto, o fator de maior relevância é o *ambiente oracional*, que dirige as tendências do *corpus* de maneira mais nítida do que o *peso de S e de C*.

	k (suj. pleno)	l (suj. pron.)	n (compl. pleno)	o (compl. pron.)
1 (or. princ.)	↗		↗	
2 (or. coord.)	↗		→	
3 (or. subord.)	↘	↘	↘	↘

Tabela P5 – tendências gerais de comportamento das classes de *natureza do sujeito* e *natureza do complemento* segundo o *ambiente oracional*.

Fixando-se o *ambiente oracional*, percebemos como evoluem as classes de *peso de S e de C* no *corpus*. Observa-se, em primeiro lugar, a preferência pelos traços **k** e **n**, cujas frequências absolutas predominam no Gráfico P9a; as retas de regressão de ambas têm inclinações positivas ($iRRL(\mathbf{k}) = 5.22$; $iRRL(\mathbf{n}) = 1.27$) em decorrência de as proporções aumentarem no intervalo de amostragem. As demais classes têm frequências absolutas muito baixas, e não passam dos 4%. Veja-se também que a escala vertical é extensa, corroborando a observação de que as orações principais vão-se equiparar, no final do intervalo, com os níveis de coordenadas e de subordinadas.

As ocorrências reproduzidas abaixo foram colhidas no *corpus* português, e representam as categorias mais freqüentes:

- k - Estes dous Reys ... se concertarão no anno do Senhor de 1478 - PrincDJoão XVIII
n - Este antigo castello tinha recordações de gloria - CastFaria XIX

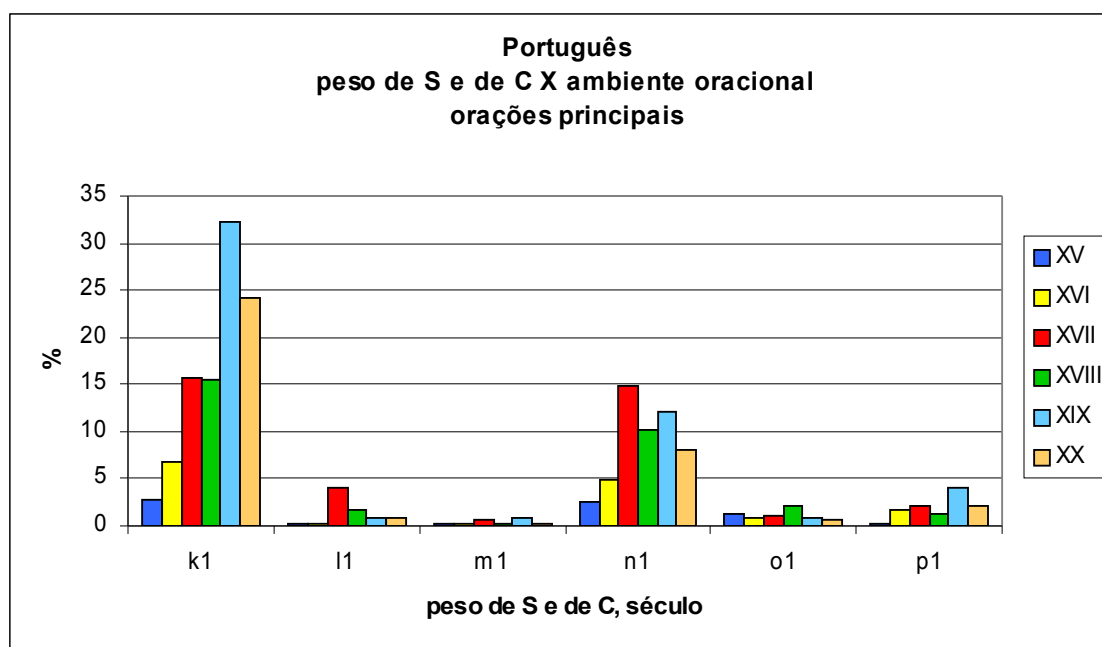


Gráfico P9a – distribuição das classes de *peso de S e de C* em orações principais.

As ocorrências das classes de *peso de S e de C* para as orações coordenadas mostram a mesma concentração em torno dos grupos **k** e **n** observada para as orações principais. Aqui, entretanto, a inclinação da reta de **k** continua positiva, mas a de **n** é negativa, em virtude do efeito provocado pelo pico no século XVI e pelo valor modesto atingido no século XX: $iRRL(\mathbf{k}) = 2.24$; $iRRL(\mathbf{n}) = -.36$. Como já assinalado para as orações principais, as classes **l**, **m**, **o** e **p** não chegam aos 5% de expressividade máxima, e não serão comentadas.

Para as classes mais frequentes, recolheram-se ocorrências como as que compõem os exemplos abaixo:

k - lugar, e as derradeiras sam postas no fundo do liçeçe quando o mestre - Ceuta XV
n - Continua a descrição e dá-se a razão do nome de mar Roxo - PeJerLobo XVII

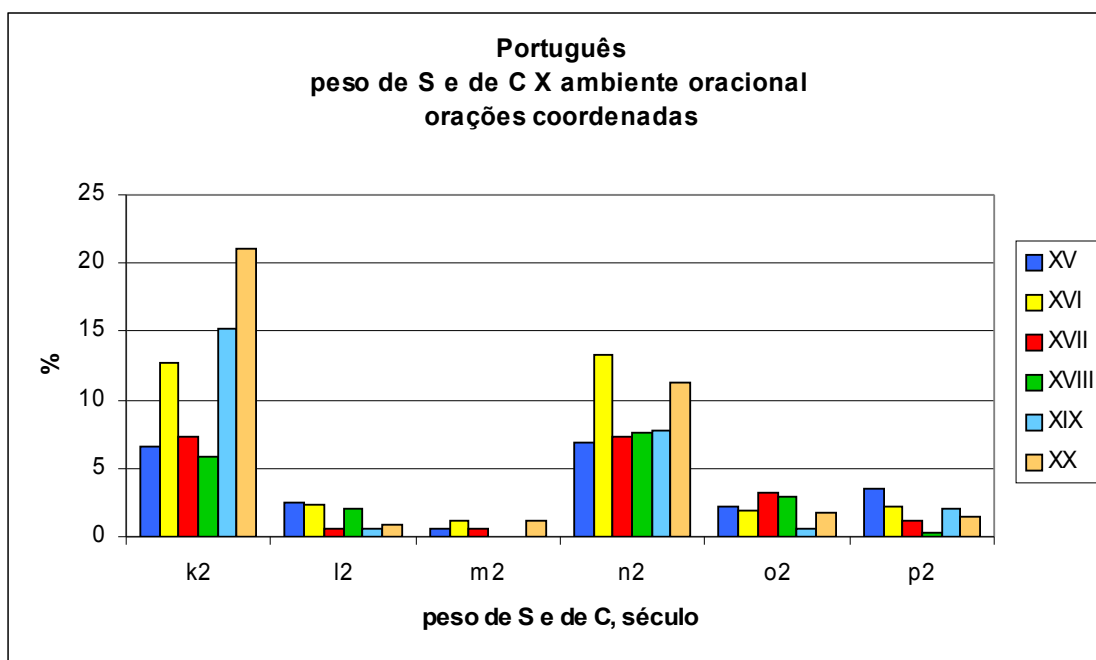


Gráfico P9b – distribuição das classes de *peso de S e de C* para orações coordenadas.

Também para as orações subordinadas há maior concentração das ocorrências em torno das classes **k** e **n**, apesar de as classes pronominais **l** e **o** já se pronunciarem em torno dos 10% de frequência absoluta. Contudo, se as classes mais comuns têm iRRL positiva em orações principais, nas subordinadas, ela é negativa: $iRRL(\mathbf{k}) = -2.64$; $iRRL(\mathbf{l}) = -1.98$; $iRRL(\mathbf{n}) = -2.49$; $iRRL(\mathbf{o}) = -.54$; $iRRL(\mathbf{p}) = -.77$. A escala do gráfico P9c é mais extensa que as duas anteriores, o que aponta para a maior frequência desse ambiente oracional, como já foi comentado.

k - que sempre essa consoante vá com a vogal seguinte - DNLião XVI
n - lenta que costuma devorar os gigantes de marmore e granito - CastFaria XIX

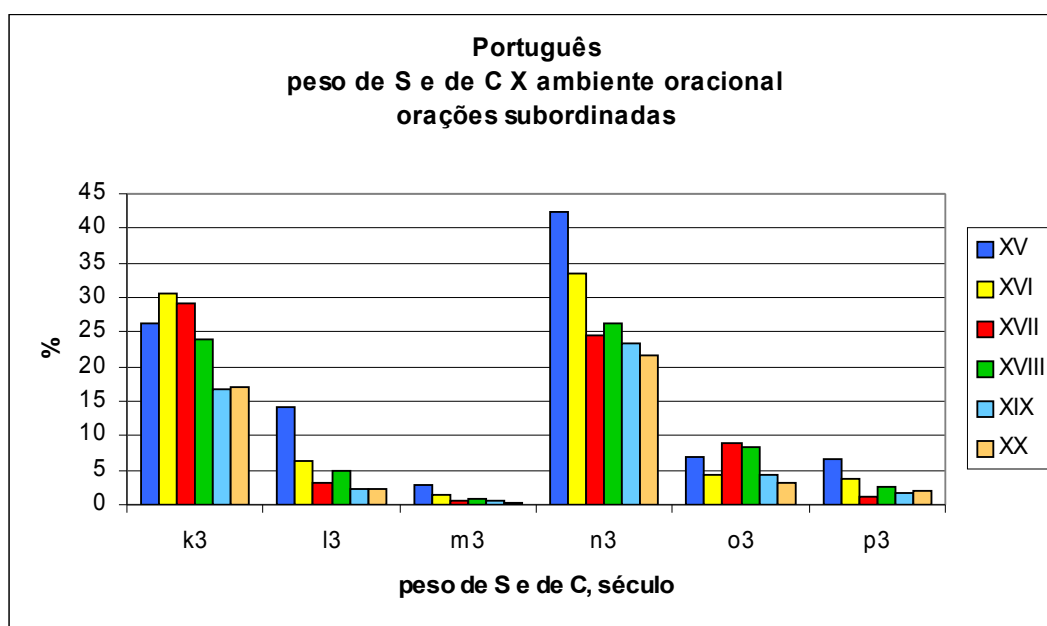


Gráfico P9c – distribuição das classes de *peso de S e de C* para *orações subordinadas*.

As frequências absolutas de orações coordenadas em subordinação não chegam aos 2%, e não é possível divisar tendências de comportamento das classes de *peso de S e de C* nesse ambiente.

Repetem-se aqui, portanto, as observações apontadas à Tabela P4: o fator determinante, *ambiente oracional*, mostra o aumento das classes mais comuns de *peso de S e de C* em orações principais e a diminuição das frequências das classes mais expressivas de subordinadas. Não descarto, com isso, a proposição de Abeillé & Godard 1997, segundo a qual o peso do SN deve

ser considerado numa análise detalhada da evolução sintática de línguas românicas; entretanto, deixo clara a necessidade de se contemplar, além desse fator, o *ambiente oracional* em que o SN se insere, uma vez que os resultados numéricos aqui obtidos são bastante sugestivos da relevância desse critério diante do *peso do SN*.

O fator *tempo do verbo* mostra uma configuração peculiar, segundo a qual há predominância absoluta da classe α sobre as demais. No caso do *corpus* português, essa classe dominante tem um perfil praticamente especular em relação ao de β , como se pode observar no gráfico abaixo; os pontos de máximo absoluto de α e de mínimo absoluto de β são coincidentes no final do século XVI, mas os pontos críticos seguintes, ainda que situados no século XIX, mostram leve defasagem.

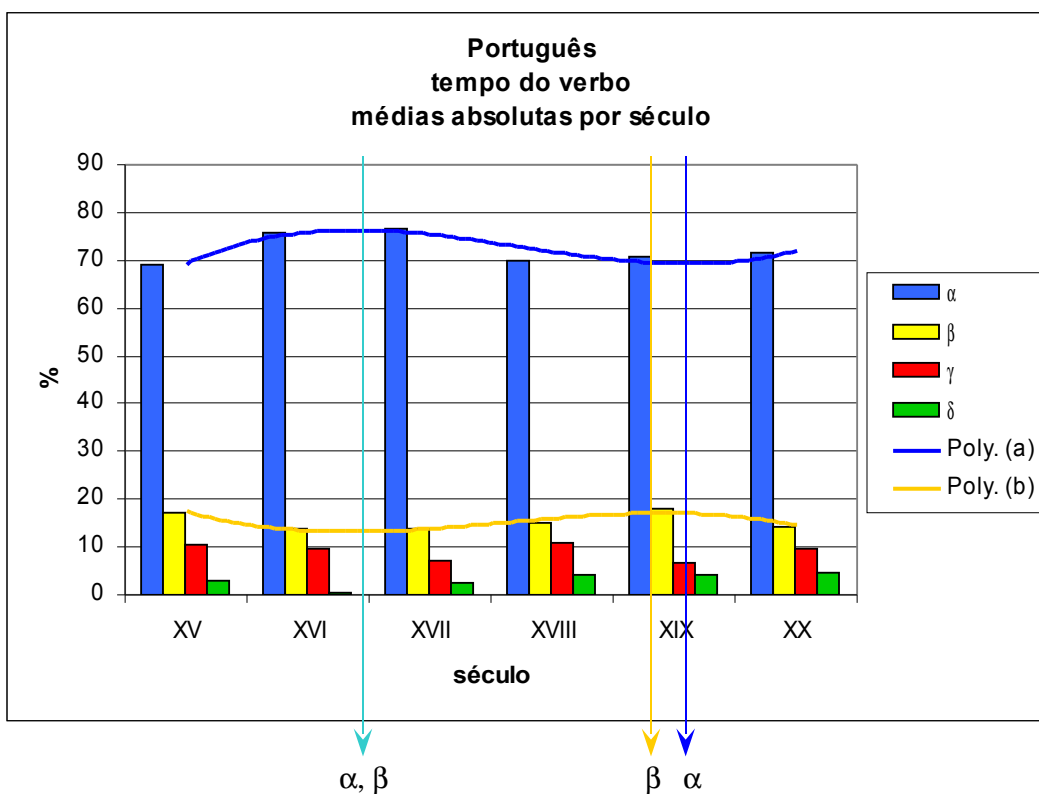


Gráfico P10 – distribuição geral das ocorrências
segundo o critério *tempo do verbo*.

A conjugação dos fatores *tempo do verbo* e *ordem* oferece os gráficos P10a a P10c, que ilustram um comportamento complexo do sistema ao longo da amostragem.

Em P10a, que trata das formas finitas do verbo (α), percebe-se claramente que as classes SVC e VC ocorrem em proporções levemente decrescentes ($iRRL(\mathbf{a}) = -.37$; $iRRL(\mathbf{i}) = -1.41$), ao passo que SV e VS, sem o complemento verbal expresso, aumentam em relação às demais ($iRRL(\mathbf{g}) = 2.29$; $iRRL(\mathbf{h}) = .75$). Portanto, o esvaziamento das ordens formadas pelos constituintes V e C é acompanhado da expansão das que não trazem este último. A classe α é também a mais freqüente das categorias em que se divide o fator *tempo do verbo* neste trabalho; algumas ocorrências colhidas no *corpus* são reproduzidas abaixo:

- a (SVC) - mas aquelles ... conhecerám bem quanto verdadeiramente screvo - LealCons XV
 c (VSC) - o alcançou, dá êle mesmo testemunho em certa doação - CondeDHenr XVII
 g (SV) - *Ides buscar, aventureiros, o verdadeiro gosto que Amor inspira* - IlhaAmor XVIII
 h (VS) - por elle, tão em vão, como passavam em vão ... annos e vendavaes - Maias XIX

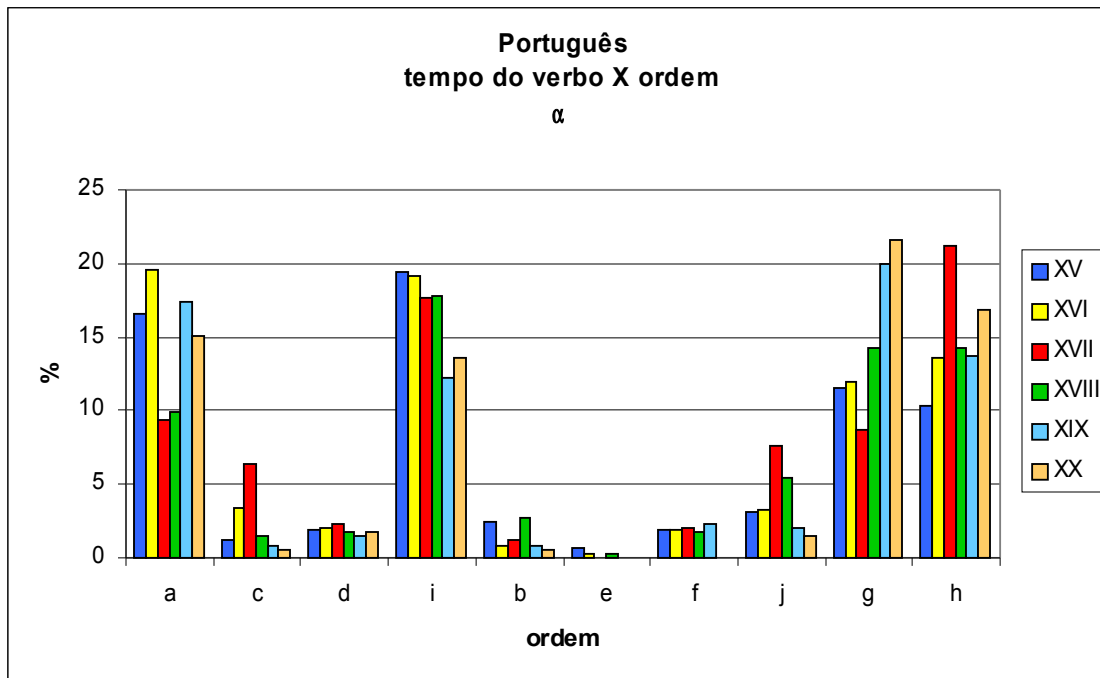


Gráfico P10a – distribuição das freqüências relativas da ordem de constituintes para formas finitas do verbo.

Para as formas reduzidas do verbo (β), a segunda categoria mais freqüente do traço em questão, verifica-se a predominância absoluta – e discretamente crescente – da ordem **i** (VC) em todo o intervalo de amostragem ($iRRL(i) = .39$); as demais ordens são minoritárias, e ocorrem segundo freqüências absolutas invariavelmente inferiores a 5%. Observe-se que os perfis obtidos para esta classe β indicam que a mesma tem comportamento diferente do apresentado pelas formas finitas do verbo, cf. gráfico P10a: α tem uma distribuição mais difusa, enquanto as formas reduzidas do verbo (β) se especializam na ordem **i** (VC).

Algumas das ocorrências que constituem o *corpus* são reproduzidas abaixo:

- g (SV) – e temperado spedimento quando cadahuu se partir - LealCons XV
h (VS) – do porto de gaya como quer Duarte galuão na estoria del rey - FernOliv XVI
i (VC) – o que notei, referindo-o com o Patriarcha de Ethiopia - PeJerLobo XVII
j (CV) – ca... se nom achar que he seruiço de Deos nom entendo de o fazer – Ceuta XV

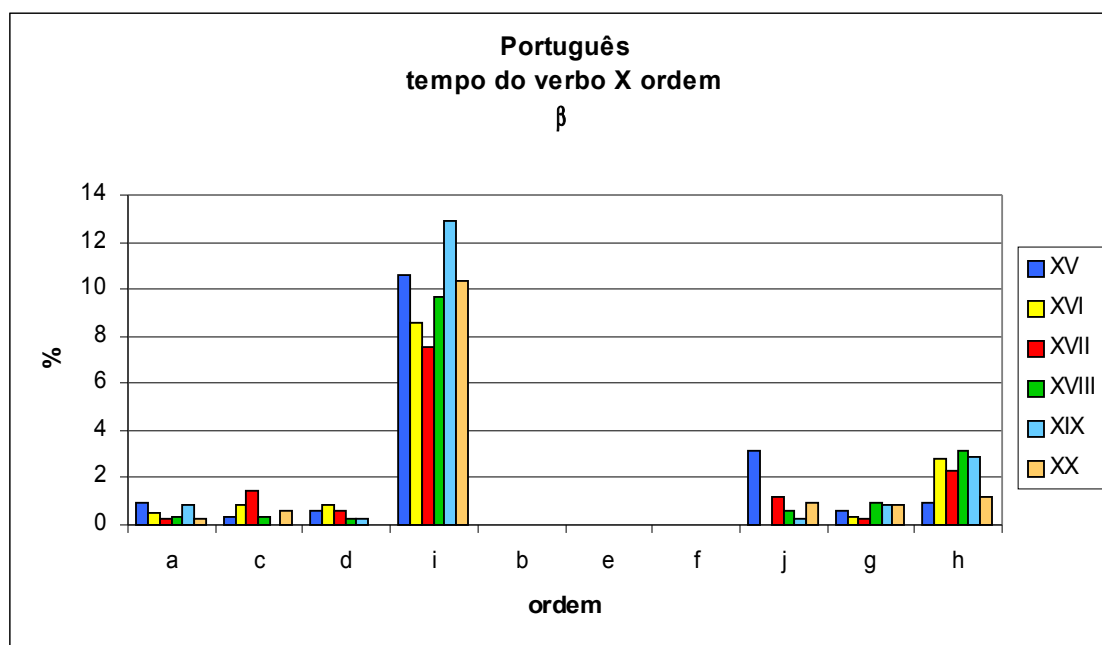


Gráfico P10b – distribuição das freqüências relativas da ordem de constituintes para formas reduzidas do verbo.

As locuções verbais, em vez de se especializarem numa ordem como o fazem as formas reduzidas, distribuem-se preferencialmente nas mesmas ordens principais adotadas pelas formas finitas – **a** (SVC), **i** (VC), **g** (SV) e **h** (VS). A extensão da escala de ordenadas é bem menor do que nas classes anteriores e, em decorrência disso, é temerário tirar conclusões sobre o comportamento dessa classe; ainda assim, os perfis ascendente de **a** (SVC) e descendente de **i** (VC) parecem claros, a despeito de a amplitude dessas mudanças corresponder a pouco mais de 1%. Tomem-se os exemplos abaixo:

a (SVC) – O Estado da fortuna pode cõceder ou tirar fauor aos estudos - FernOliv XVI

i (VC) – que pudesem comonicar o nome a todo o mar - PeJerLobo XVII

g (SV) – De como EIRey desembarcou com sua gente, e mandou - PrincDJoão XVIII

h (VS) – Era esta lagea a mesa em que costumava comer Salat-ibn-Salat - CastFaria XIX

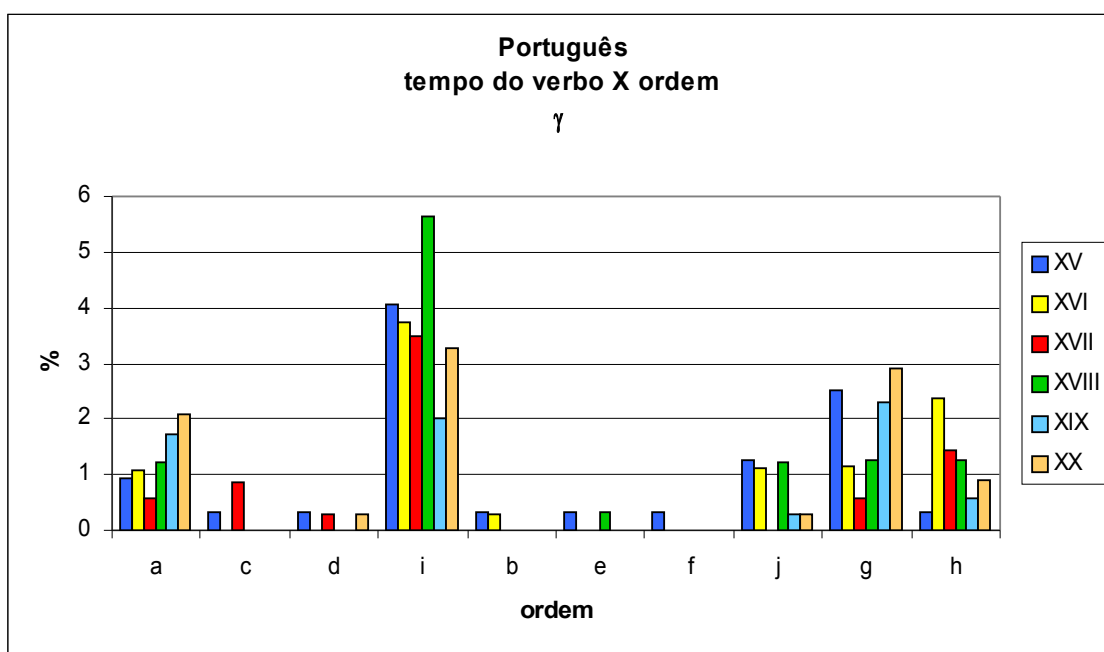


Gráfico P10c – distribuição das frequências relativas da ordem de constituintes relacionadas a locuções verbais.

Os tempos compostos ocorrem em proporções extremamente baixas no *corpus*, e os máximos absolutos para essa classe não ultrapassam os 2%; por esse motivo, não serão feitos comentários adicionais sobre o comportamento dessa categoria.

Em resumo, o fator *tempo do verbo* mostra, para as ordens mais frequentes, as tendências gerais apresentadas à Tabela P5 abaixo (em cinza, as tendências identificadas estatisticamente, mas que devem ser consideradas com cautela devido à baixa frequência absoluta da classe γ); pode-se notar que o *tempo do verbo* oferece correlações mais claras do que o fator *ordem de constituintes*. Esses resultados oferecem novos argumentos para corroborar as conclusões de Bybee *et al.* 1994, que destacam o papel de fatores como o *tempo verbal* como indicadores de mudança lingüística.

<i>ordem</i>	<i>tempo do verbo</i>		
	α	β	γ
a	↘	→	↗
i	↘	↗	↘
g	↗	↗	↗
h	↗	↗	↘

Tabela P5 – tendências gerais de comportamento das classes de *tempo do verbo*.

O cruzamento dos fatores *tempo do verbo* e *ambiente oracional* mostra que as frequências absolutas das formas verbais finitas aumentam em orações principais (1), mas diminuem nas orações subordinadas (3). As orações coordenadas (2) flutuam no intervalo de amostragem, e apresentam leve tendência a aumentar ($iRRL(2) = 1.22$); percebe-se que o perfil dessa classe é coincidente com o apresentado no Gráfico P6 acima, no qual a curva tem um ponto de mínimo no século XVIII. Em outras palavras, a distribuição das ocorrências segundo o fator *tempo do verbo* é dependente do *ambiente oracional*. Algumas das ocorrências do *corpus* estão reproduzidas abaixo:

- 1 – nom sey pera as semelhantes ... mais proveitosa regra - LealCons XV
- 2 – Mas em nossa lingoa há XVI. diphthongos .s. ãa, ãe ... - DNLião XVI
- 3 – pois diz que, quando fêz aquêlê testamento, estava de caminho - CondeDHenr XVII
- 4 – de rogar para que ele ... nelas entrasse e derramasse - JangPedra XX

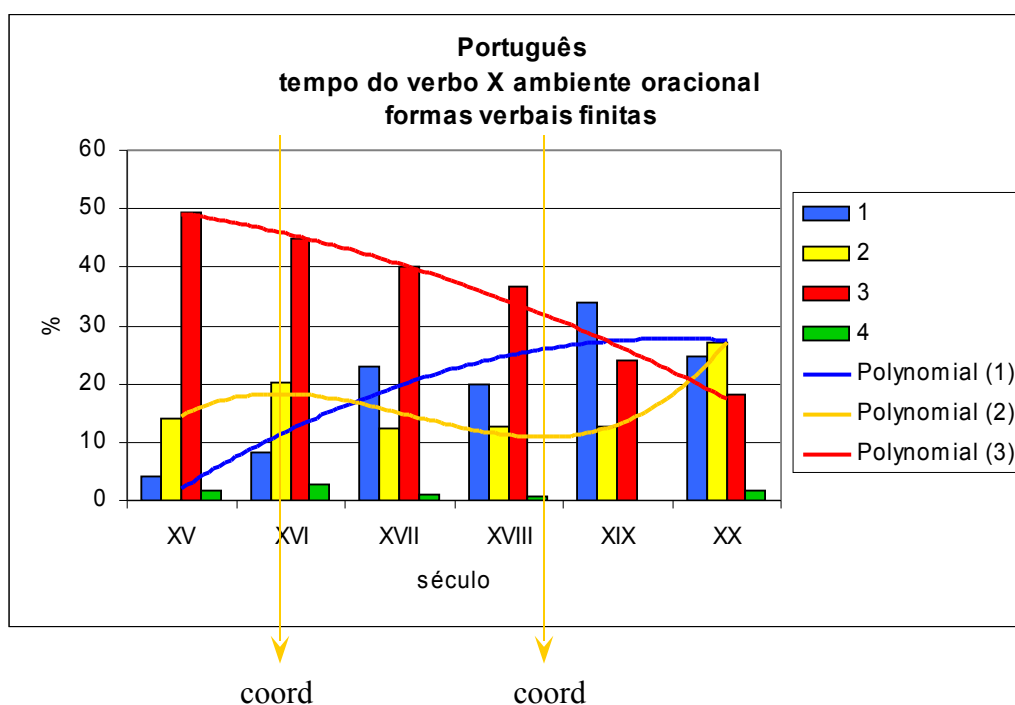


Gráfico P11a – distribuição de ocorrências da classe α segundo o *ambiente oracional*.

As formas reduzidas do verbo (classe β) especializam-se, definitivamente, em orações subordinadas; o perfil evolutivo das frequências absolutas dessa classe é dado por uma senóide deslocada para a direita e invertida em relação à que corresponde às coordenadas (curva 2, em amarelo) no gráfico anterior. Os demais ambientes oracionais não comportam proporções expressivas de formas reduzidas do verbo, como se pode observar a partir do exame do Gráfico P11b. Vamos aos exemplos e ao gráfico de distribuição:

- 1 – a graça do senhor deos por bem e sagesmente o bem fallar praticarmos - LealCons XV
- 2 – per .am., & succeder aquelle diphthongo, em lugar das dictas letras - DNLião XVI
- 3 – que, o que eu mandar e ordenar, seja firme e estável - CondeDHenr XVII
- 4 – ou por defender ou por descercar Lisboa - CastFaria XIX

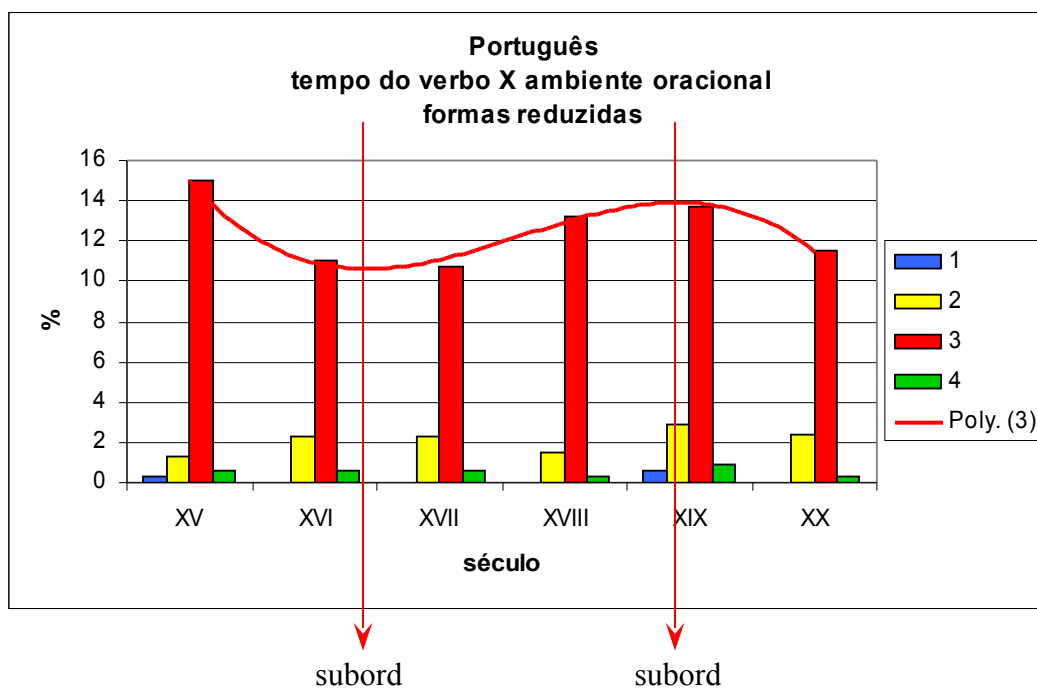


Gráfico P11b – distribuição de ocorrências da classe β segundo o *ambiente oracional*.

As locuções verbais ocorrem preferencialmente associadas às orações subordinadas, a despeito da flutuação das médias absolutas dessa classe no *corpus* e da tendência levemente decrescente ($iRRL = -.52$) desse subconjunto de dados. O ponto excêntrico dessa curva no século XIX é certamente devido ao efeito provocado por um dos textos – tanto mais sensível em virtude de a escala do gráfico ser reduzida.

- 1 – eu quero ordenar huuas festas rreaes que durem todo hum anno - Ceuta XV
- 2 – Mas bem podem ocorrer estas duas vogaes, sem formar diphthongo - DNLIão XVI
- 3 – e mandou subir huma pessoa à gavea a qual notou o mesmo - PeJerLobo XVII
- 4 – a Villa naõ pudesse ser soccorrida, nem della pudesse sahir pessoa - PrinDJoão XIX

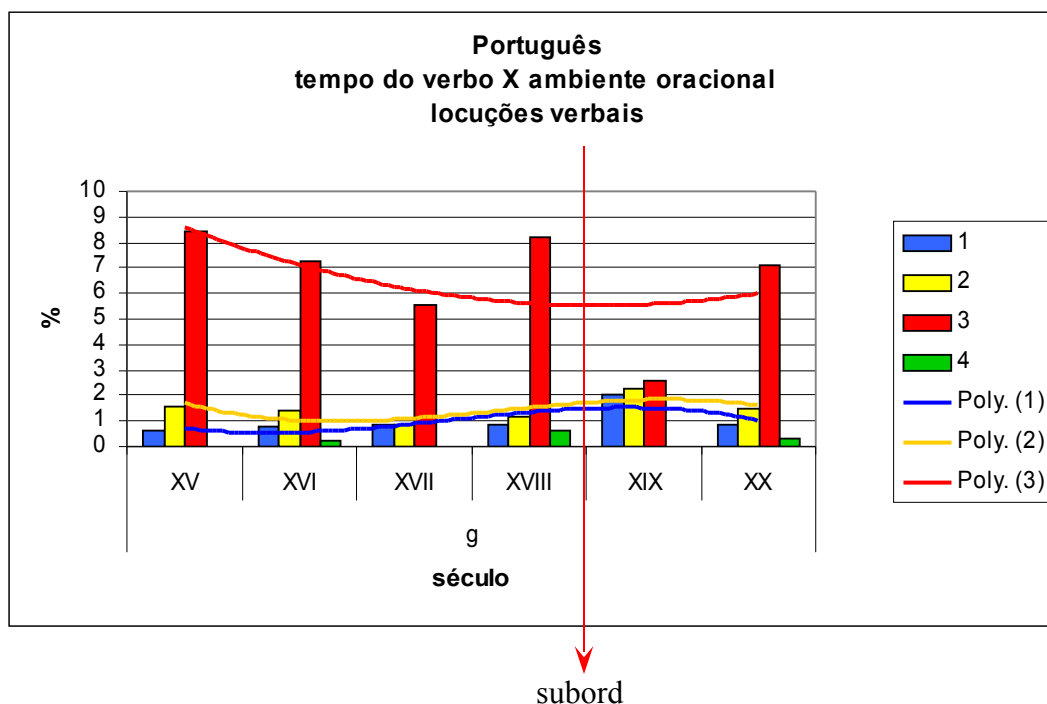


Gráfico P11c – distribuição de ocorrências da classe γ segundo o *ambiente oracional*.

Isolando-se apenas as orações principais (classe 1 do fator *ambiente oracional*), observa-se que as ocorrências associadas a esse ambiente oracional estão absolutamente concentradas nas formas finitas do verbo; o máximo da curva que representa o perfil dessa classe está situado na passagem do século XIX para o XX. Os exemplos abaixo são referentes a esse traço:

α - desta derradeira chamada Lisibona/apõta somẽte Plinio no terceiro liuro - FernOliv XVI

β - a graça do senhor deos por bem e sagesmente o bem fallar praticarmos - LealCons XV

γ - No ano de 1080 mandou fazer a igreja de S. Martinho - CondeDHenr XVII

δ - A venda da Tojeira fõra realmente aconselhada por Villaça - Maias XIX

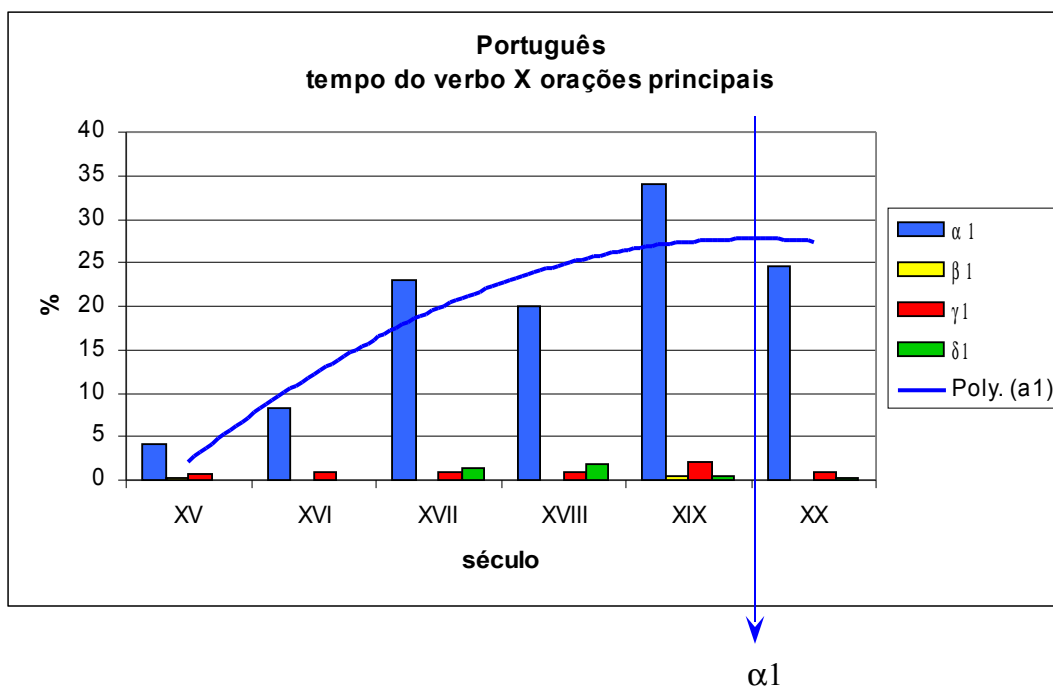


Gráfico P11d – distribuição de ocorrências de orações principais segundo o *tempo do verbo*.

Apesar de as formas finitas do verbo dominarem o gráfico correspondente às ocorrências de orações coordenadas, o perfil dessa classe é diferente do observado para as orações principais. Temos, aqui, uma curva que apresenta máximo (séc. XVI) e mínimo (séc. XVIII) locais – ambos situados antes do máximo das frequências para as orações principais, como se pode observar comparando-se os Gráficos P11d e P11e –, e, portanto, coincidente com a forma da curva 2 (orações coordenadas) no Gráfico P6. Tomem-se os exemplos:

- α - E com esto darey tantas e tam grandes dadiuas prinçipalmente aaquelles - Ceuta XV
- β - Turdugal:& agora mudãdo alghuas letras Portugal/nã do porto - FernOliv XVI
- γ - e tôdas estas coisas confirmam ser o conde secular e não eclesiástico - CondeDHenr XVII
- δ - e o exercito castelhano foi constringido a levantar o cerco - CastFaria XIX

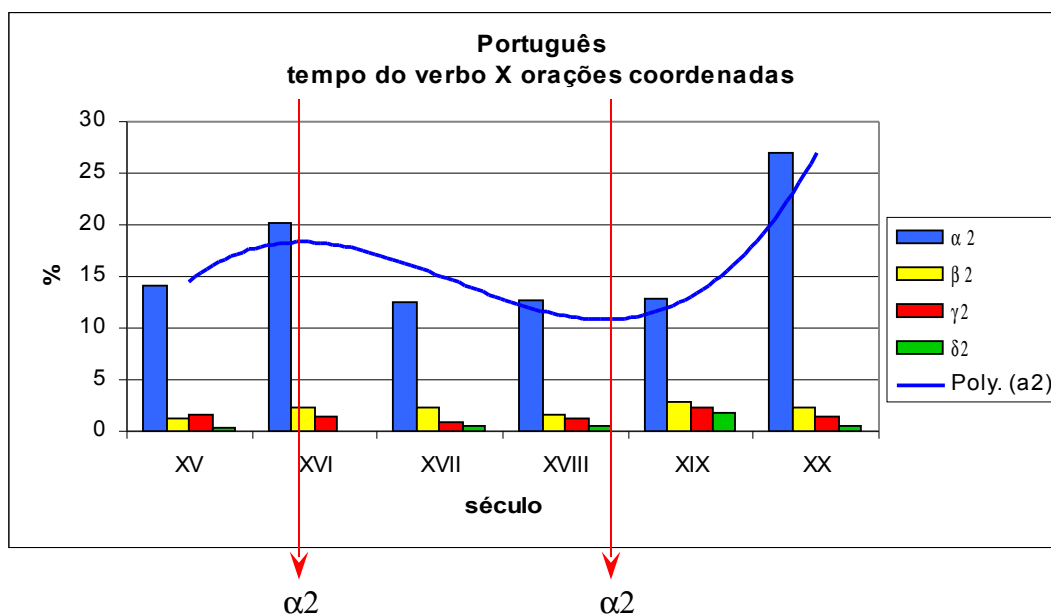


Gráfico P11e – distribuição das ocorrências de orações coordenadas segundo o *tempo do verbo*.

De maneira semelhante, a classe mais freqüente no gráfico relativo às orações subordinadas, α_3 , obedece ao perfil sempre decrescente apresentado para esse ambiente oracional no Gráfico P6, e tem $iRRL = -6.35$. As freqüências absolutas das formas reduzidas mostram variação estável no intervalo estudado. O *corpus* português é constituído de ocorrências semelhantes às dispostas abaixo:

α – a quẽ deos quis dar aqlla bem auenturança de viuer - FernOliv XVI

β – parece querem encontrar na restauração de Braga - CondeDHenr XVII

γ – lenta que costuma devorar os gigantes de marmore e granito - CastFaria XIX

δ – que o sol fazia a trajectória do costume nem podia fazer outra - JangPedra XX

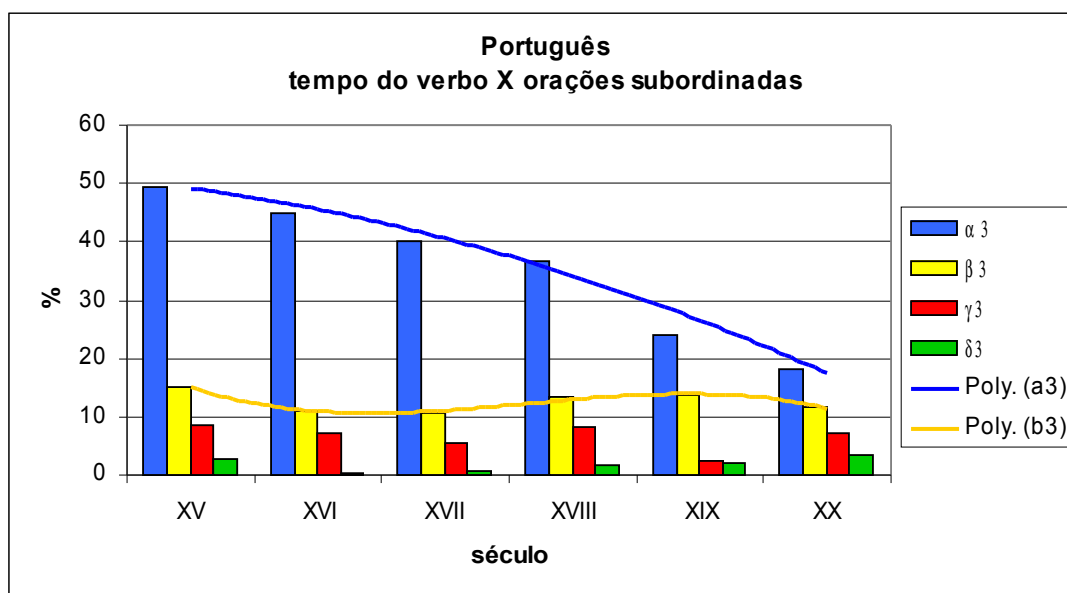


Gráfico P11f – distribuição das ocorrências de orações subordinadas segundo o *tempo do verbo*.

O fator *regência do verbo* também foi sobreposto a *ordem dos constituintes*, permitindo a construção da série de gráficos P12; as categorias mostram um comportamento especializado em relação à ordem, o que contribui para a caracterização sintática da história do português e retoma, a exemplo do que foi feito para o *tempo do verbo*, as observações de Bybee *et al.* 1994 sobre a importância da *regência verbal* para o estudo da mudança lingüística.

O Gráfico P12a mostra que uma das classes é dominante em todo o intervalo de amostragem (assemelhando-se, nesse sentido, ao Gráfico P10), e que todas têm um perfil em que há pequena variação.

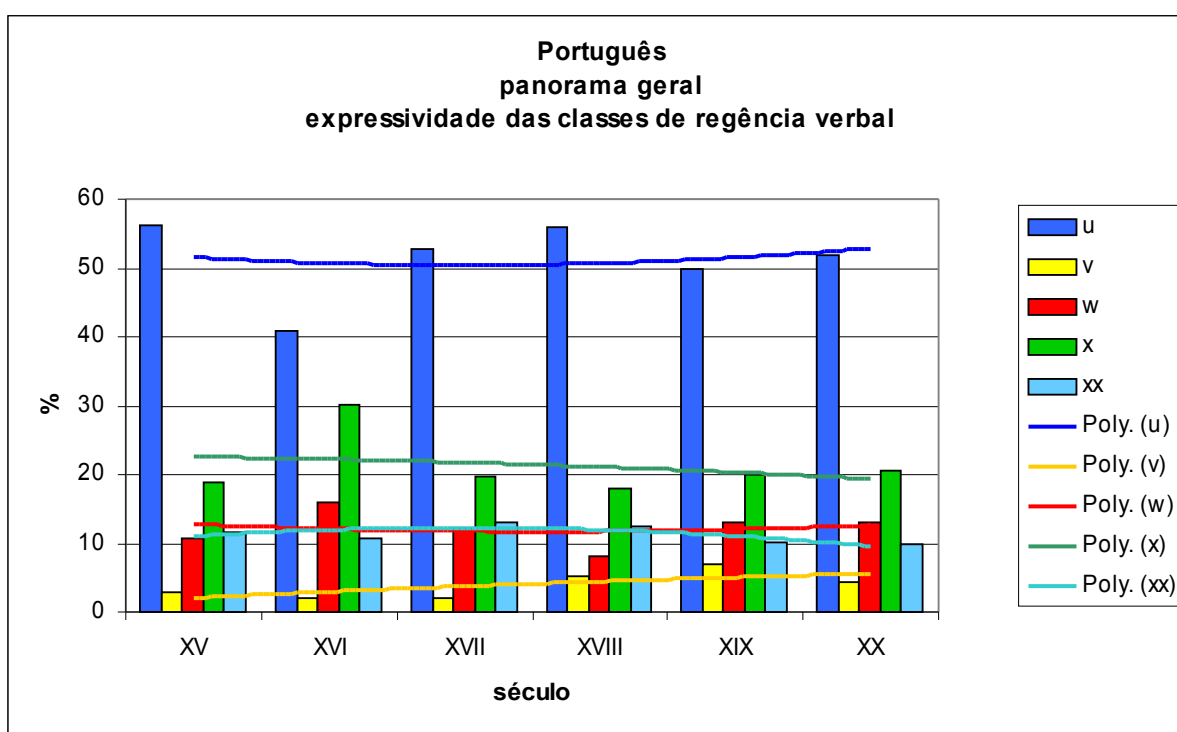


Gráfico P12a – distribuição das ocorrências segundo o fator *regência do verbo*.

A classe **u** (transitivos diretos), cuja distribuição em relação à ordem dos constituintes é apresentada em P12b, é a mais freqüente de todas e tende a organizar-se preferencialmente em torno das ordens **i** (VC) e **a** (SVC); enquanto a primeira é resolvida por reta de regressão com inclinação negativa (-.87), a regressão da segunda fornece um valor próximo de zero (iRRL = .20), a despeito da flutuação moderada que nela se verifica. Ainda que as outras ordens tenham

freqüências relativas bem mais reduzidas, podem-se divisar a) uma tendência à horizontalidade para a reta de regressão da classe **g** ($iRRL(SV) = -.10$), b) um perfil decrescente para a classe **j** ($iRRL(CV) = -.75$) e c) uma discreta elevação dos valores de **h** ($iRRL(VS) = .38$) ao longo do tempo. Vejam-se os exemplos colhidos para várias ordens de realização do traço **u** no *corpus* português:

- a (SVC) - dos benefícios que lhes eu assi fazer lhes ponha neçessidade - Ceuta XV
 c (VSC) – mas ã muitas outras cousas tem anossa lingoa auantagẽ - FernOliv XVI
 d (VCS) – com sua espada e poder real, favorecendo-o Deus - CondeDHenr XVII
 i (VC) – valor para vos fazer uma relação exacta da minha viagem - IlhaAmor XVIII
 g (SV) – e dois longos bancos feudaes que Carlos trouxera de - Maias XIX
 h (VS) – propunha-se a adopção, pelos dois países peninsulares - JangPedra XX

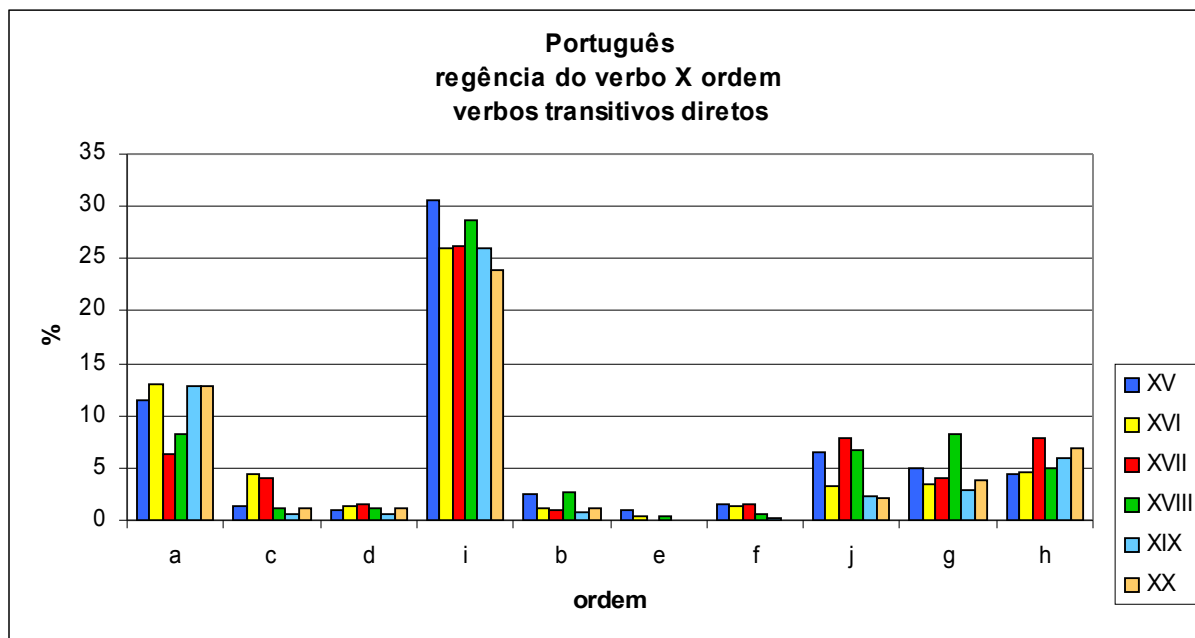


Gráfico P12b – frequências absolutas da classe u segundo a *ordem de constituintes*.

Para os verbos de ligação (classe **x**), cuja frequência absoluta é menor que a da classe **u** (transitivos diretos), nota-se também uma preferência pelas ordens **a** (SVC) e **i** (VC), ambas com perfis decrescentes; entretanto, as ordens **g** (SV) e **h** (VS), apesar de muito menos comuns, mostram franca elevação nos textos dos séculos XV a XX. Tomem-se os exemplos seguintes:

a (SVC) – os maridos das boas molheres... serem sempre amados - LealCons XV

d (VCS) – de Arouca eram senhores ... Egas Ermiguio, Odorio - CondeDHenr XVII

i (VC) – o qual sendo já Rey veyo cercar duas, ou tres vezes - PrincDJoão XVIII

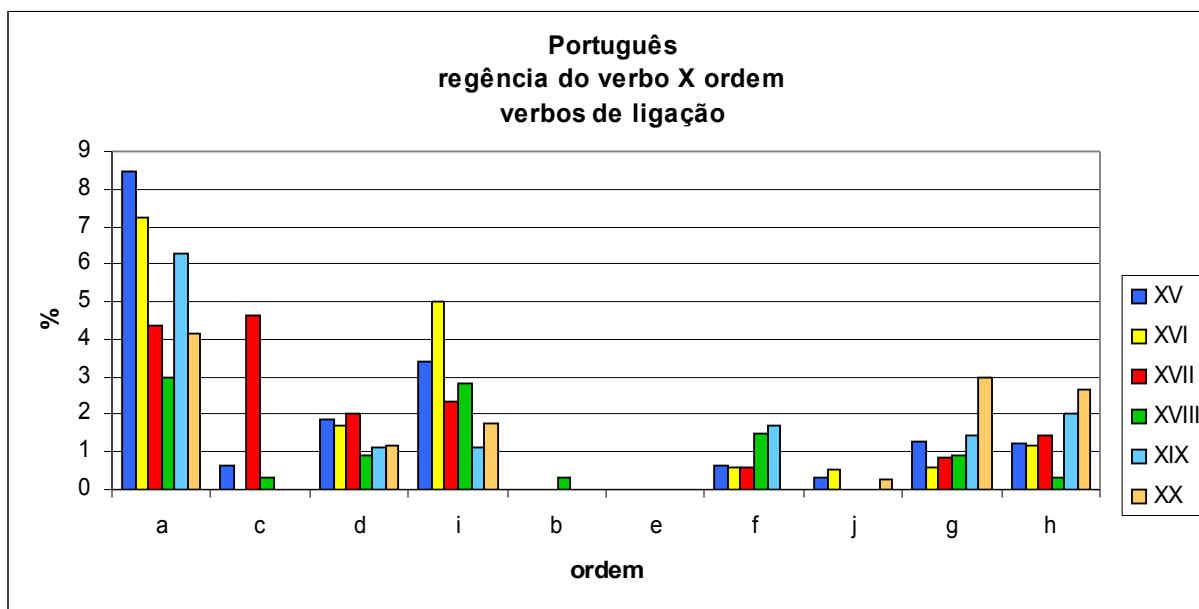


Gráfico P12c – distribuição das frequências absolutas da *ordem de constituintes* relacionadas a *verbos de ligação*.

Os intransitivos (classe **xx**), por sua vez, têm perfis crescentes tanto para **g** (iRRL(SV) = 1.35) quanto para **h** (iRRL(VS) = .16); trata-se também de uma das categorias mais frequentes no *corpus* português. Os exemplos abaixo ilustram as duas ordens em que essa classe se realiza:

g (SV) – copia na paragem por onde as naos vem da India - PeJerLobo XVII

h (VS) – que as mais das vezes nace a pryncipal culpa de nos - LealCons XV

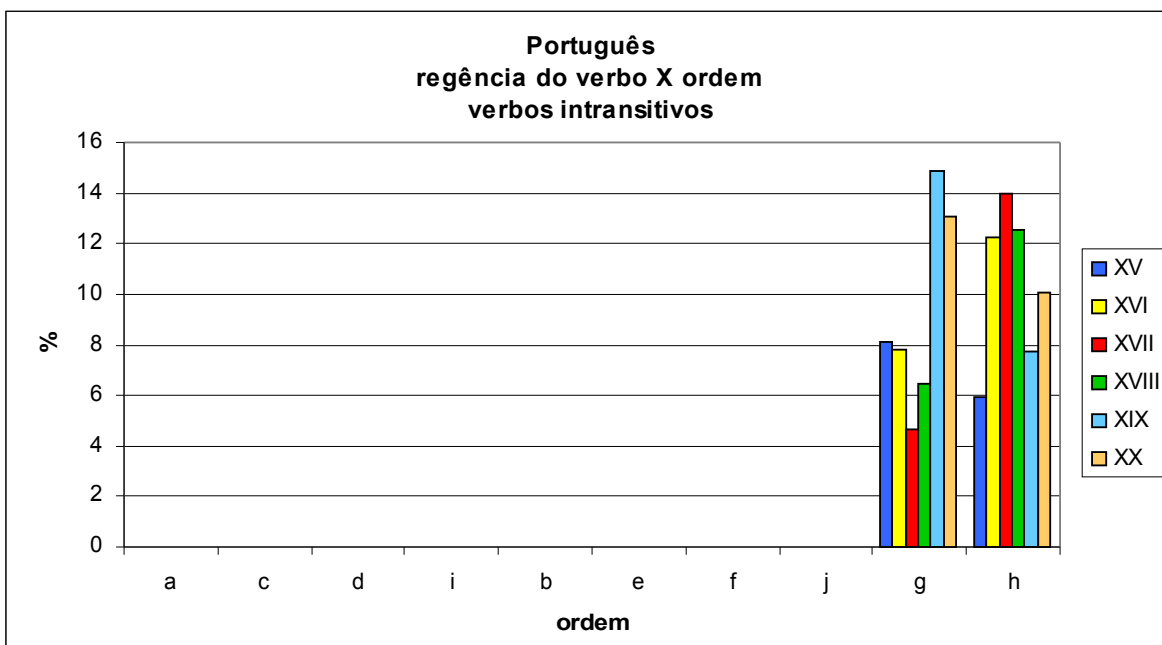


Gráfico P12d – frequências absolutas da classe xx segundo a *ordem de constituintes*.

A classe **v** (transitivos indiretos) ocorre preferencialmente na ordem SV: os valores de frequência para essa classe compõem um perfil ascendente no grupo **g** (iRRL(SV) = 6.48), mas os demais não oferecem segurança para que se tomem conclusões sobre seu comportamento. Eis os exemplos:

g (SV) – Isso, como dizia Villaça, acontecia a todos os muros - Maias XIX

g (SV) – o autor da idéia olhou para o mapa e verificou que - JangPedra XX

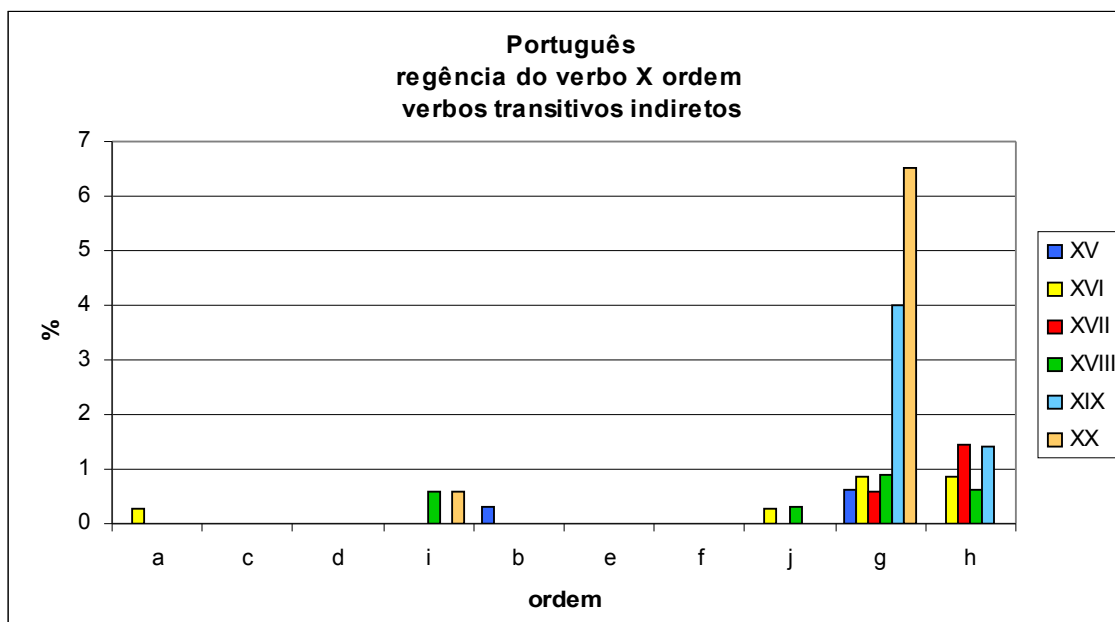


Gráfico P12e – frequências absolutas da classe v segundo a *ordem de constituintes*.

A classe **w** (bitransitivos, ou transitivos diretos e indiretos), por fim, é a menos comum no *corpus* português, e as frequências absolutas das categorias associadas a esse tipo de verbo não passam dos 4%; as ordens mais frequentes, **a** (SVC), **i** (VC) e **j** (CV), mostram uma flutuação cuja regressão fornece uma inclinação muito próxima de zero, enquanto **g** (SV) e **h** (VS) esboçam uma tendência crescente.

Em resumo, o comportamento das classes do fator *regência do verbo* indica, para as ordens **g** (SV) e **h** (VS), uma tendência generalizada ao crescimento, e as outras ordens têm perfis estáveis ou decrescentes.

Através do cruzamento dos fatores *regência do verbo* e *ambiente oracional*, podem-se considerar com mais detalhe alguns aspectos evolutivos do sistema português – como, por exemplo, a relação entre esses critérios de recorte da base de dados e a relevância de cada fator. Começamos com o isolamento das classes de regência verbal, a fim de observarmos como cada uma delas se comporta em termos de frequência absoluta nos ambientes oracionais propostos neste trabalho.

Para os verbos transitivos diretos (classe **u**), as curvas tendem a repetir o padrão geral definido no Gráfico P6: as orações subordinadas (classe **3** de *ambiente oracional*), mais frequentes que as demais, exibem perfil decrescente ($iRRL(\mathbf{u3}) = -4.15$), ao passo que as orações

principais aumentam ($iRRL(\mathbf{u1}) = 1.94$) e as coordenadas flutuam em torno dos 10% ($iRRL(\mathbf{u2}) = .02$).

- 1 – desta derradeira chamada Lisibona/apõta somête Plinio - FernOliv XVI
- 2 – e dá-se a razão do nome de mar Roxo ou Vermelho - PeJerLobo XVII
- 3 – que nos lançou com violência à outra parte oposta àquela - IlhaAmor XVIII
- 4 – que o sol fazia a trajectória do costume nem podia fazer outra - JangPedra XX

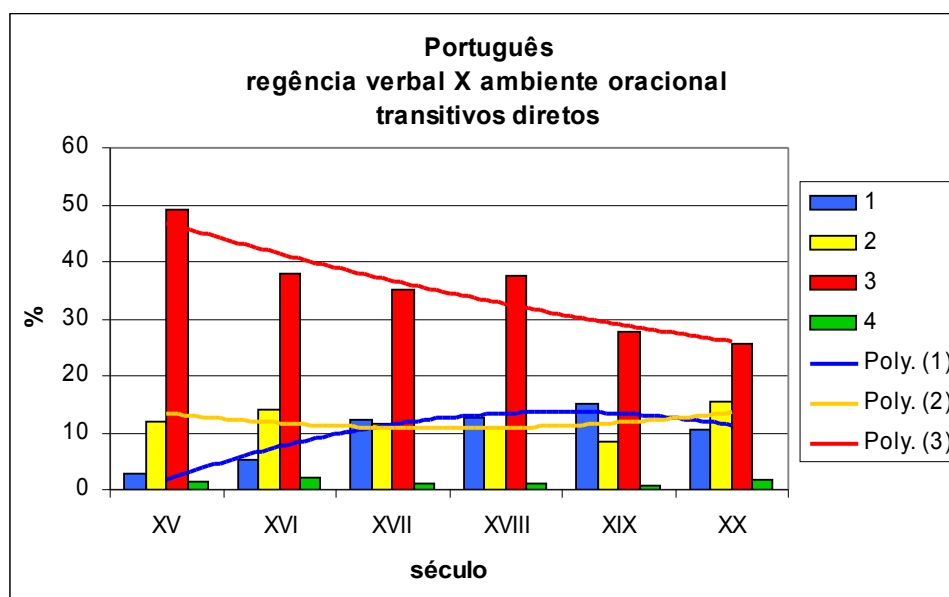


Gráfico P13a – distribuição das ocorrências de *verbos transitivos diretos* segundo o *ambiente oracional*.

Os verbos de ligação (P13b) e intransitivos (P13c) apresentam o mesmo padrão básico já disposto no Gráfico P6; entretanto, nestes dois casos, as curvas referentes aos ambientes oracionais 1 (principais), 2 (coordenadas) e 3 (subordinadas) se cruzam, indicando que esses verbos têm um comportamento diferenciado em relação aos demais. As setas indicam, para ambos os gráficos, o ponto em que esses tipos verbais passam a ser mais frequentes em orações principais que em subordinadas: o século XVIII. Exemplos dessas classes precedem os respectivos gráficos:

- 1 – O conde de Barçelos era mais velho que nenhum delles - Ceuta XV
- 2 – Mas o mais frequente vso desta abbreuiatura, he servir - DNLião XVI
- 3 – só de parte onde as arvores ficão descubertas aos raios do sol - PeJerLobo XVII
- 4 – agora que he tempo & somos senhores porque melhor he - FernOliv XVI

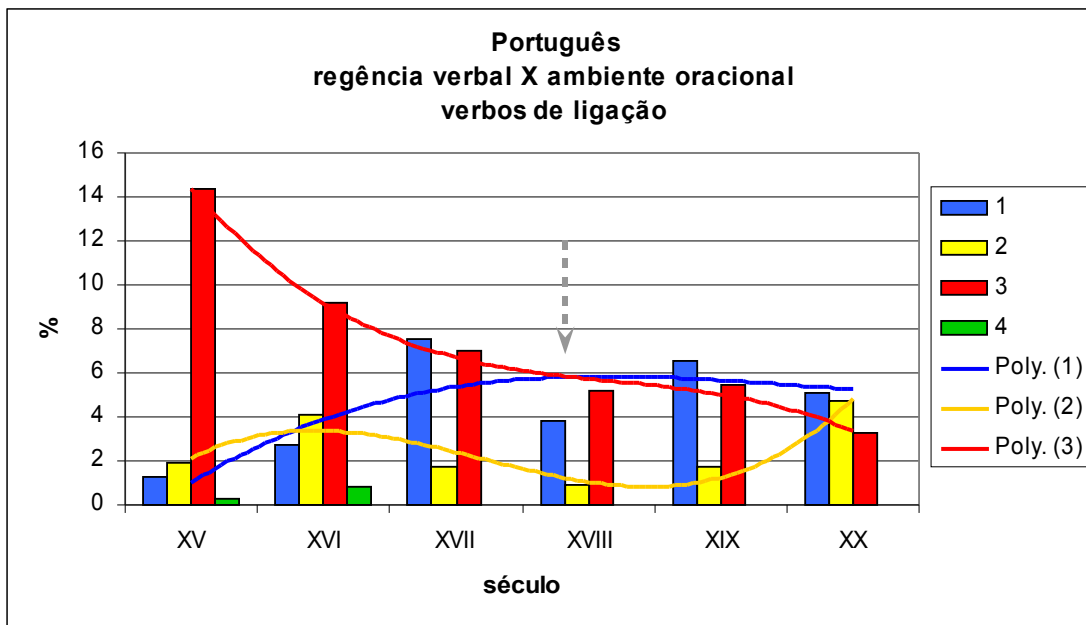


Gráfico P13b – frequências absolutas de *verbos de ligação* segundo o *ambiente oracional*.

- 1 – Trá-los Montes, estavam outros capitães e governadores - CondeDHenr XVII
- 2 – Mas bem podem ocorrer estas duas vogaes, sem formar diphthongo - DNLIão XVI
- 3 – que ... parece que descansa ali a mesma Natureza, dando repouso - IlhaAmor XVIII
- 4 – formidáveis parangonas, Nasceu a Nova Atlântida - JangPedra XX

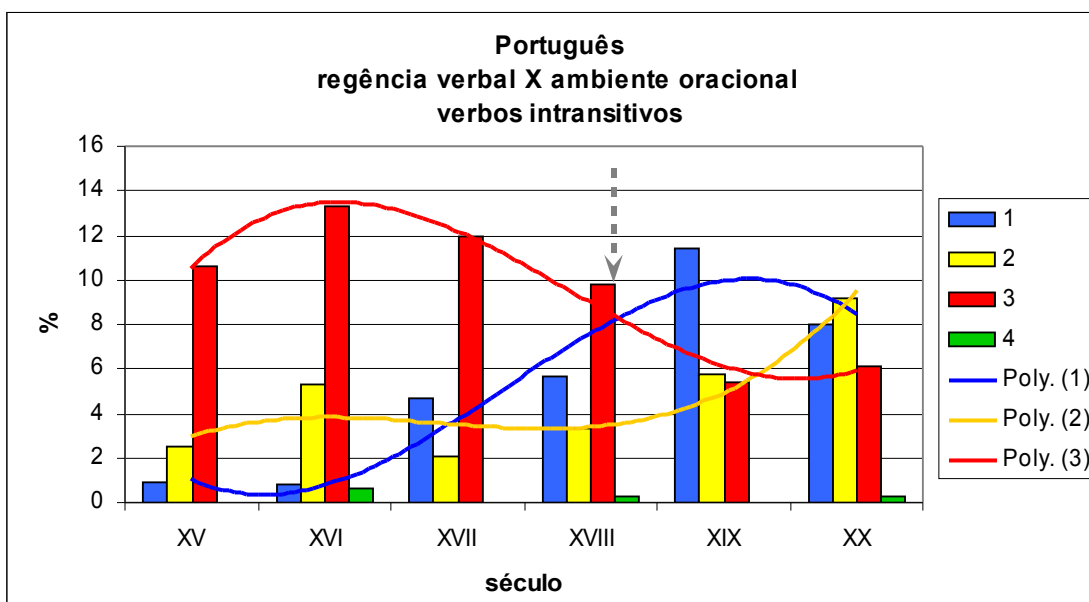


Gráfico P13c – frequências absolutas de *verbos intransitivos* segundo o *ambiente oracional*.

Outro ponto-de-vista sobre o mesmo par de fatores é dado pelo isolamento do *ambiente oracional*, permitindo-se a variação das classes de *regência do verbo* – i.e., pode-se observar como as classes de regência se comportam em ambientes oracionais diferentes. No caso das orações principais, verifica-se a predominância de verbos transitivos diretos (classe **u**), seguida de um grupo sempre crescente de intransitivos (classe **xx**) e dos verbos de ligação (classe **x**). Em termos de regressão linear, os valores obtidos para os três grupos são $iRRL(\mathbf{u1}) = 1.94$, $iRRL(\mathbf{x1}) = .77$ e $iRRL(\mathbf{xx1}) = 1.95$. Temos, como exemplos, os seguintes:

u – Esto digo segundo mynha tençom - LealCons XV

v – A outra parte do meio serve de sepultura aos turquos e mouros - PeJerLobo XVII

w – Disse-me um Cupido que se me tinha dado desde - IlhaAmor XVIII

x – Longos annos o Ramalhete permanecera deshabitado - Maias XIX

xx – O ruido dos combates calou no alto do monte - CastFaria XIX

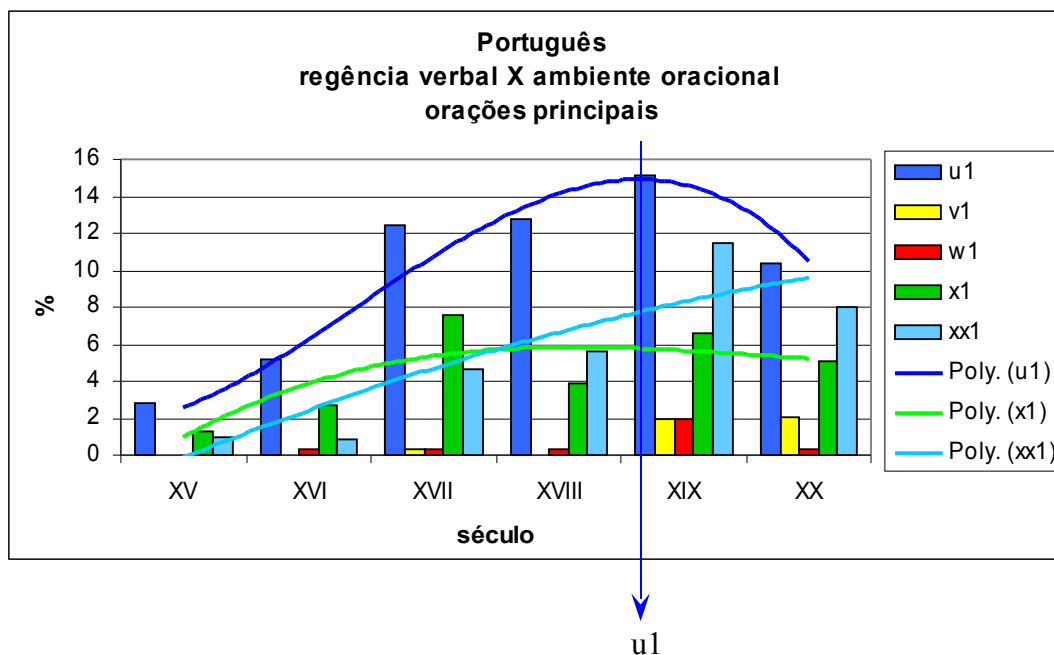


Gráfico P13c – frequências absolutas de *orações principais* segundo a *regência do verbo*.

Em orações coordenadas, as classes de regência verbal mostram ainda preferência pela classe **u** (transitivos diretos), cujas freqüências variam em torno dos 12% segundo um padrão senoidal e com RRL horizontal (inclinação = .02); o mesmo vale para os verbos de ligação (classe **x**), apesar de as freqüências absolutas desse tipo verbal serem bem mais modestas; para os intransitivos (classe **xx**), tem-se $iRRL = 1.03$. Veja-se também que o ponto de mínimo local da curva de **u2** (transitivos diretos em orações coordenadas) é praticamente coincidente com o máximo local de **u** (transitivos diretos) em orações principais (classe **1**, cf. Gráf. P13c). Nesse ambiente oracional, encontram-se exemplos como os seguintes:

- u – que as pazes foram firmadas entenderam elles que nom ficaua hy cousa - Ceuta XV
v – & ficara com maior eternidad a memoria delle - FernOliv XVI
w – e dotam ambos à mesma casa a vila de Treixede - CondeDHenr XVII
x – E o Ramalhete, continuou deshabitado. - Maias XIX
xx – Constante, tremia também o animal, como então mesmo - JangPedra XX

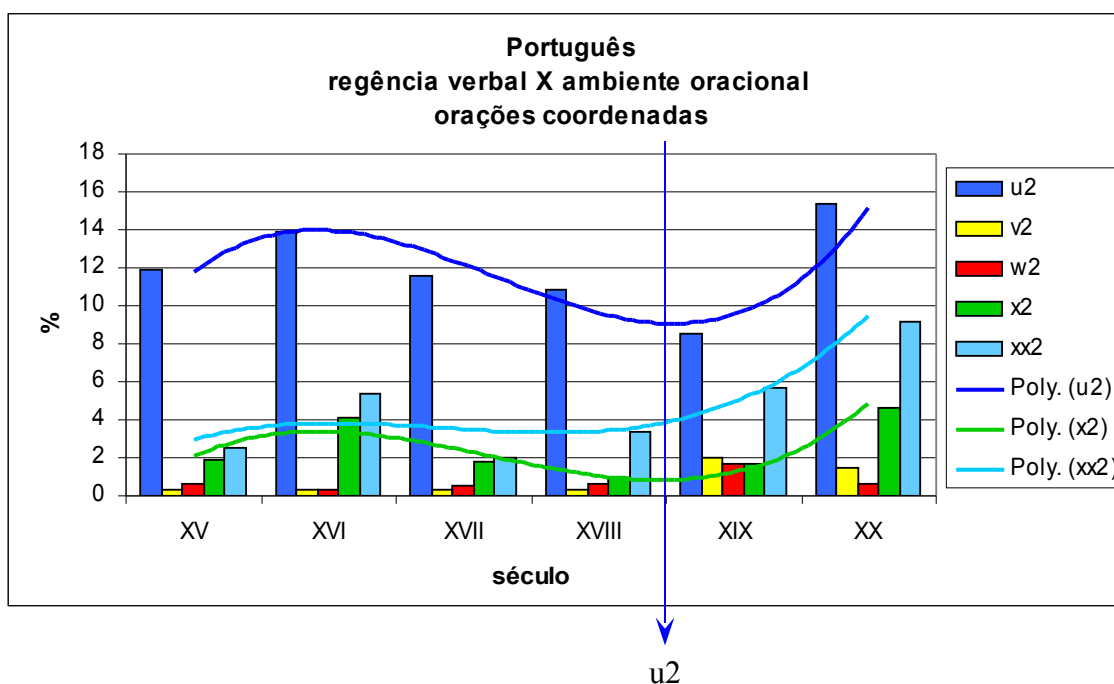


Gráfico P13d – freqüências absolutas de orações coordenadas segundo a regência do verbo.

As orações subordinadas não reservam surpresas para o comportamento das classes de regência verbal: as três categorias mais proeminentes, **u** (transitivos diretos), **xx** (intransitivos) e **x** (verbos de ligação), apresentam tendências decrescentes ($iRRL(u3) = -4.15$; $iRRL(xx3) =$

-1.37; $iRRL(x3) = -1.97$). Em todo o intervalo de amostragem, a classe **u3** domina absolutamente o gráfico, indicando frequências consistentemente mais elevadas que as das demais classes; exmplos da mesma encontram-se reproduzidos abaixo:

- u – quanto em el for, consiirando sua hidade - LealCons XV
- v – de noe & de Zubal/diz Beroso estoreador de Babilonia - FernOliv XVI
- w – Sam Jeronimo chama ao dito mar na força da lingua hebreá Sansuf - PeJerLobo XVII
- x – que depouys foy Condestavel destes reynos - PrincDJoão XVIII
- xx – de ramagens verde-garrafa, esvoaçavam cegonhas prateadas - Maias XIX

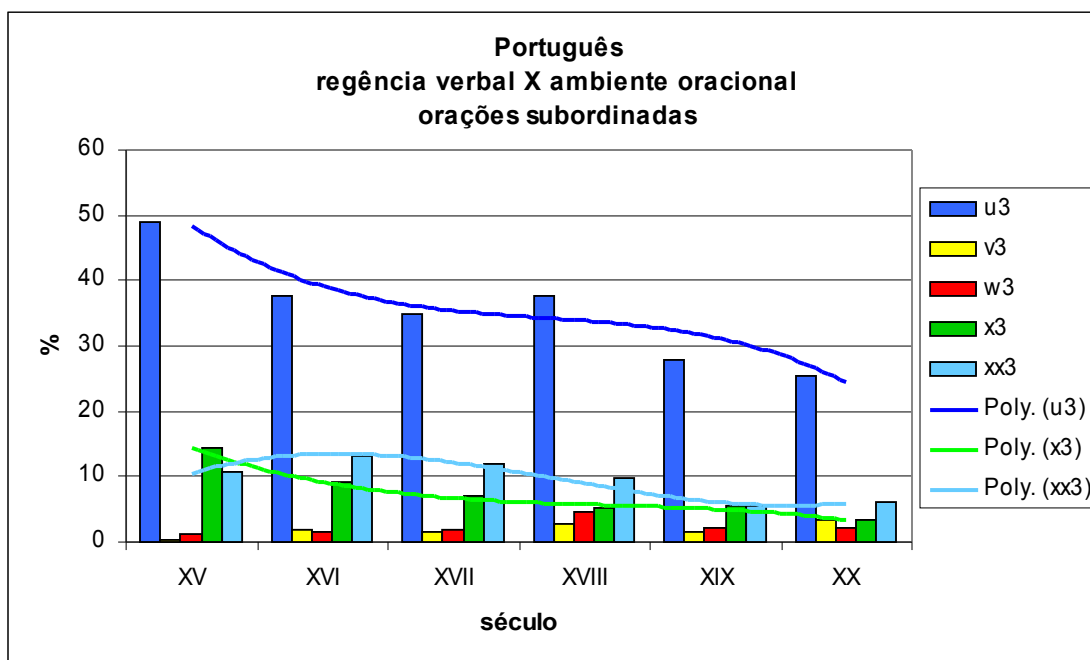


Gráfico P13e – frequências absolutas de orações subordinadas segundo a regência do verbo.

As três conjugações verbais do *corpus* português se comportam segundo padrões característicos: a primeira conjugação, **_a_**, cresce em frequência absoluta (iRRL = 3.77), enquanto a segunda cai quase 30 pontos percentuais (iRRL = -6.05); a terceira conjugação, que ocorre em proporções mais baixas que as outras duas, também verifica um aumento moderado (iRRL = 2.29). A seta indica o ponto em que, no final do século XVII, as curvas de primeira e de segunda conjugações se cruzam – i.e., o ponto de equilíbrio entre as duas classes de verbos no *corpus* português; observe-se que a frequência absoluta da classe **_e_** cai abaixo dos 50% no início do século XVII, mas a primeira conjugação só atinge esse mesmo nível no século XX. Em tempo: o argumento de que a primeira conjugação torna-se cada vez mais freqüente em decorrência da regularidade mais elevada de seu paradigma flexional indica a necessidade de se investigar a natureza dos fatores que determinaram a preponderância da segunda conjugação nas fases iniciais da amostragem.

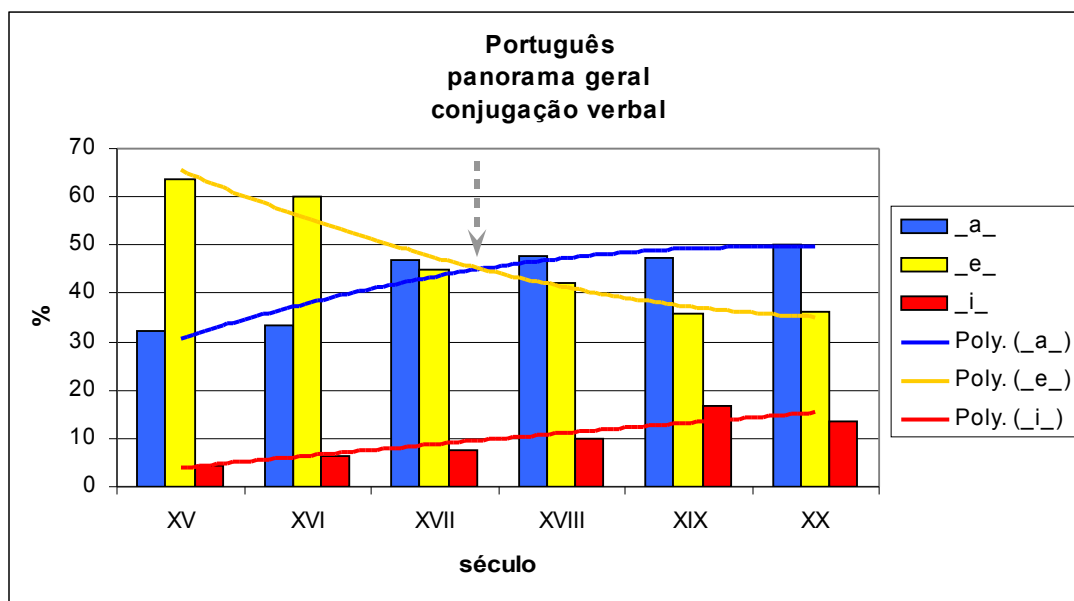


Gráfico P14a – distribuição geral das ocorrências segundo o fator *conjugação do verbo*.

A primeira conjugação se distribui preferencialmente nas ordens **a** (SVC), **i** (VC), **j** (CV), **g** (SV) e **h** (VS) – todas com perfil crescente, alinhadas com o comportamento geral delineado para essa classe no gráfico anterior: $iRRL(\mathbf{a}) = .81$; $iRRL(\mathbf{i}) = .45$; $iRRL(\mathbf{j}) = .80$; $iRRL(\mathbf{g}) = .16$; $iRRL(\mathbf{h}) = 1.54$. Também neste caso, a ordem **i** (VC) é a mais freqüente no *corpus* português; vejamos os exemplos e o gráfico de distribuição dos dados:

a (SVC) - muitas rreçoões a concrusam detreminaram de o falarem a elRey - Ceuta XV
g (SV) - possuía êste senhorio e o sustentava com grande valor - CondeDHenr XVII
h (VS) - e reynando em Mauritania os Senhores, e Pontifices de Cordova - PrincDJoão XVIII
i (VC) - do seu paletot, e confiando os pés à impermeabilidade dos seus sapatos - Serões XX

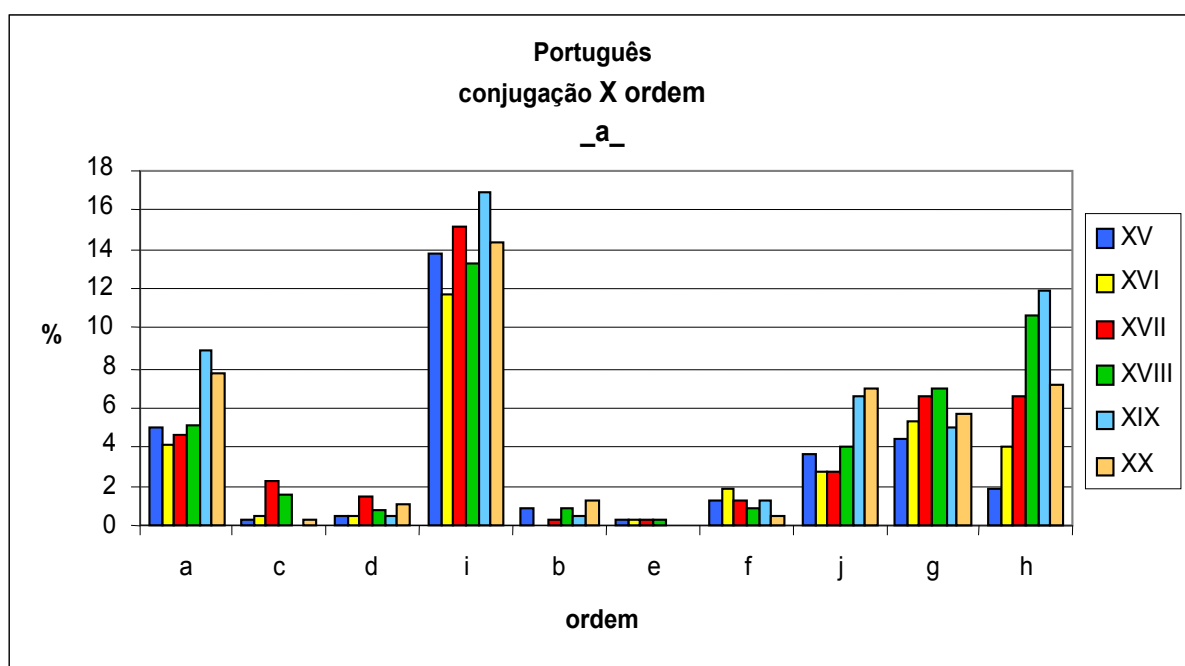


Gráfico P14b – distribuição da I conjugação verbal segundo a ordem de constituintes.

A segunda conjugação também se especializa nas ordens **a** (SVC), **i** (VC), **g** (SV) e **h** (VS), dentre as quais apenas a última contraria a tendência decrescente do grupo: $iRRL(\mathbf{a}) = -1.49$; $iRRL(\mathbf{i}) = -2.24$; $iRRL(\mathbf{g}) = -.85$; $iRRL(\mathbf{h}) = .46$. Tomem-se os exemplos:

a (SVC) - por as quaes alguus screuem, grand, & sanct - DNLião XVI

g (SV) - esteve, Granada ao sul de Madrid desde que Madrid nasceu - JangPedra XX

i (VC) - vendo-as mais yunto de nós en grande coipia i exzaminando - PeJerLobo XVII

j (CV) - clamou então : - Pois se o sabes, cumpre o teu dever - CastFaria XIX

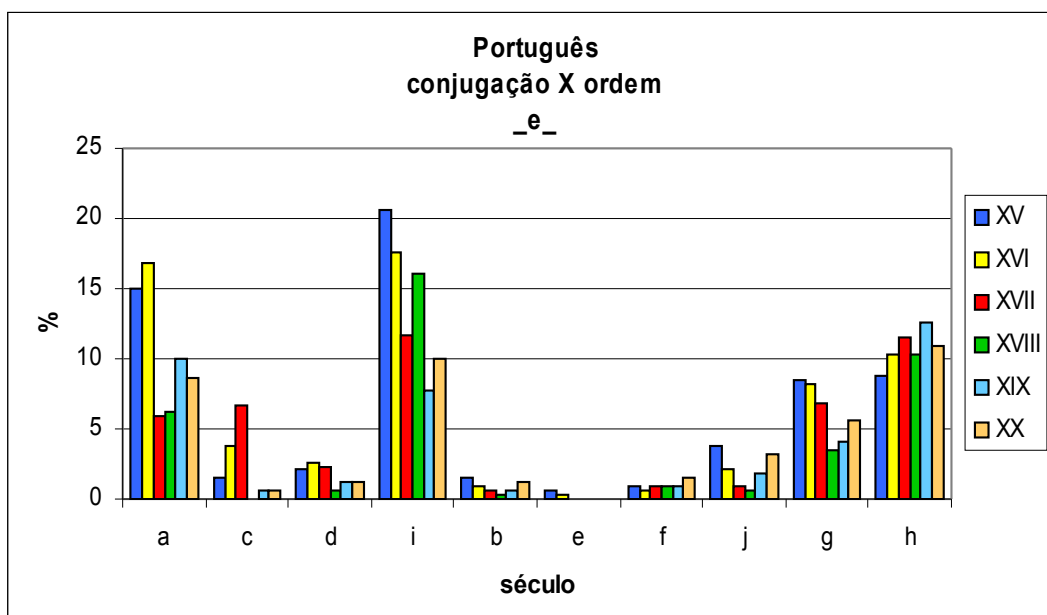


Gráfico P14c – distribuição da II conjugação verbal segundo a ordem de constituintes.

A terceira conjugação verbal ocorre em frequência bastante inferior às demais – daí a reduzida escala vertical do Gráfico P14d. Concentra-se também preferencialmente nas ordens **a** (SVC), **i** (VC), **g** (SV) e **h** (VS), como a segunda conjugação, mas exhibe a mesma tendência crescente da primeira conjugação: $iRRL(\mathbf{a}) = .43$; $iRRL(\mathbf{i}) = .71$; $iRRL(\mathbf{g}) = .16$; $iRRL(\mathbf{h}) = .72$. Os exemplos abaixo fazem parte do *corpus* português:

a (SVC) – ainda que el nom senta em sy muyta perfeiçom pera seer amado - LealCons XV

i (VC) – E por assi ser liquido este .m. & não ferir a vogal seguinte - DNLião XVI

g (SV) – tudo ri, tudo agrada, tudo parece magnífico - IlhaAmor XVIII

h (VS) – Serviram os fragmentos do castello de Faria para se construir - CastFaria XIX

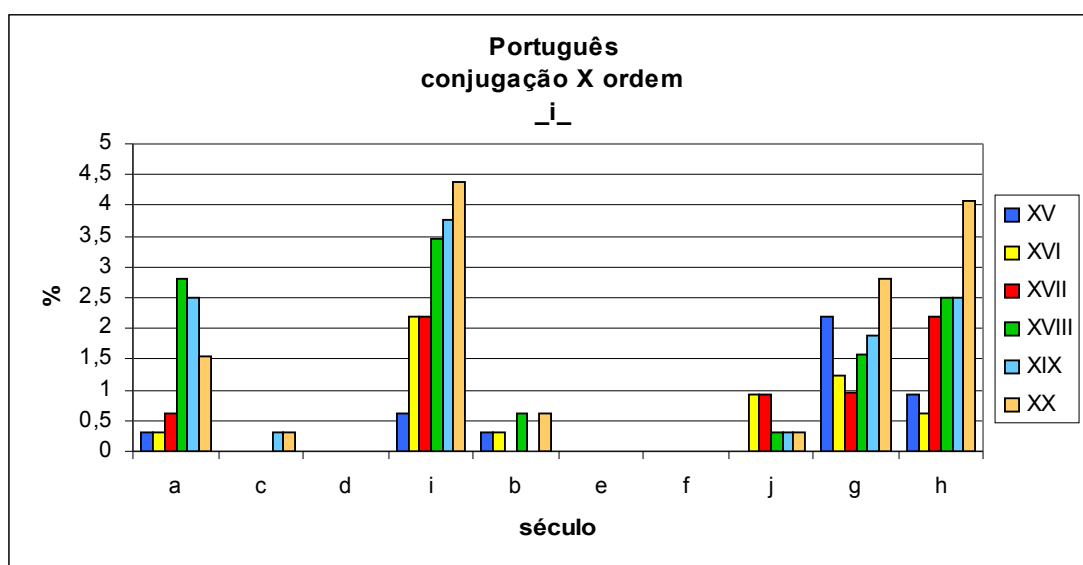


Gráfico P14d – distribuição da III conjugação verbal segundo a ordem de constituintes.

Em linhas gerais, as tendências de comportamento das ocorrências apresentadas nos Gráficos P14b-d segue um padrão mais claramente definido pela *conjugação verbal* – i.e., podem-se fazer afirmações mais abrangentes sobre a distribuição das ocorrências segundo esse fator, que tem prevalência sobre a *ordem de constituintes*. Assim é que a primeira conjugação, **_a_**, participa do *corpus* em proporções crescentes em todas as ordens mais usuais, ao passo que a segunda, **_e_**, diminui em quase todas elas (excetuando-se a ordem **h** (VS)); a terceira

conjugação, *_i_*, também mostra leve tendência positiva, embora ocorra em proporções muito baixas diante das outras duas e os resultados a ela referentes devam ser tomados com cautela.

<i>conjugação</i>	<i>a (SVC)</i>	<i>i (VC)</i>	<i>g (SV)</i>	<i>h (VS)</i>
<i>a</i>	↗	↗	↗	↗
<i>e</i>	↘	↘	↘	↗
<i>i</i>	↗	↗	→	↗

Tabela P6 – tendências de distribuição das ocorrências de *ordem de constituintes* segundo a *conjugação do verbo*.

O cruzamento dos fatores *conjugação do verbo* e *ambiente oracional* mostra, inicialmente, como cada grupo de verbos se distribui nos diversos tipos oracionais. Nesse sentido, observa-se que a curva de orações principais (classe 1) é secante à de coordenadas (classe 2); as concavidades dessas curvas são opostas, mas os dois conjuntos têm iRRLs positivas (respectivamente, 2.45 e 1.49). As orações subordinadas (classe 3) são resolvidas por um arco cujo máximo está no final do século XVII.

Para a primeira conjugação, tomem-se os exemplos:

- 1 – Do grande amor se gera huu formento no coração - LealCons XV
- 2 – e arte q ensina a bem ler & falar:saybamos quem primeiro - FernOliv XVI
- 3 – porque o acho com grandes heranças em Portugal, nas terras - CondeDHenr XVII
- 4 –que lavavam os brasões das suas armas e reformavam as librés - Serões XX

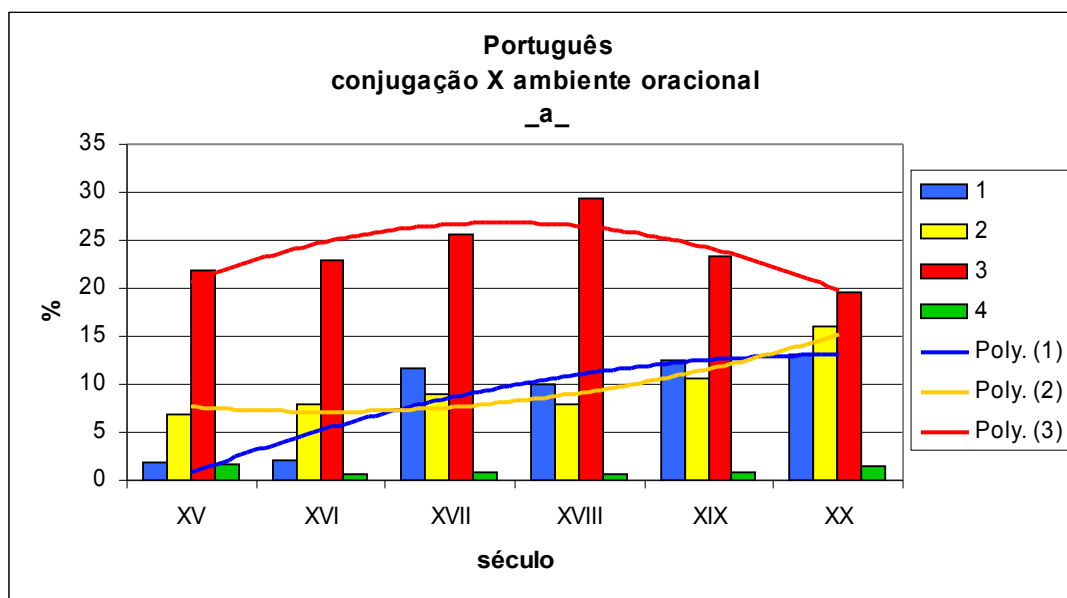


Gráfico P14e – distribuição da I conjugação verbal segundo o ambiente oracional.

Para a segunda conjugação, o perfil é diferente, como assinalado no Gráfico P14a: há um aumento moderado ($iRRL = 1.86$) das freqüências de orações principais (classe 1) associadas a essa conjugação verbal, mas o movimento mais sensível é o decréscimo ($iRRL = -7.24$) das proporções de orações subordinadas (classe 3), cujo valor é bastante próximo ao encontrado para as subordinadas no Gráfico P6 ($iRRL = -6.83$). Esse paralelismo mostra que, em termos de ambiente oracional, a queda na freqüência absoluta da segunda conjugação é determinada pelas orações subordinadas.

- 1 – Luso que também enobreceu esta terra não foy Grego - FernOliv XVI
- 2 – riu muito da phrase, e respondeu que aquellas razões eram excellentes - Maias XIX
- 3 – por ser o lugar por onde entramos a Ithiopia e ter muito que contar delle - PeJerLobo XVII
- 4 – da guerra, e que sofre mal as perdas, e affrontas, que nella recebe - PrincDJoão XVIII

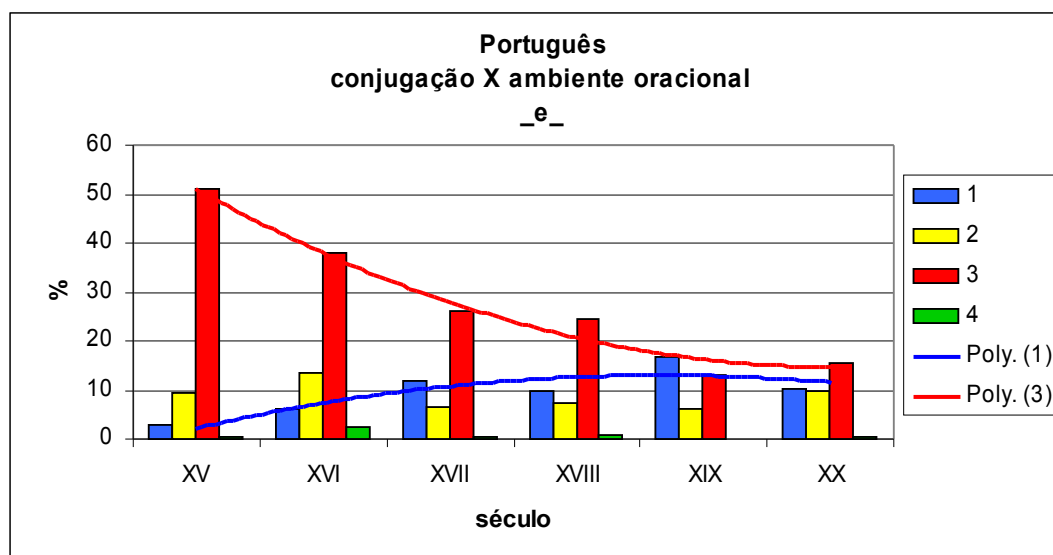


Gráfico P14f – distribuição da II conjugação verbal segundo o ambiente oracional.

Para a terceira conjugação, assinalada como _i_, temos uma extensão mais modesta da escala do eixo das ordenadas, como já visto no Gráfico P14a. Veja-se que as curvas referentes às orações principais (classe 1) e subordinadas (classe 3) têm máximos praticamente coincidentes no século XIX; contudo, o incremento ponto a ponto da curva de subordinadas é menor que o de principais. Além disso, os perfis dessas curvas mostram um decréscimo após o século XIX, enquanto a curva de coordenadas (classe 2) é ascendente no final do intervalo de amostragem. Os exemplos abaixo correspondem a ocorrências associadas a esse traço:

- 1 – Bem ouuy rrespondeo elRey o que ata ora me dissestes. - Ceuta XV
- 2 – e por este respeito saem muitos por toda a costa de Melinde - PeJerLobo XVII
- 3 – uma fita onde se distinguiam letras e numeros d'uma data - Maias XIX
- 4 – da espanha & seguilos ate a sua ytalia - FernOliv XVI

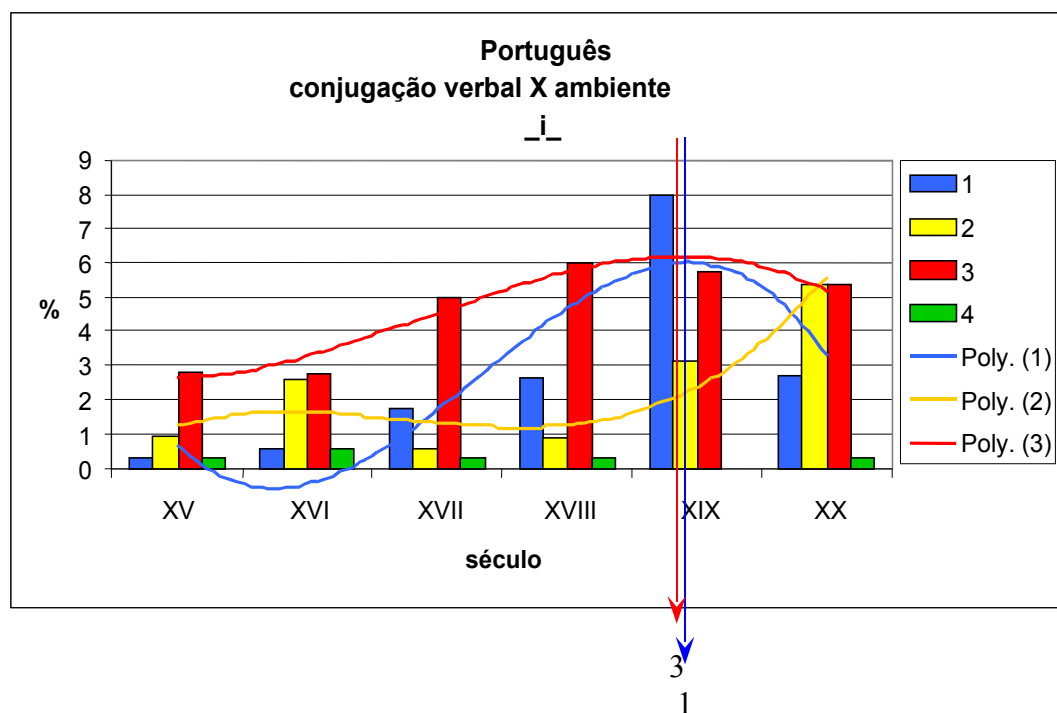


Gráfico P14g – distribuição da III conjugação verbal segundo o ambiente oracional.

Outra possibilidade de se cruzarem os fatores *conjugação verbal* e *ambiente oracional* implica deixar livre o primeiro e fixar o segundo: dessa forma, pode-se apreciar como as conjugações se distribuem em cada ambiente oracional. No caso das orações principais (classe 1), os traçados das curvas correspondentes às três conjugações são semelhantes, ainda que tenham alturas (i.e., freqüências absolutas) diferentes; além disso, apesar de os máximos dessas curvas estarem todos compreendidos no século XIX, há alguma defasagem entre eles. Para as orações principais, o subconjunto de ocorrências que tem a inclinação mais elevada é o de verbos de primeira conjugação ($iRRL(_a_)= 2.45$), seguido dos verbos de segunda ($iRRL(_e_)= 1.86$) e dos de terceira ($iRRL(_i_)= 1.00$).

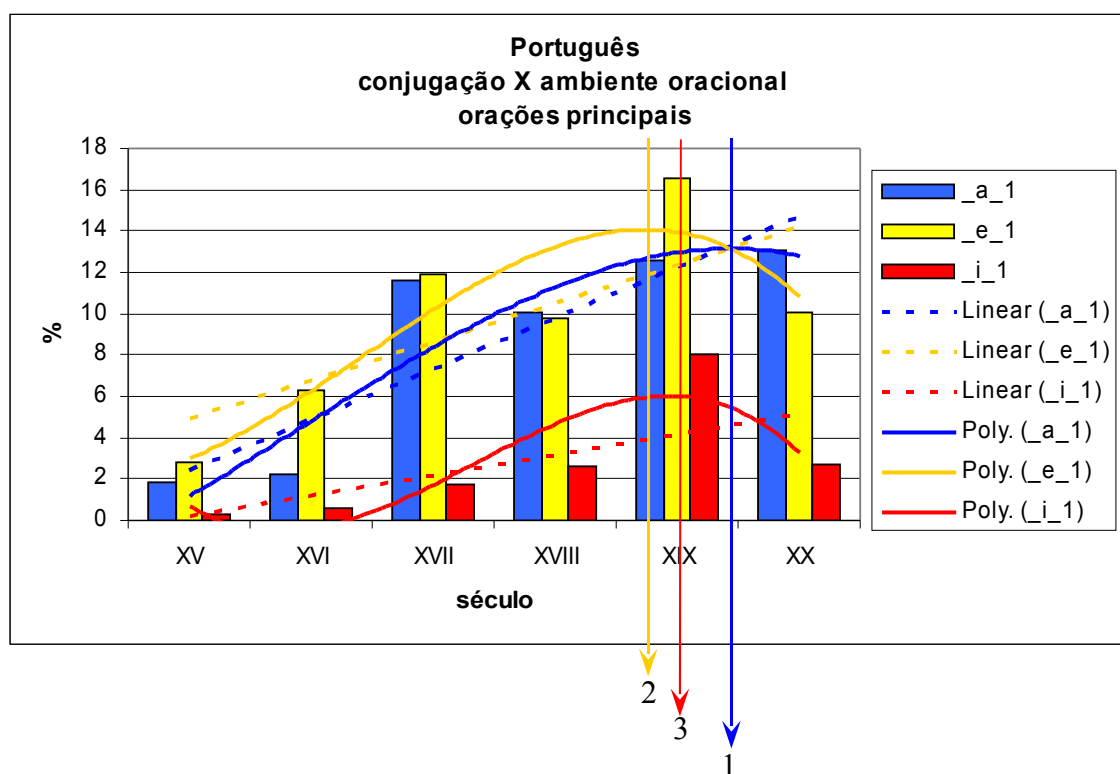


Gráfico P14h – distribuição das orações principais segundo a conjugação verbal.

As orações coordenadas (classe 2) mostram uma distribuição diferente das principais, segundo o critério da conjugação verbal: a primeira e a terceira conjugações têm retas de regressão linear de inclinação positiva ($iRRL(_a_2)= 1.49$; $iRRL(_i_2)= .69$), enquanto a freqüência da segunda conjugação cai ($iRRL(_e_2)= -.52$) ao longo do intervalo em estudo.

Observe-se que há vários pontos coincidentes entre as curvas, como os máximos locais de segunda e de terceira conjugações no século XVI, e os mínimos locais das curvas de primeira e de terceira conjugações entre os séculos XVII e XVIII.

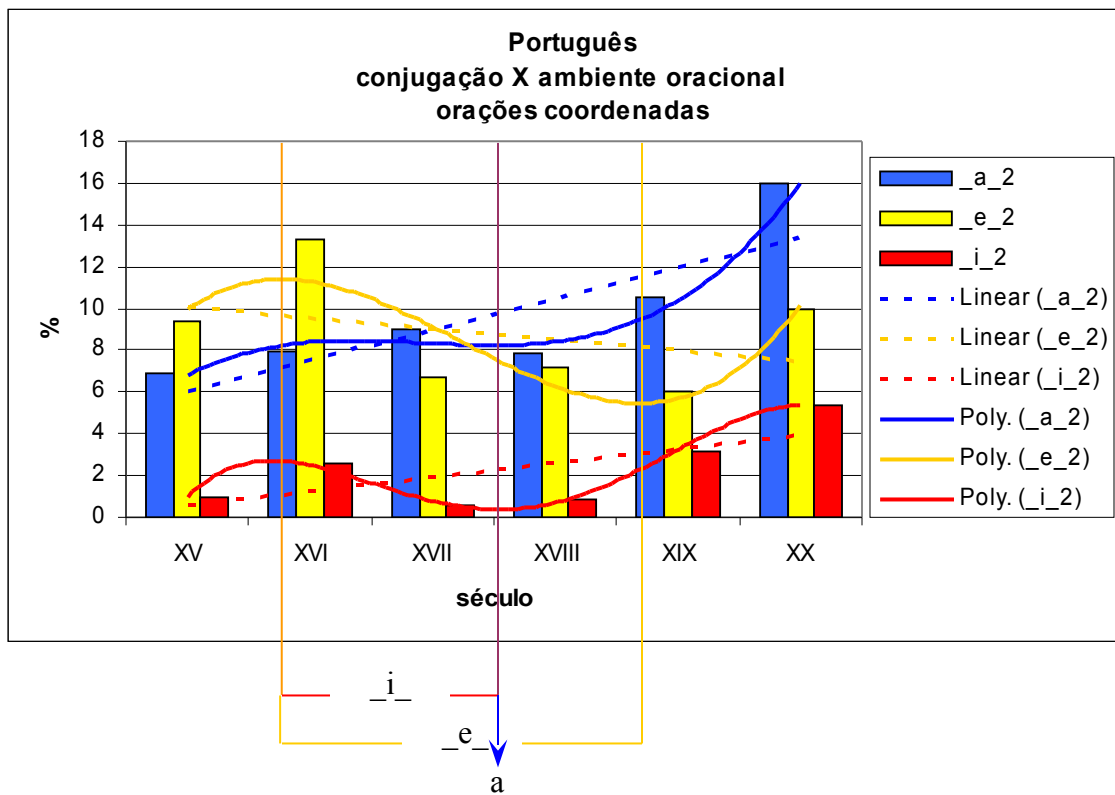


Gráfico P14i – distribuição das orações coordenadas segundo a conjugação verbal.

Para as orações subordinadas, já comentamos (q.v. Gráfico P14f) que há diminuição sensível da frequência de verbos de segunda conjugação (classe e); em particular, note-se, aqui, o paralelismo entre as retas de regressão dos subconjuntos de frequências de verbos de primeira e de terceira conjugações (respectivamente assinalados como a e i). Isso mostra que, nesse ambiente oracional, há diminuição das proporções de verbos de segunda conjugação, mas manutenção daquelas de verbos de primeira e de terceira conjugações.

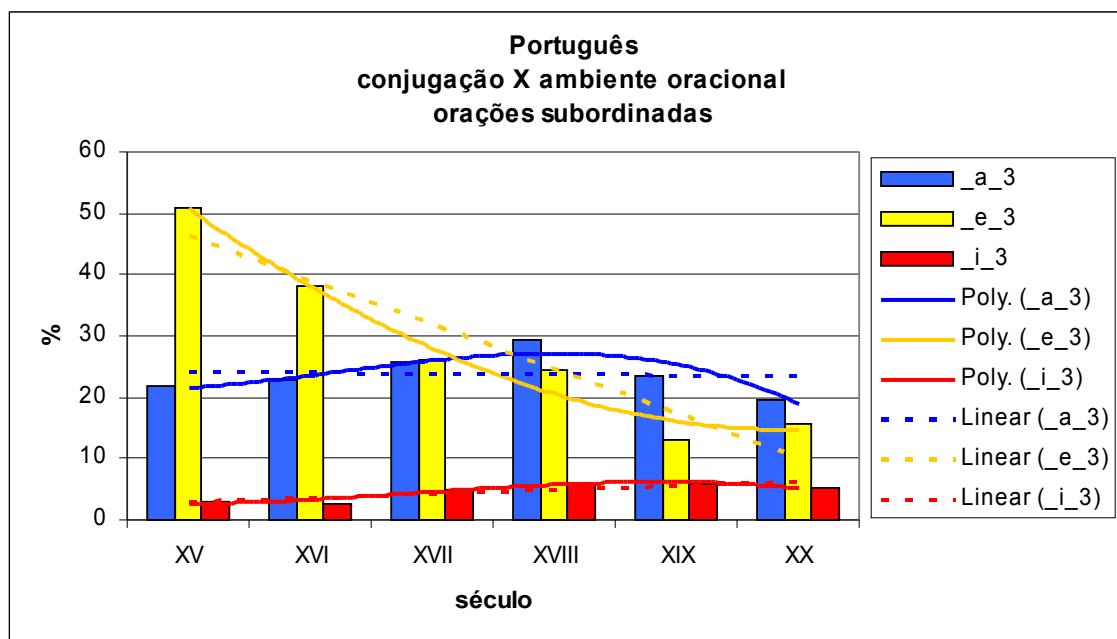
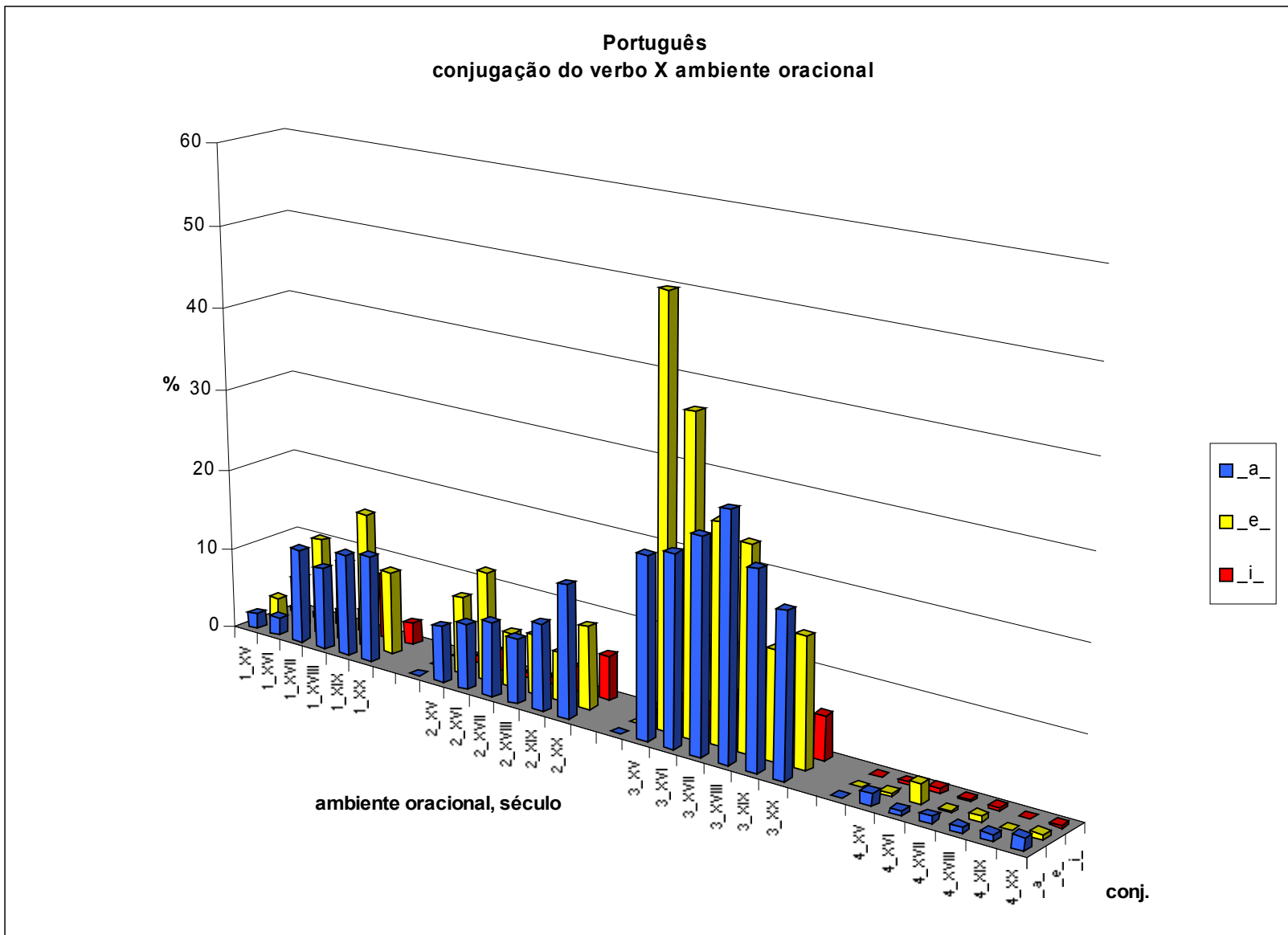


Gráfico P14j – distribuição das orações subordinadas segundo a conjugação verbal.

Também para o cruzamento dos fatores *conjugação verbal* e *ambiente oracional*, não faremos comentários sobre as freqüências em orações coordenadas em subordinação (classe 4) – que não atingem os 2.5% de freqüência absoluta no *corpus* português.

Uma vez que as escalas dos gráficos imediatamente anteriores são diferentes, apresentamos, para efeito de comparação, o Gráfico P14k, o qual sintetiza todos os dados lançados em P14h-j mais os de coordenadas em subordinação (classe 4 de *ambiente oracional*). Destaca-se nele a preponderância das orações subordinadas (classe 3), além da já mencionada franca diminuição das proporções de verbos de segunda conjugação (classe *_e_*) nesse ambiente oracional.

Gráfico P14k – comparação entre as frequências absolutas das ocorrências para os quatro ambientes oracionais e segundo as três classes de conjugação verbal.



Com respeito às frequências dos itens lexicais, temos aqui a clássica curva hiperbólica de Zipf, a qual indica que a distribuição das ocorrências encontradas nos textos portugueses obedece ao padrão usualmente associado a bases de dados textuais.

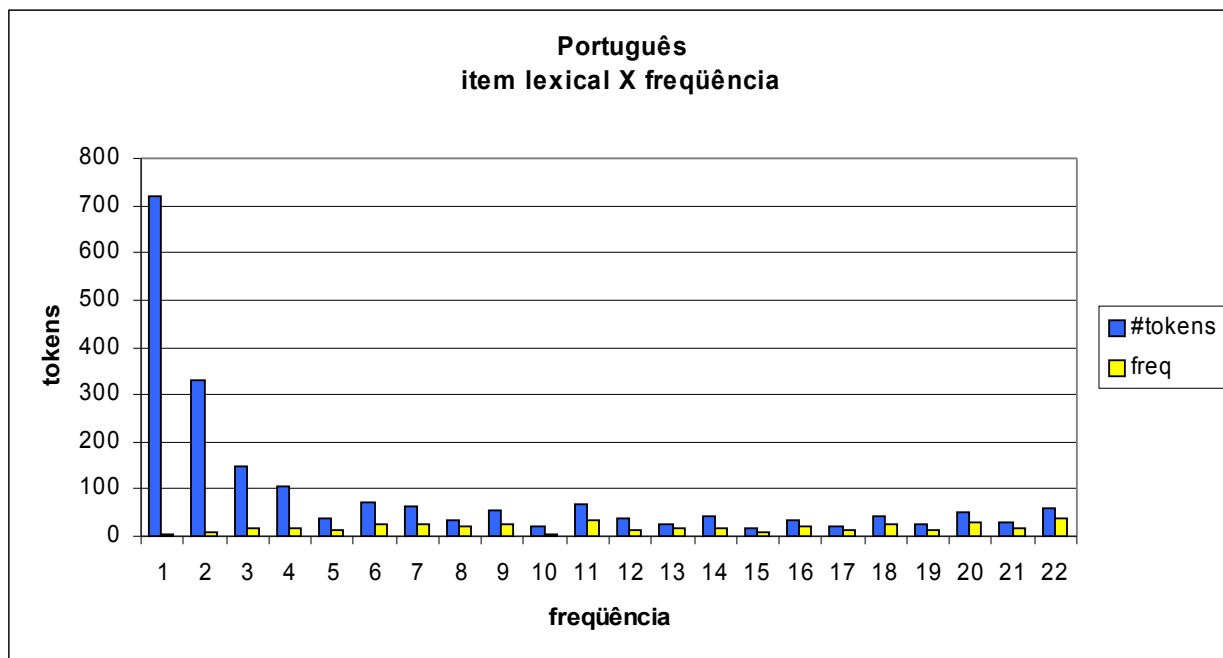


Gráfico P14l – distribuição dos verbos segundo a sua frequência no *corpus*.

O verbo mais freqüente no *corpus*, 'ser', apresenta uma distribuição de freqüências que também obedece ao padrão já delineado para o *ambiente oracional* no Gráf. P6, e reproduz, inclusive, as concavidades das curvas encontradas nesse gráfico. Esse comportamento é, naturalmente, mais claro para os itens lexicais mais freqüentes, mas se estende aos demais de maneira semelhante. Os perfis dos itens menos freqüentes sofrem, obviamente, a interferência de outros fatores, como os discursivos e estilísticos, mas o tratamento dos mesmos em conjunto (como feito acima para o fator *conjugação*) permite divisar aspectos condicionadores de seu comportamento.

As ocorrências reproduzidas abaixo compreendem o verbo *ser* nos ambientes sintáticos distintos neste trabalho:

1 – Esta erva he a mesma a que no mar chamão sargaso - PeJerLobo XVII

2 – Porem pera taaes revesada sospeita... he muyto scusada - LealCons XV

3 – porque o til, que se põe em irmão, não he sobre o .o. que he a derradeira letra - DNLião XVI

4 – agora que he tempo & somos senhores porque melhor he que ensinemos a Guine ca - FernOliv XVI

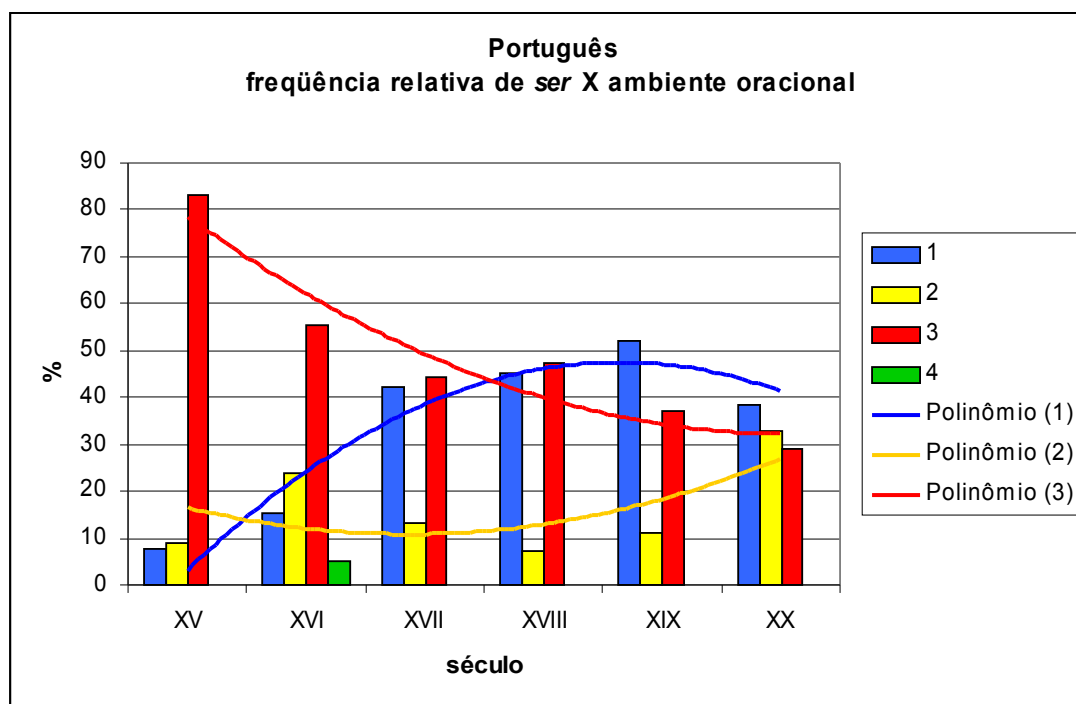


Gráfico P14m – distribuição do verbo *ser* no *corpus* português segundo o ambiente oracional.

Por fim, a distribuição das ocorrências segundo a classe semântica do verbo mostra, da mesma forma que se observou para o italiano e para o espanhol, uma concentração nas classes ATTR, COGN, DISP e SEM. Os valores de inclinação das RRLs para os subconjuntos [VC] dessas categorias mais freqüentes apontam para a diminuição das freqüências da maioria: $iRRL(ATTR) = -1.22$; $iRRL(COGN) = -.11$; $iRRL(DISP) = .99$; $iRRL(SEM) = -.54$.

Tomem-se os exemplos abaixo, seguidos do gráfico P14n, que sumariza o comportamento do conjunto:

- ATTR - Como .a que he vogal:& .b. que he cõsoante: & nam tẽ voz ao menos -
FernOliv XVI
- COGN - O mesmo se colhe de outras muitas escrituras - CondeDHenr XVII
- DISP - as contas, se despenderaõ cento e quinze mil dobras, gasto taõ moderado -
PrincDJoão XVIII
- SEM - É tempo de dizermos o motivo de tanta e tão excepcional agitação - Serões XX

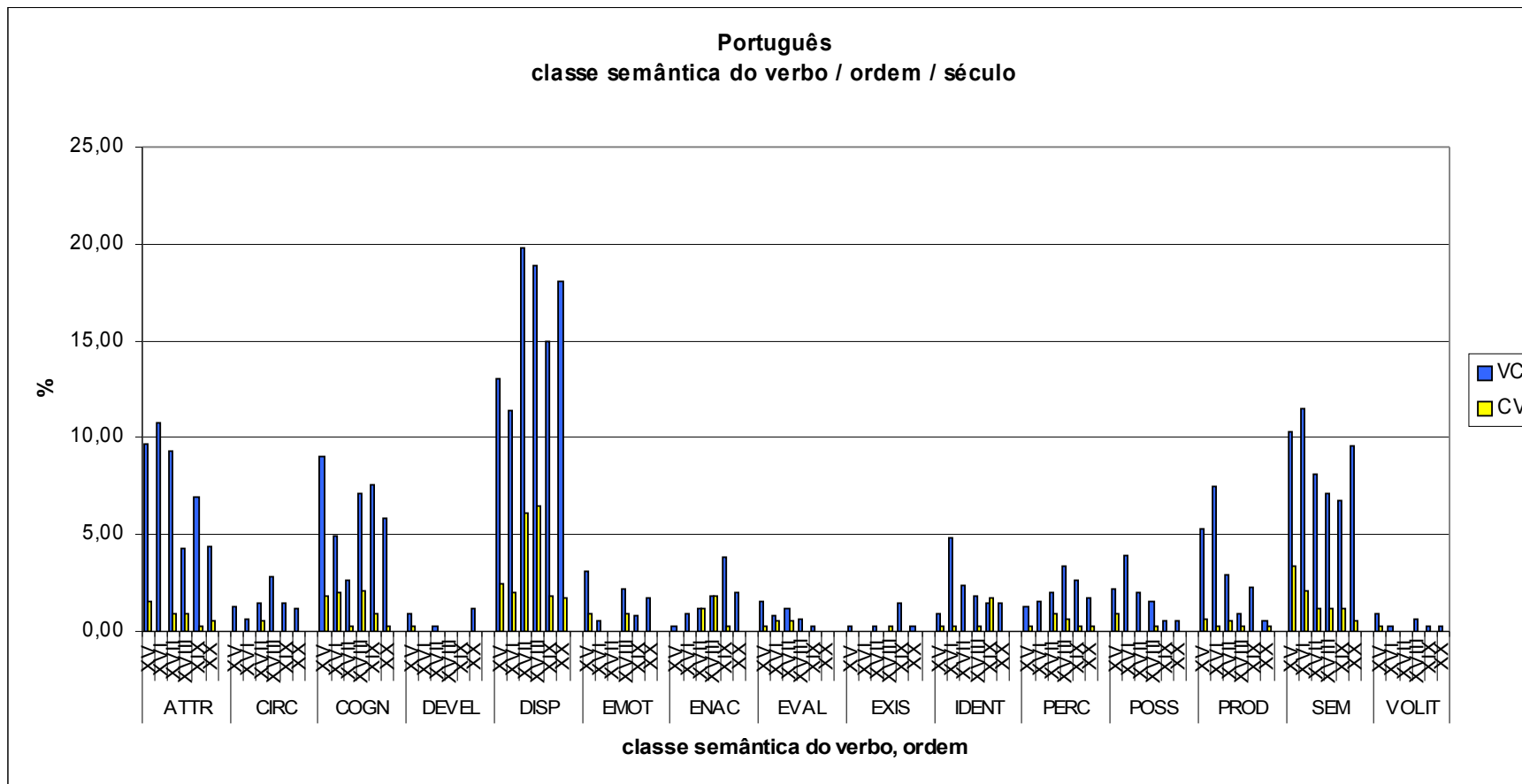


Gráfico P14n – distribuição das ocorrências segundo a *classe semântica do verbo*.

Em suma, os dados do *corpus* português indicam que alguns fatores são dependentes de outros, os quais lhes determinam o perfil de distribuição ao longo do tempo. Assim é que, por exemplo, a própria *ordem de constituintes* não é apenas função do tempo (cf. gráfico P1), como também do *peso de S e de C* (cf. gráficos P7 e tabela P4) e do *ambiente oracional*, AO (gráficos P6a-c e tabela P3). Da mesma forma, funções da *ordem de constituintes* e, por conseguinte, do *ambiente oracional* são os traços do verbo – *tempo* (cf. gráficos P10a-c e tabela P5, para a *ordem*, e a série de gráficos P11, para o fator AO), *regência* (cf. gráficos P12, para a *ordem*, e os gráficos P13, para o *ambiente oracional*) e *conjugação* (cf. gráficos P14b-d e tabela P6, para a *ordem de constituintes*, e os gráficos P14e-k, para o fator AO) – e a *agentividade* (cf. tabela P1, para a *ordem*, e gráficos P4a, P5a e tabela P2, para o *ambiente oracional*). Demonstra-se também que os perfis dos fatores *peso de S e de C* e *item lexical* (avaliada apenas para o verbo mais freqüente, *ser*) são dependentes de AO nos gráficos P8-P9 e P14m, respectivamente; em P14n, por fim, observa-se que a *classe semântica do verbo* também está ligada à *ordem de constituintes*. Dentre todos os fatores elencados neste trabalho, o único que se mostra exclusivamente dependente do tempo é o *ambiente oracional*; como se depreende da análise dos dados, há mudança nas proporções relativas de vários traços, de maneira a obedecer à alteração nas freqüências de ocorrência dos ambientes oracionais. Esse é, portanto, um indício de que, no *corpus* português, vários movimentos de mudança são controlados pelo *ambiente oracional*.

Tratamento numérico das ocorrências: o *corpus* espanhol

Os vários recortes feitos na base de dados de espanhol a partir da consideração dos fatores de análise levam à obtenção de perfis de comportamento do *corpus*, apresentados nos gráficos e nas tabelas seguintes. Os fatores são tratados, de princípio, isoladamente e em relação à ordem dos constituintes, e, em seguida, combinados com o critério *ambiente oracional* para observar sua evolução em diferentes tipos oracionais. Os comentários que acompanham os gráficos serão sumarizados, posteriormente, na seção de comparações entre os *corpora*.

As ocorrências de espanhol de distribuem preferencialmente nas ordens **a** (SVC), **i** (VC), **j** (CV), **g** (SV) e **h** (VS), com predominância da ordem **i** (VC). Os comportamentos de algumas dessas ordens são bem definidos em termos de tendências, como **a** (SVC) e **h** (VS), cujas proporções aumentam (iRRL = .58 e 1.17, respectivamente), e **i** (VC) e **j** (CV), em decréscimo (iRRL = -.57 e -.96, respectivamente).

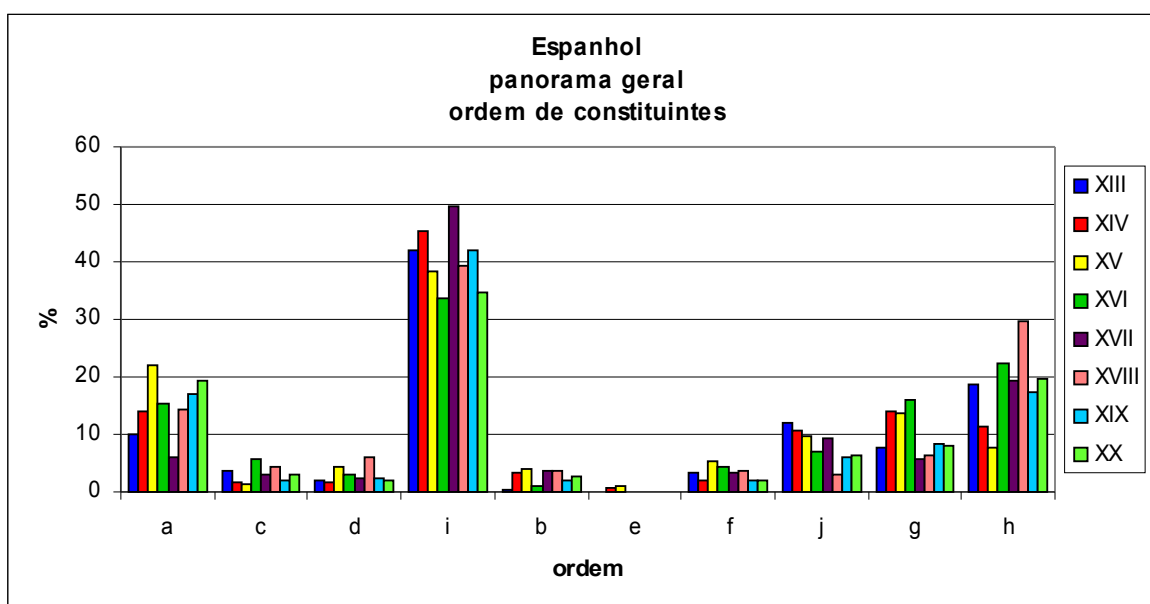


Gráfico E1 – tendências gerais de comportamento do *corpus* espanhol para a combinação dos traços *ordem* e *agentividade de S e de C*.

Com respeito à *agentividade de S e de C*, observa-se, em primeiro lugar, a concentração das ocorrências em torno das classes prototípicas Ah e Pn nas ordens [VC].

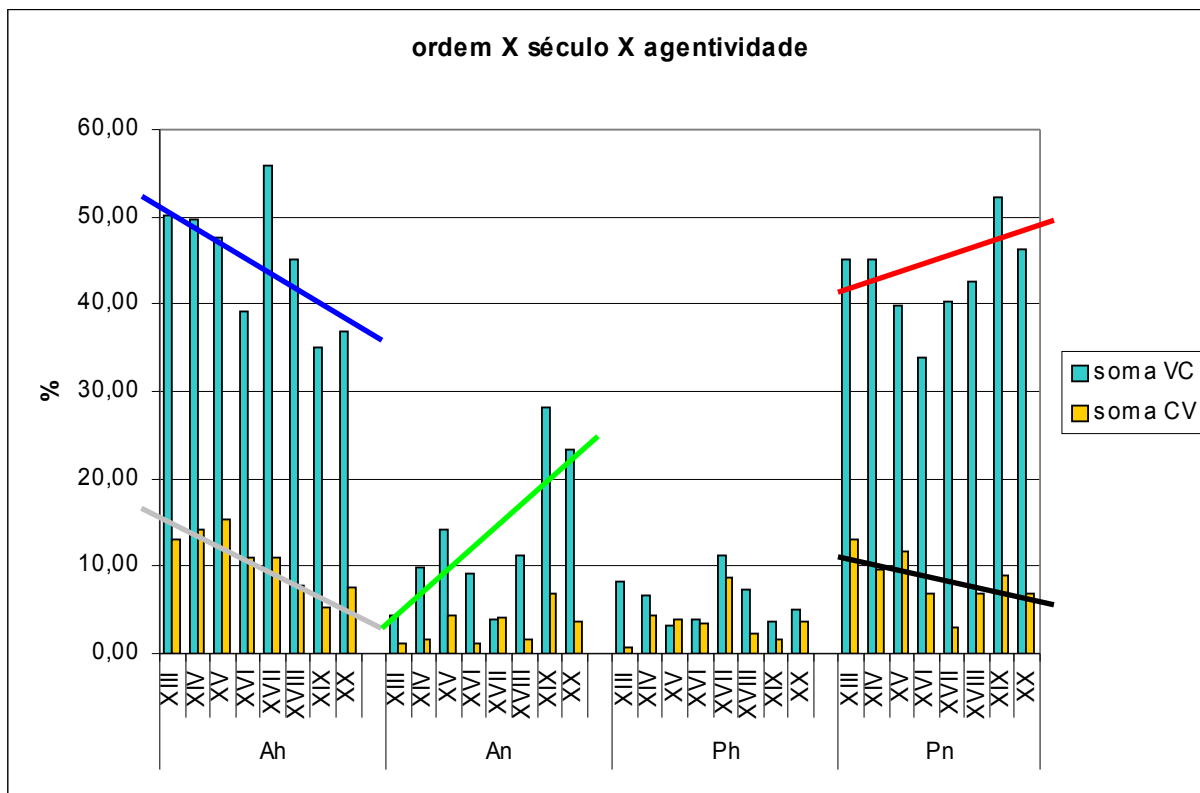


Gráfico E1 – tendências gerais de comportamento do *corpus* espanhol para a combinação dos traços *ordem* e *agentividade de S e de C*.

Em termos gerais, o cruzamento dos fatores *ordem de constituintes* e *agentividade* aponta para o fato de que as classes Ah (VC), Pn (VC) e An (VC), as mais abundantes, têm comportamentos diferentes: as proporções relativas de Ah diminuem (iRRL = -1,87), enquanto as de Pn e de An tendem a aumentar: iRRL (Pn) = 0,69 e iRRL (An) = 2,50. Some-se a isso a redução das freqüências de Ah (CV) (iRRL = -1,24) e de Pn (CV) (iRRL = -0,80), e tem-se o indicativo de uma possível tendência do sistema, condicionada pela posição relativa do complemento verbal: se a ordem é [VC], o traço [+humano] tende a diminuir em freqüência, enquanto [-humano] tende a aumentar; se o complemento é pré-verbal, o agente [+humano] e o paciente [-humano] têm suas freqüências reduzidas.

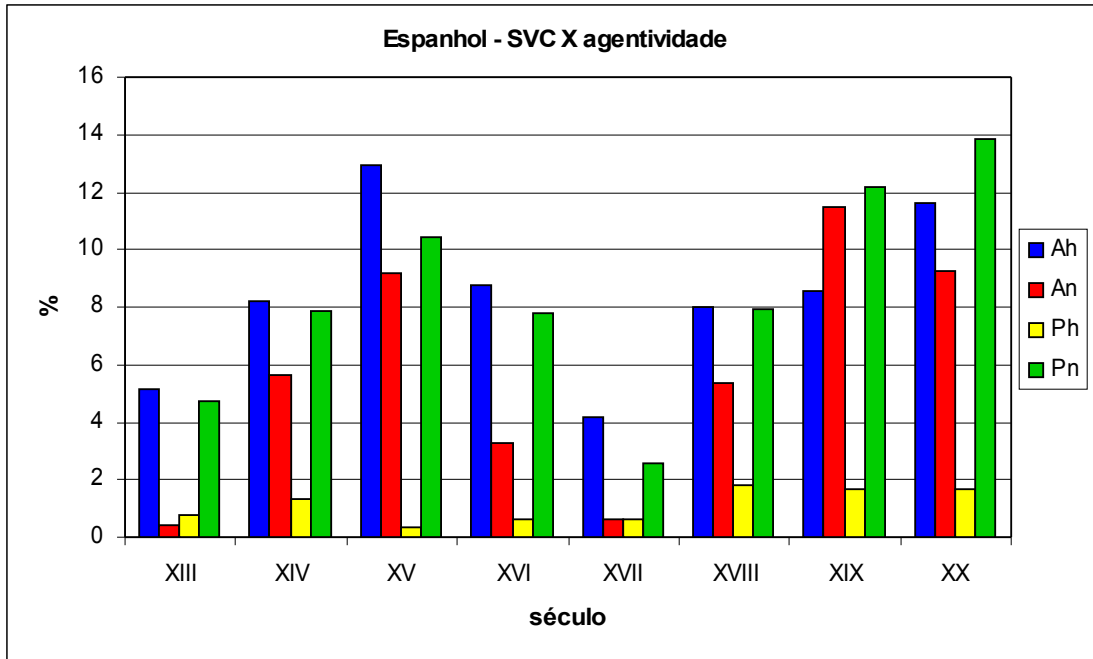


Gráfico E2a – distribuição das ocorrências de ordem a (SVC) segundo a agentividade de S e de C.

O cruzamento do traço SVC com os de agentividade identificados para o *corpus* espanhol mostra uma flutuação senoidal em todo o espaço de variação destes últimos. Esse comportamento é mais patente nas combinações mais frequentes (SVC-Ah e SVC-Pn), mas é também sugerido nas menos numerosas. Os perfis estão sempre abaixo dos 15% totais, o que torna o grupo SVC um dos mais expressivos de todo o conjunto. Veja-se ainda que as associações mais frequentes são justamente aquelas em que haveria maior coerência na atribuição de papéis: agente tende a ser associado ao traço humano, e paciente a referentes não-humanos. São exemplos:

Ah – En cuanto a los peligros, yo no negaré que esta guerra (PAZ)

An – ningún influjo derribará sus intenciones de hacer la paz (PAZ)

Ph – de commo los troyanos lo dieran sus donas que oferesçiese (TRO)

Pn – vuestra gente destroyeron ya aquella tierra toda otra vegada (TRO)

Apesar do perfil flutuante apresentado no gráfico E2a, a resolução de todas as tendências nele sumarizadas corresponde a retas de inclinação positiva, como pode ser observado no gráfico abaixo:

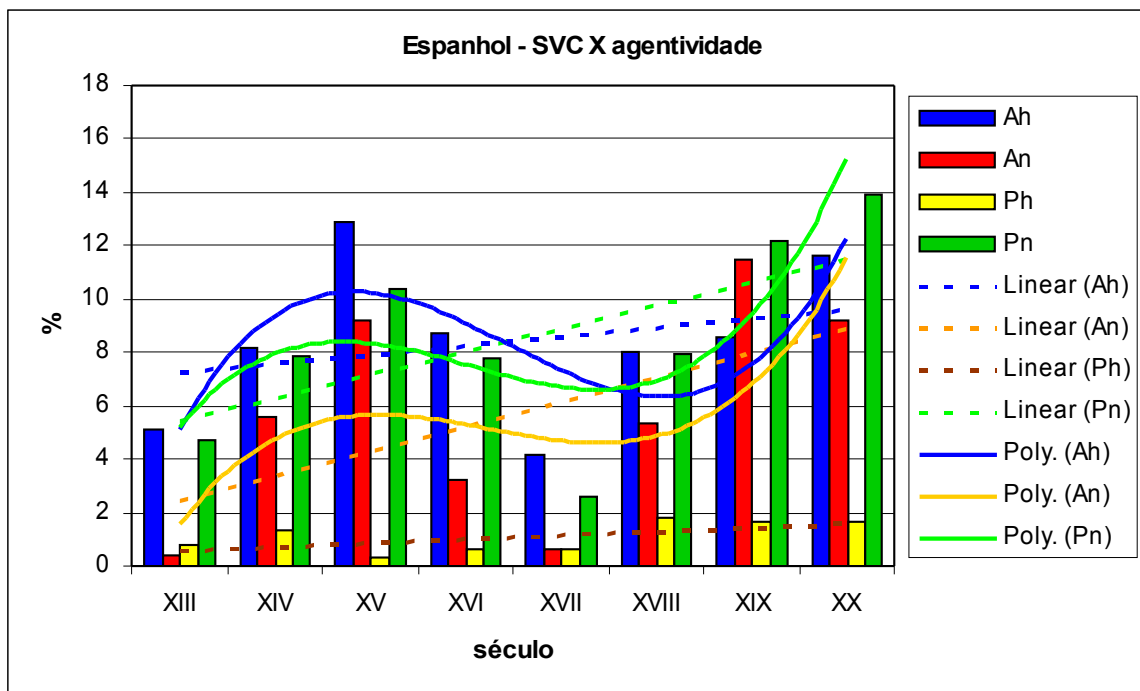


Gráfico E2b – tendências de comportamento das ocorrências de ordem a (SVC) segundo a agentividade de S e de C.

O fato de se terem retas ascendentes indica, em linhas gerais, o aumento da preferência pela ordem básica **a** (SVC). As subclasses temáticas que apresentam maior expressividade na mudança, conforme os resultados oferecidos pela amostra, são An e Pn, relacionadas às retas que têm mais elevadas taxas de incremento. Em termos relativos, é interessante notar que as proporções de Pn nos textos adotados ultrapassaram as de Ah no início do século XVIII, e que as curvas e RRLs também indicam o possível cruzamento das classes Ph e Ah de maneira a refletir o avanço da primeira sobre a última.

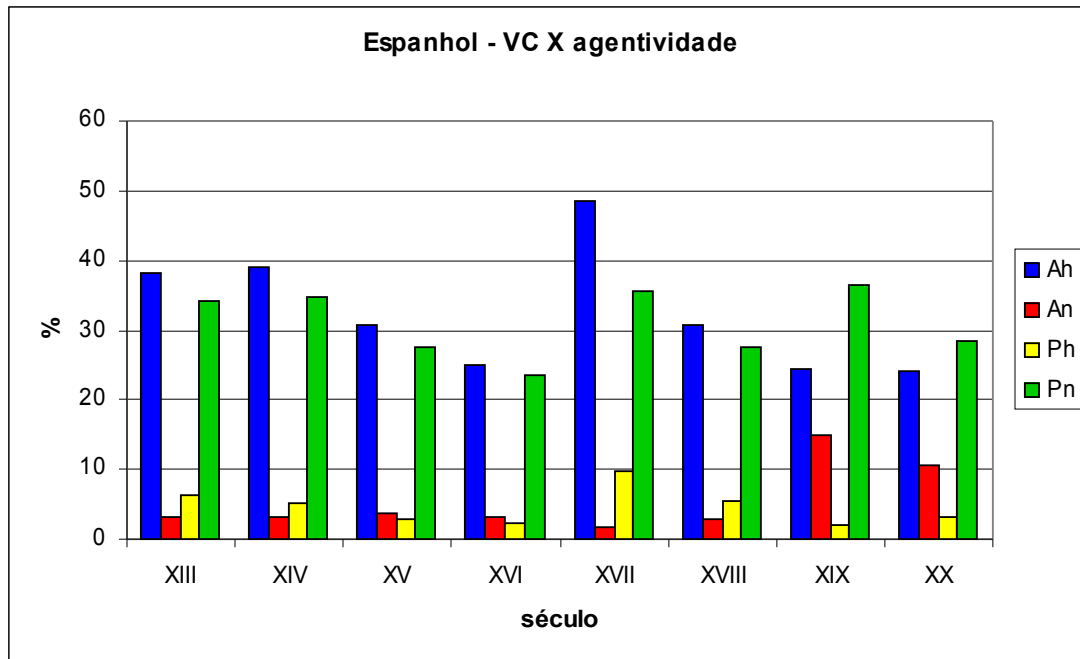


Gráfico E3a – distribuição das ocorrências de ordem i (VC) segundo a agentividade de S e de C.

Situação diversa é a que se apresenta para a combinação dos traços i (VC) e agentividade: o perfil das subclasses mais freqüentes, Ah e Pn, é descontínuo, mostrando decréscimo das proporções de VC até o século XVI, seguido de aumento brusco no século XVII e decréscimo ou flutuação até o final do intervalo de amostragem. Não obstante todos os valores estarem compreendidos abaixo dos 50% de freqüência absoluta, esse é o traço mais freqüente no corpus espanhol. Os exemplos abaixo ilustram a combinação dos traços:

Ah – A ésta presenta tu obra, y sabes ende qué ganarás (MUG)

An – Lo primero que le llevó los ojos fué el cielo (AGR)

Ph – entre aquéstas no olvidaré las virtuosas mugeres de los de Menia (MUG)

Pn – y concibiendo un horror implacable al pecado (AGR)

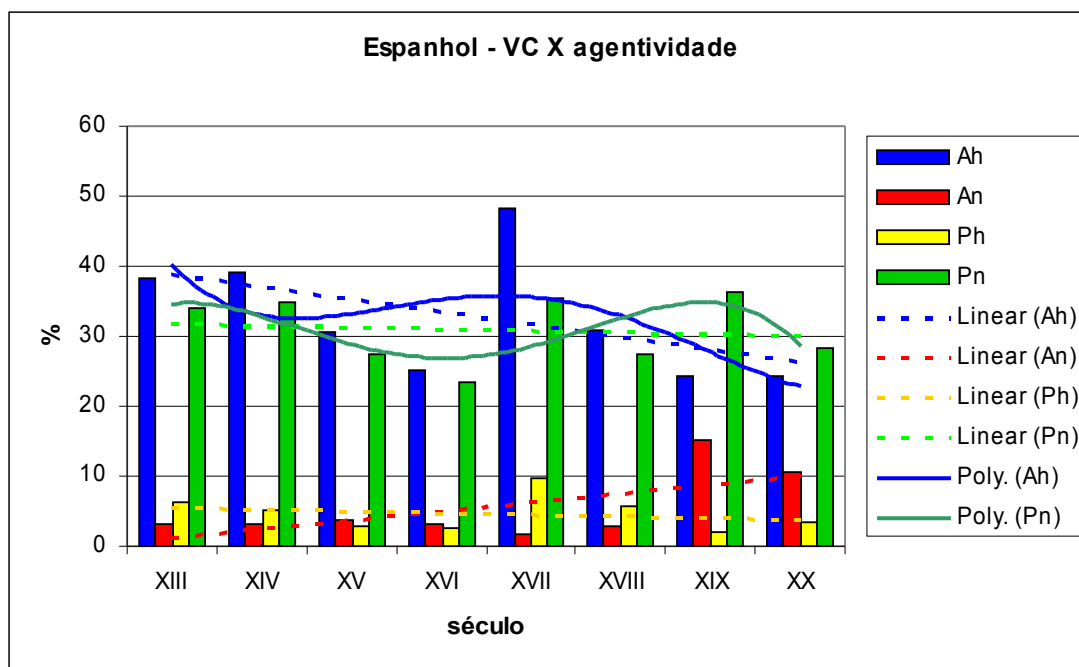


Gráfico E3b – tendências de comportamento das ocorrências de ordem i (VC) segundo a agentividade de S e de C.

A regressão linear dos perfis apresentados no gráfico E3b fornece resultados interessantes. Em primeiro lugar, nota-se que os conjuntos relativos ao traço P (tanto **h** quanto **n**) são praticamente estáveis, oscilando entre valores cuja diferença máxima está em torno dos 15%; em segundo, a reta correspondente à combinação An tem a mesma inclinação de sua semelhante para o subconjunto **a** (SVC) (q.v. gráfico E2b acima); por fim, observa-se facilmente que a tendência geral para Ah é decrescente, ao contrário do que se dá para os dados associados ao traço SVC.

O isolamento da classe **i** (CV) permitindo-se a variação do traço *agentividade* apontou, para o *corpus* espanhol, comportamentos semelhantes para os pares de combinações a) [CV-Ah] e [CV-Pn] e b) [CV-An] e [CV-Ph]: o primeiro par se apresenta em proporções relativas cada vez menores no sentido positivo do eixo que representa o tempo, ao passo que o segundo tende a uma leve flutuação ascendente. Ilustre-se a combinação desses traços:

Ah – e guisad commo nos acorrades con vino e con farina e con carne (TRO)

An – una es la luz que los ilumina (CAS)

Ph – los que lo salieron a rresçebir (TRO)

Pn – a las casetas, la concurrencia que las usa y llena (CAS)

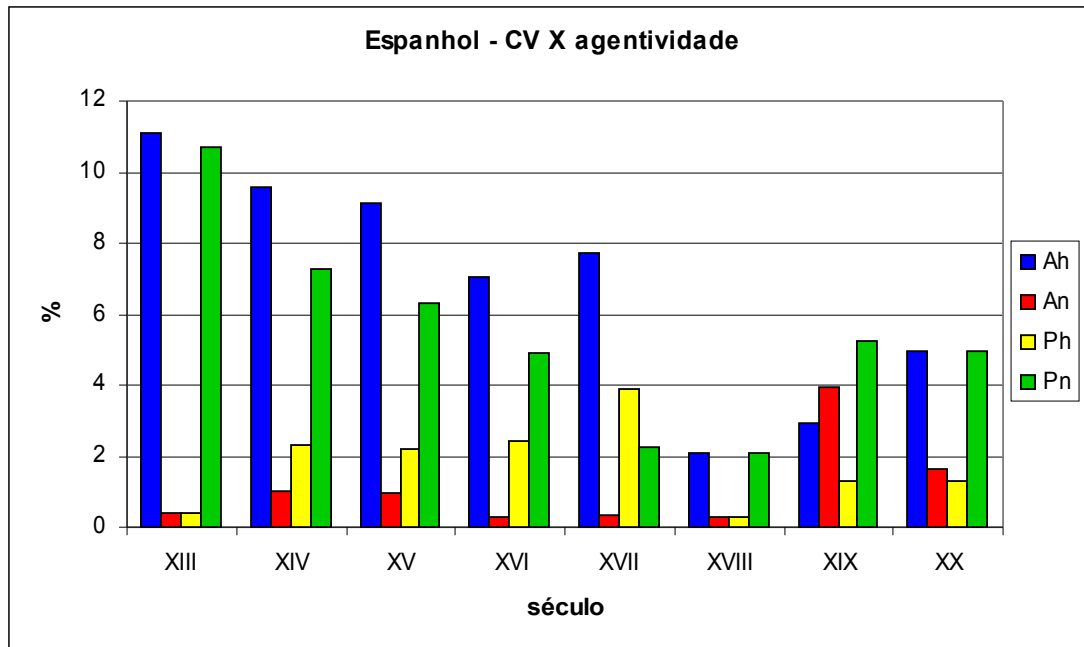


Gráfico E4a – distribuição das ocorrências de ordem j (CV) segundo a agentividade de S e de C.

As linhas de tendência para as quatro subcategorias confirmam o exposto acima; observe-se, contudo, que o declínio de Ah é um pouco mais acentuado que o da preferência por Pn, ao passo que há um leve aumento (da ordem de 2%) nas proporções relativas de An. Ph, por sua vez, oscila timidamente, num movimento cuja amplitude máxima não chega aos 4%.

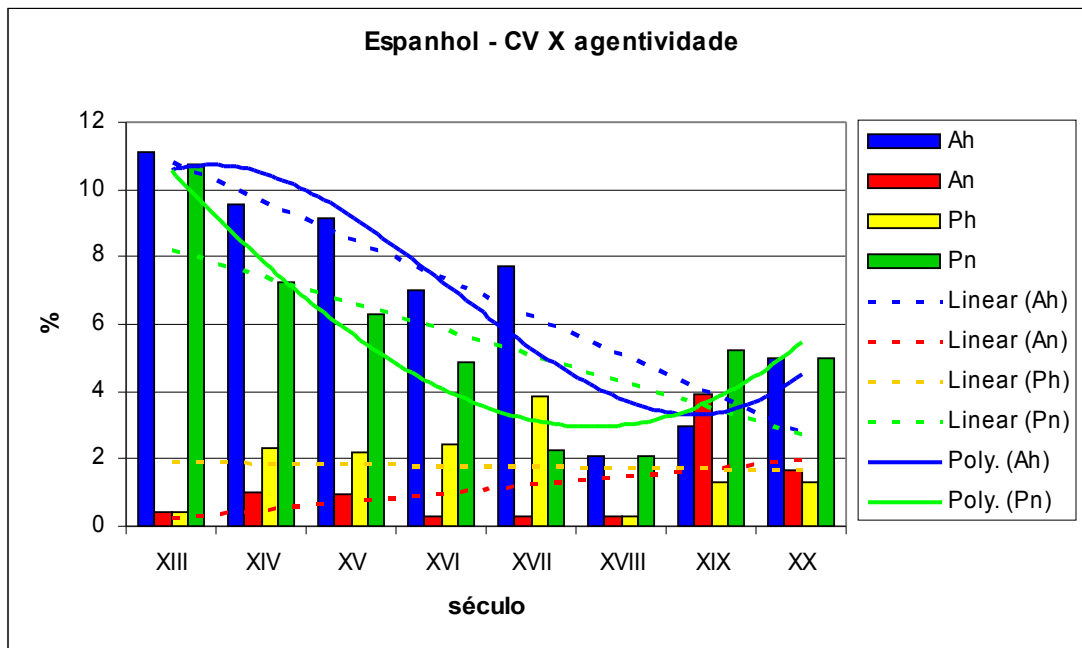


Gráfico E4b – tendências de comportamento das ocorrências de ordem j (CV) segundo a agentividade de S e de C.

Com respeito aos casos em que o complemento do verbo é elíptico, as subclasses referentes à agentividade para a ordem **g** (SV) mostram, em primeiro lugar, um comportamento aproximadamente paralelo, com expressividades diferentes; assim é que Ah, o grupo mais comum, tem pico no século XVI (exatamente como também acontece para Pn, seu antípoda e o segundo mais freqüente). An, por seu turno, tem distribuição senoidal de período igual a 500 anos, e Ph tem comportamento decrescente, apesar de não se apresentar em proporções superiores a 2%. Os exemplos abaixo são relativos a essas características:

Ah – Mi amo quedó desesperado, y como no lo podía remediar (CON)

An – Los montes surgían con suavidades doradas (SIL)

Ph – mi cabo de escuadra, a quien yo respetaba como al Rey (CON)

Pn – El automóvil va jadeante y antipático (SIL)

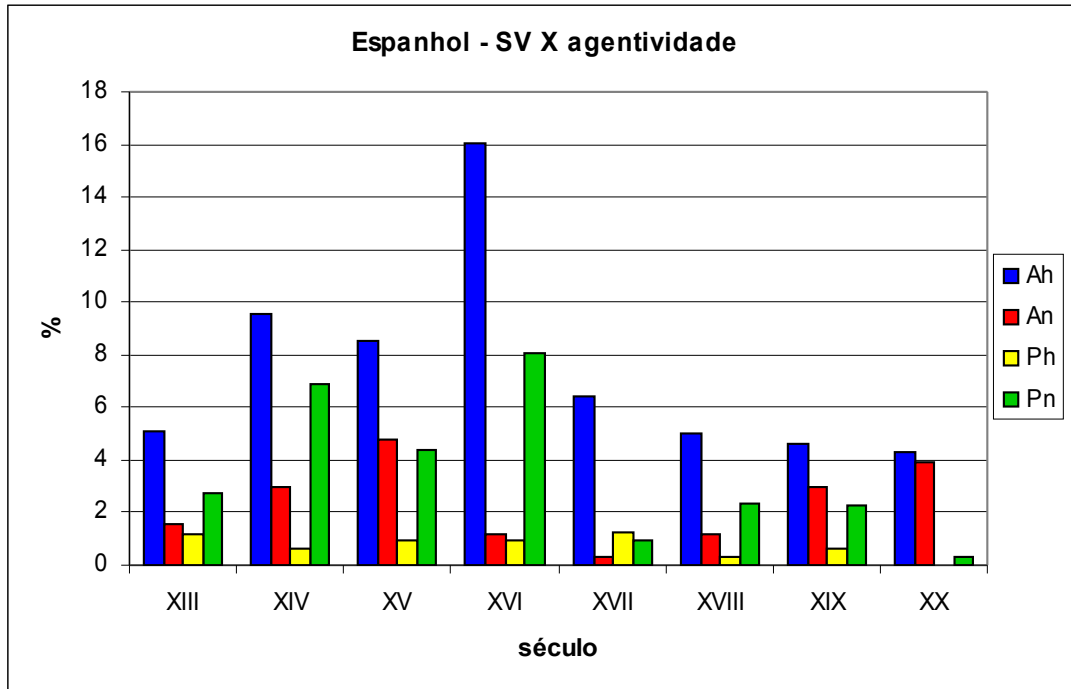


Gráfico E5a – distribuição das ocorrências de ordem g (SV) segundo a agentividade de S e de C.

É interessante notar, para a combinação dos traços g (SV) e *agentividade*, um paralelismo entre as subclasses prototípicas Ah e Pn, cujas frequências diminuem à mesma taxa – apesar de a primeira ser mais numerosa que a segunda; as duas subclasses minoritárias exibem comportamentos levemente dispares, com aumento discreto para An e leve declínio para Ph.

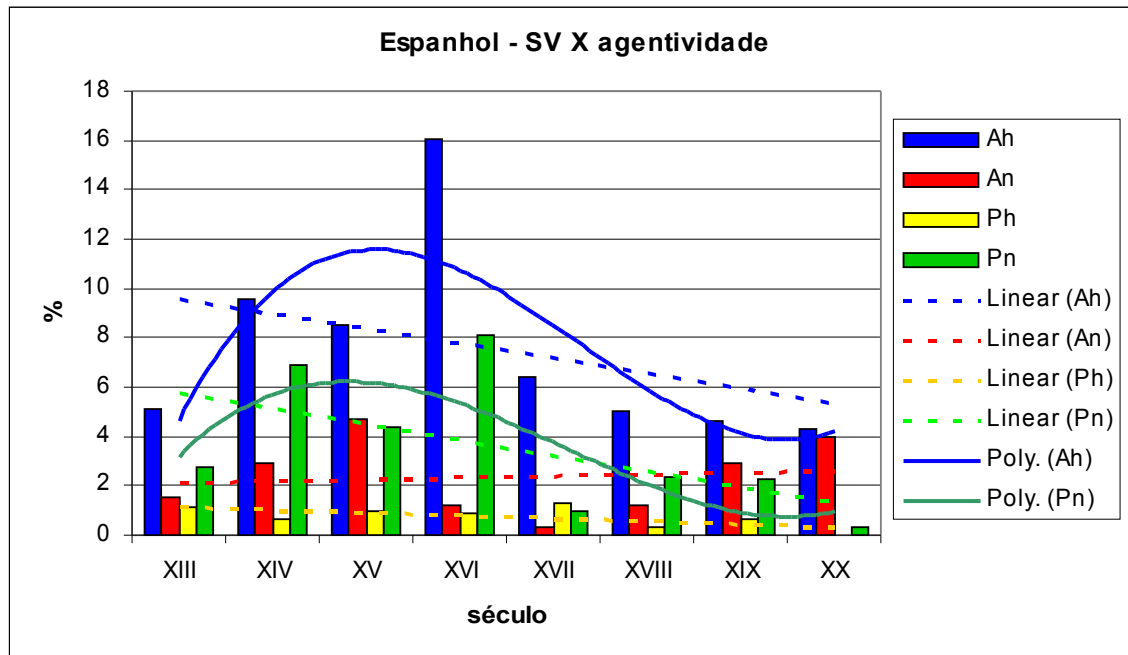


Gráfico E5b – tendências de comportamento das ocorrências de ordem g (SV) segundo a agentividade de S e de C.

Para a ordem especular à anterior, **h** (VS), é interessante notar que a tendência de Ah é acompanhada por Pn, mas com uma oscilação muito mais pronunciada; observa-se ainda que o século XVI corresponde a uma espécie de reativação da ordem **h** (VS), que parecia estar em declínio até o século XV. An e Ph, por sua vez, apresentam um perfil sempre ascendente, e as proporções relativas de An chegam mesmo a ultrapassar as das subclasses Ah e Pn no século XVIII. Vejamos exemplos colhidos no *corpus*:

Ah – Cede el viajero al cariñoso impulso (CAS)

An – et los fechos que fazen las estrellas (CGE)

Ph – en tal guisa ques non entendien unos a otros (CGE)

Pn – Largos rizos que despeinaba el viento (CAS)

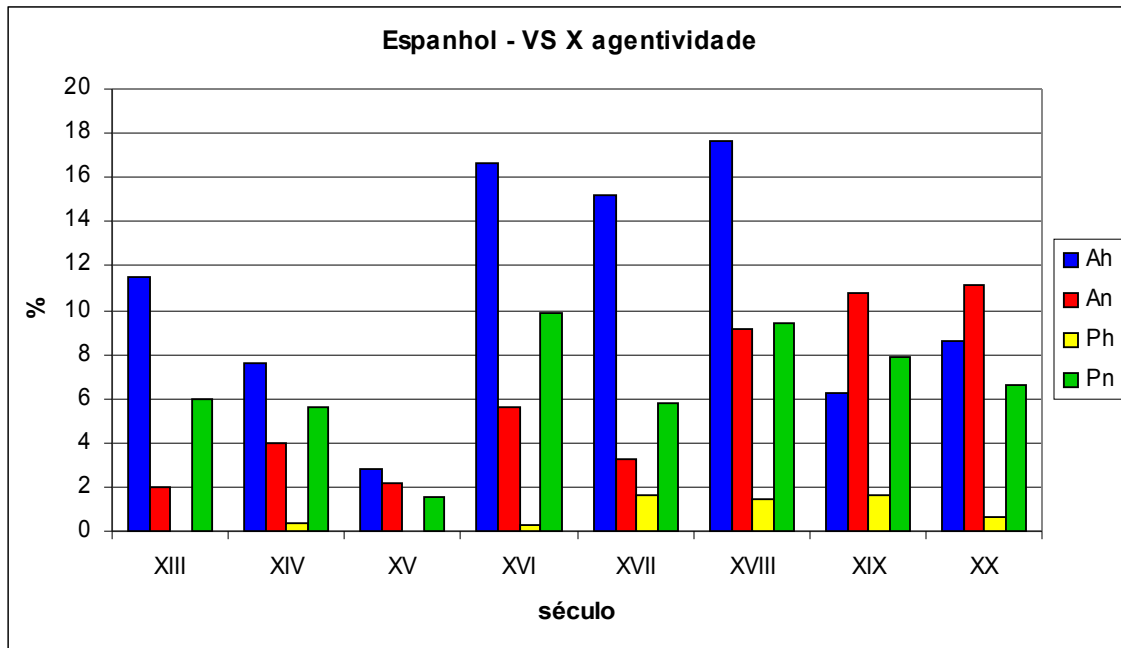


Gráfico E6a – distribuição das ocorrências de ordem h (VS) segundo a agentividade de S e de C.

A resolução das tendências para as subclasses de agentividade do gráfico E6a acima corroboram o caráter ascendente da subclasse An associada à ordem **h** (VS). Veja-se que, nesse caso, há uma ascensão mais pronunciada desse traço An do que para a categoria **g** (SV); por outro lado, essa mesma ordem verifica um decréscimo das proporções relativas de Ah, sem que isso seja acompanhado de aumento apreciável das subclasses minoritárias.

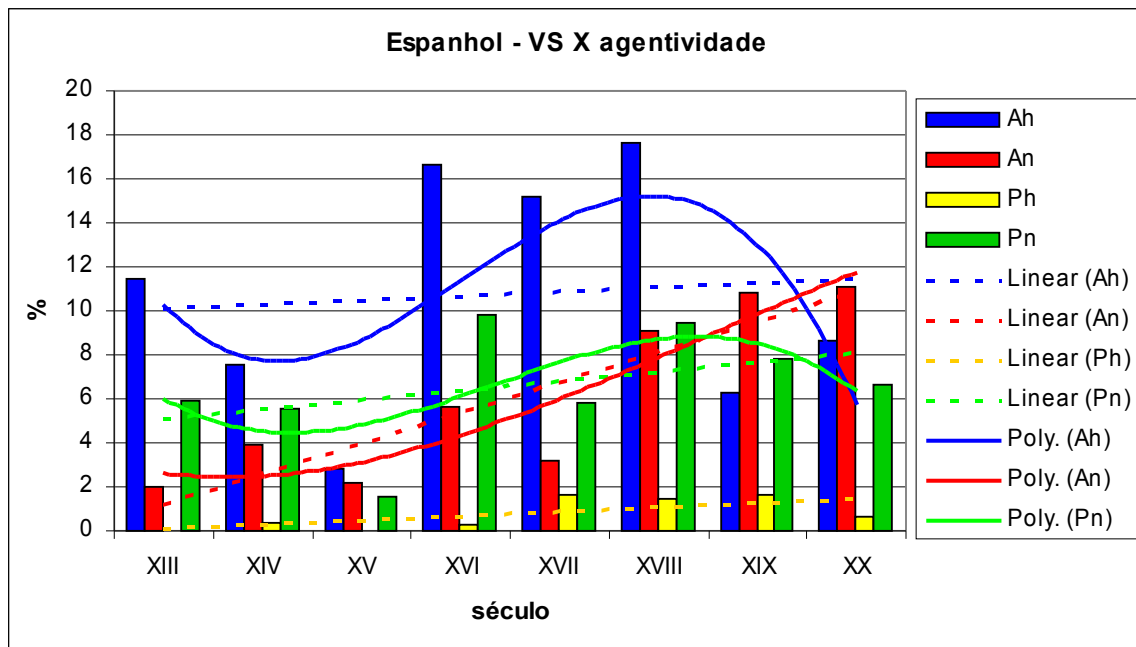


Gráfico E6b – tendências de comportamento das ocorrências de ordem h (VS) segundo a agentividade de S e de C.

A tabela E1 sumariza as observações feitas acima acerca do tratamento do fator *agentividade* no *corpus* espanhol; destacam-se nela as tendências gerais para esse critério de classificação das ocorrências, como o aumento das frequências do traço [-humano] nas ordens SVC e a relativa estabilidade das frequências associadas ao grupo Ph.

ordem	Ah	An	Ph	Pn
SVC	→	↗	→	↗
VC	↘	↗	→	→
CV	↘	→	→	↘
SV	↘	→	→	↘
VS	→	↗	→	→

Tabela E1 – comportamento geral das ordens mais comuns no *corpus* espanhol segundo a agentividade de S e de C.

No que respeita ao ambiente oracional, o conjunto de ocorrências mostra um comportamento extremamente interessante. Nos gráficos apresentados a seguir, apresenta-se a evolução individual de cada ordem em todo o intervalo estudado, o que permite acompanhar o perfil diacrônico do sistema para diversas possibilidades de ordenamento dos constituintes.

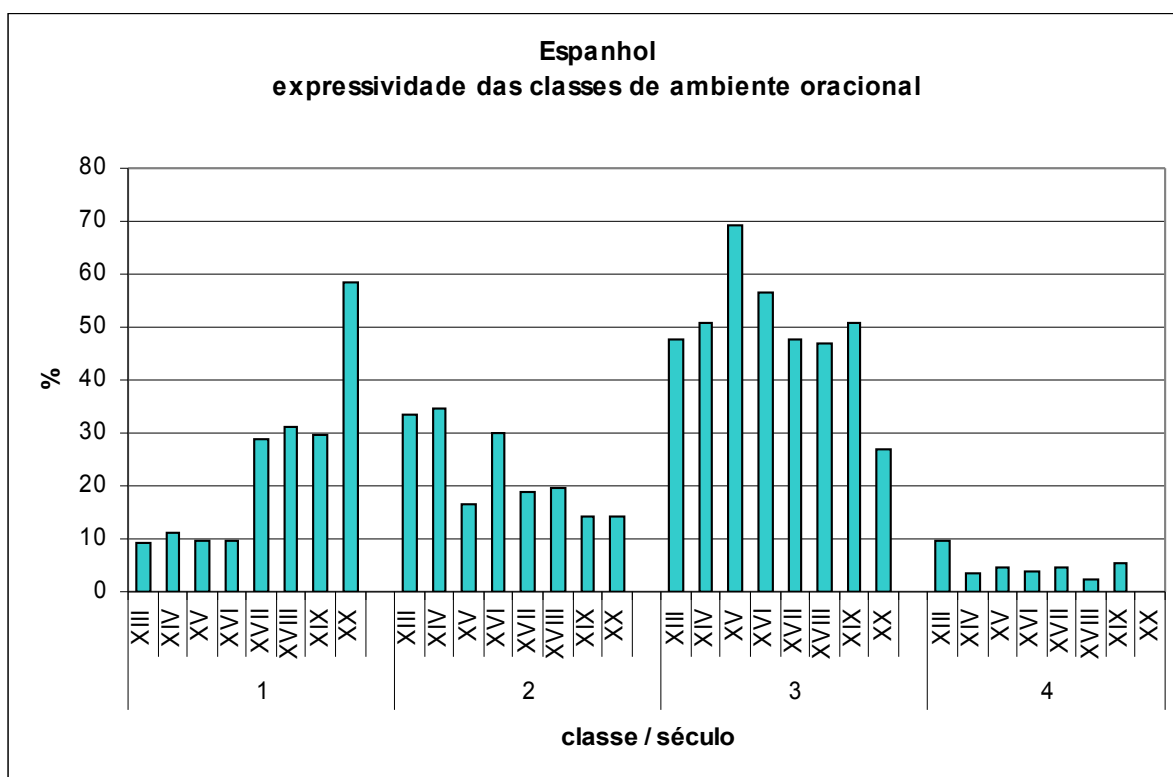


Gráfico E7 – distribuição geral das ocorrências segundo o ambiente oracional.

Verifica-se, inicialmente, um aumento substancial na proporção relativa de orações principais ($iRRL(1) = 6,23$), o qual é acompanhado de decréscimo nas frequências absolutas de orações coordenadas ($iRRL(2) = -2,82$), de subordinadas ($iRRL(3) = -2,63$) e, em menor escala, de coordenadas em subordinação ($iRRL(4) = -0,78$). Esse perfil vai-se repetir diversas vezes adiante, no cruzamento do fator *ambiente oracional* com os demais.

A distribuição das ocorrências no *corpus* espanhol indica, por exemplo, uma coincidência entre os perfis das ordens mais comuns e o das orações principais. Nota-se claramente, no Gráfico E8, o predomínio da classe genérica [VC] em relação a [CV] e às

ordens de complemento não-expresso, **g** (SV) e **h** (VS). Os valores numéricos obtidos para a superclasse [VC] são francamente liderados pelas ordens SVC (**a**) e VC (**i**), as quais apresentam distribuição semelhante à do grupo [VC] como um todo. É importante ressaltar, no entanto, que o mesmo perfil ascendente é observado nas categorias **j** (CV), **g** (SV) e **h** (VS), mas este é sensivelmente menos expressivo do que o de **i** (VC). Essas afirmações indicam a) que as orações principais estão-se especializando na ordem [VC], b) que as proporções das demais ordens também aumentaram nesse ambiente, o que só se pode ter processado às expensas das frequências relativas dos outros tipos oracionais, e c) que a mudança de ordem dos constituintes S-V-C pode ser observada concomitantemente à que atinge o SN espanhol – cf. Totaro 1998.

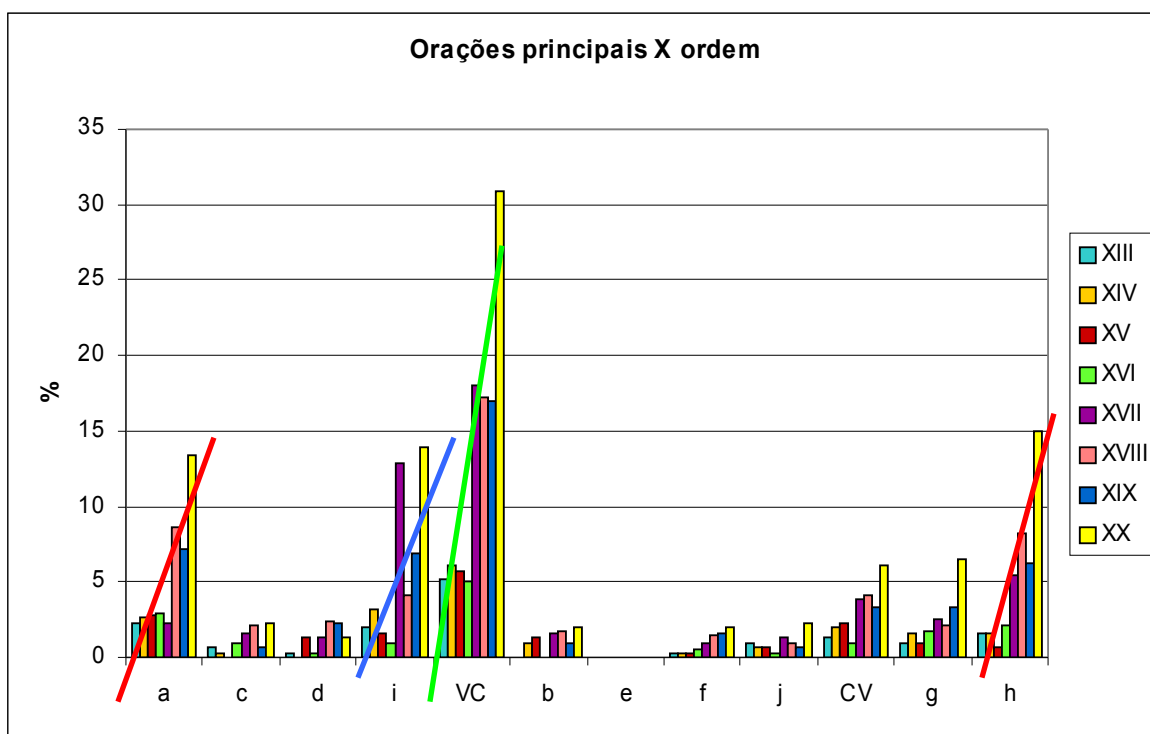


Gráfico E8 – distribuição geral das ocorrências de orações principais segundo a ordem dos constituintes.

Compare-se o gráfico acima com o obtido para as orações coordenadas (classe 2), apresentado a seguir. Nota-se imediatamente a curva descendente da superclasse [VC], liderada absolutamente pela ordem homônima, **i** (VC); além disso, a diferença entre a máxima e a mínima está em torno dos 15%, o que responderia pela metade do aumento

observado nas orações principais. Para as demais ordens, a variação é pouco expressiva, mas ainda assim os perfis de **i** (VC), **g** (SV) e **h** (VS) são decrescentes; mesmo a superclasse [CV], que corresponde à soma das freqüências relativas de quatro arranjos possíveis dos constituintes S-V-C, praticamente não ultrapassa os 5% de freqüência relativa total no intervalo de amostragem. O gráfico sugere, portanto, um esvaziamento das orações coordenadas, o qual é mais marcante para as ordens [VC] e, nesse grupo, para as orações em que o sujeito é elíptico, **i** (VC).

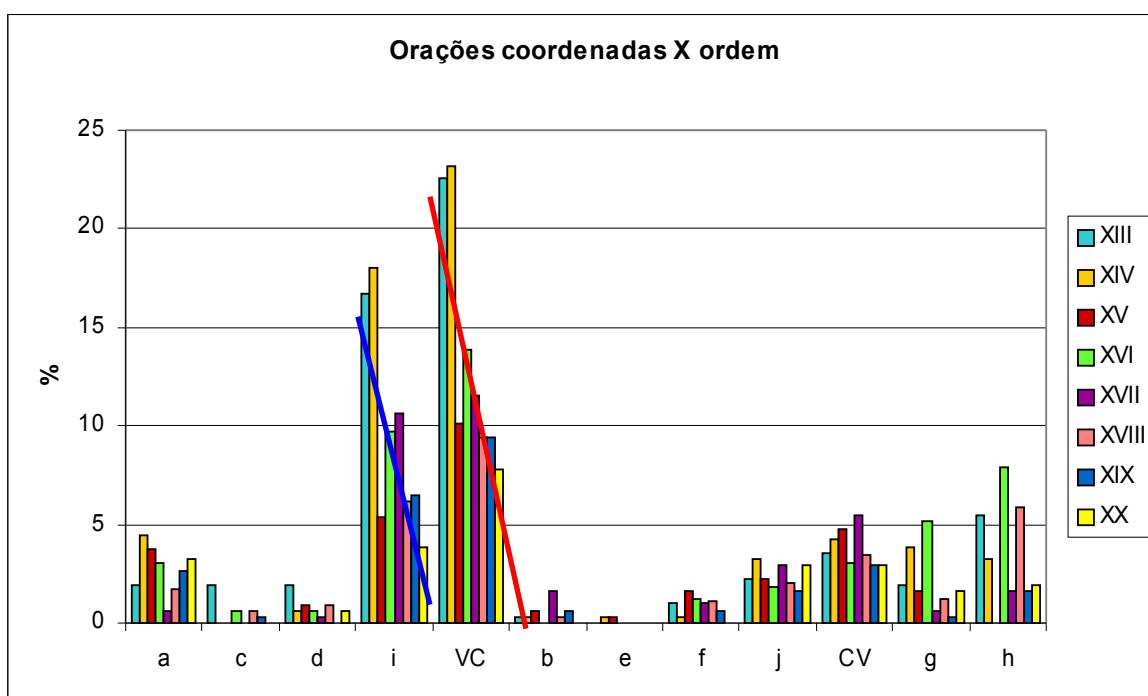


Gráfico E9 – distribuição geral das ocorrências de orações coordenadas segundo a ordem dos constituintes.

Observações semelhantes podem ser estendidas às construções hipotáticas: o comportamento da classe genérica [VC] é dominado pelas ordens VC (**i**) e SVC (**a**), e apresenta um perfil levemente decrescente. As demais ordens – [CV], liderada pelas construções com sujeito não-expresso, **j** (CV); SV (**g**) e VC (**h**) – também mostram tendências à diminuição da freqüência com que ocorrem nesse ambiente. O máximo absoluto de **i** (VC) / **a** (SVC) também aparece em **j** (CV), **g** (SV) e **h** (VS), coincidindo com o século XV, o que sugere apenas que o caráter dos textos utilizados na amostragem é algo excêntrico em relação ao padrão da época.

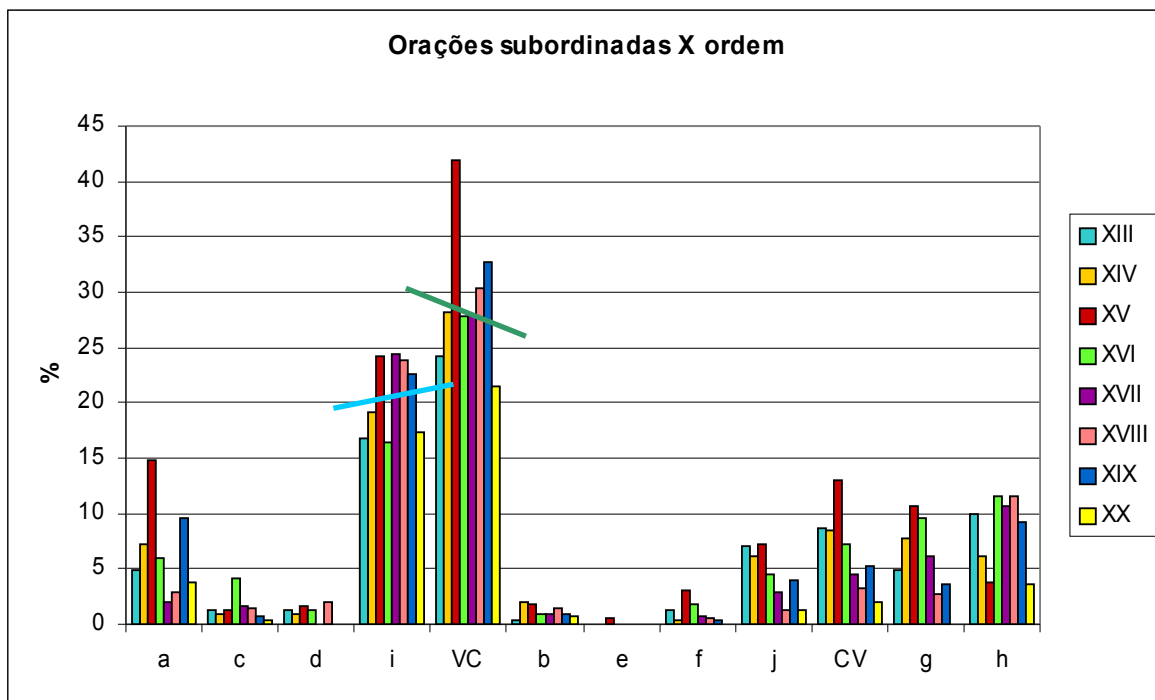


Gráfico E10 – distribuição geral das ocorrências de orações subordinadas segundo a ordem dos constituintes.

Por fim, as orações coordenadas em subordinação (classe 4 de *ambiente oracional*), bem menos freqüentes que os demais tipos oracionais, apresentam perfis levemente decrescentes a partir do tratamento numérico do *corpus* – especialmente para as ordens VC (i) e CV (j); a ordem VS (h) é a única que mostra aumento das proporções absolutas, mas os valores de freqüência desse grupo praticamente não ultrapassam 1%.

Somando-se, portanto, os efeitos de decréscimo generalizado nas proporções relativas de praticamente todas as ordens nas orações coordenadas, subordinadas e coordenadas em subordinação, percebe-se que, de fato, as orações principais aumentariam com a mesma expressividade.

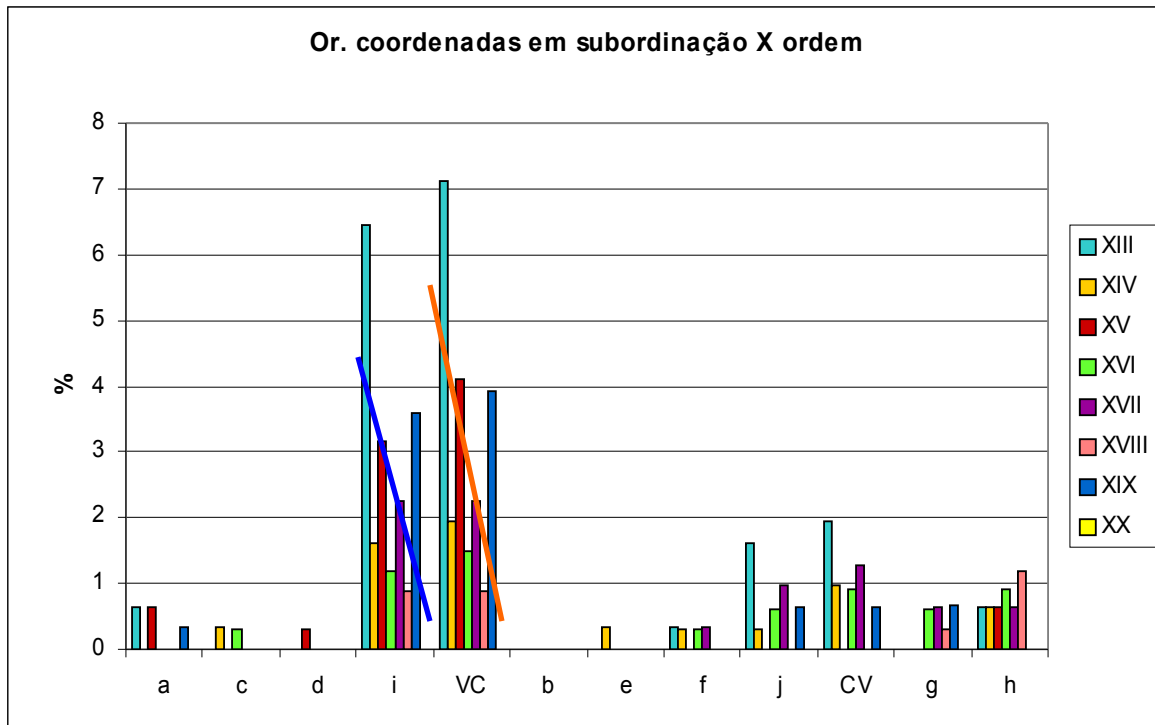


Gráfico E11 – distribuição geral das ocorrências de orações coordenadas em subordinação segundo a ordem dos constituintes.

Encontram-se resumidas, à Tabela E2, as tendências de comportamento das ordens mais expressivas e dos grupos VC e CV diante do fator *ambiente oracional*. Observe-se que, em termos de conjunto, as ordens que constam da Tabela mostram aumento de frequência em orações principais, mas tendem a ser menos frequentes ou a se manterem estáveis nos demais ambientes.

<i>ambiente</i>	<i>a</i>	<i>i</i>	<i>g</i>	<i>h</i>	<i>VC</i>	<i>CV</i>
<i>1 (principais)</i>	↗	↗	↗	↗	↕	↗
<i>2 (coordenadas)</i>	→	↘	↘	↘	↘	→
<i>3 (subordinadas)</i>	↘	→	↘	→	→	↘
<i>4 (coord. em subordinação)</i>	.	↘	.	.	↘	.

Tabela E2 – visão geral dos comportamentos das ordens mais comuns segundo o ambiente oracional.

As classes do fator *agentividade* são mais sensíveis ao *ambiente oracional*, obedecendo ao padrão já delineado no Gráfico E7 para este fator. Observe-se, nos gráficos seguintes, a estreita correlação formal existente entre os perfis dessas classes e o do Gráfico E7. Para as orações principais, por exemplo, os perfis mais salientes – Ah, An e Pn – mostram, todos, uma tendência crescente paralela à curva de orações principais em E7. Isso significa que a distribuição dos traços de *agentividade* está condicionada ao *ambiente oracional*, como outros fatores discutidos adiante. Tomem-se os exemplos:

Ah – Estonce otorgaronlo todos e - TRO XIII

An – Estos ella los da, ella los tira a quien quiere - FOR XV

Ph – servíle con voluntad, y él me quería bien - CON XVII

Pn – El perdulario desconocia sus derechos - Prod XX

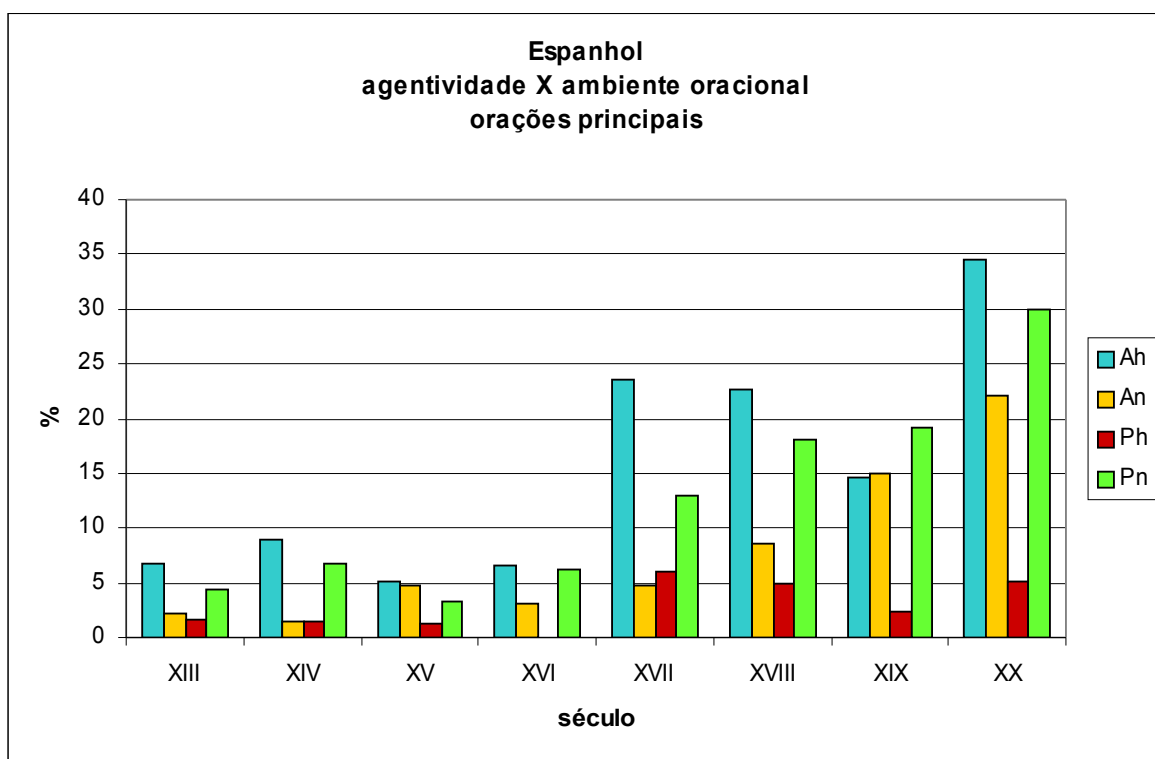


Gráfico E12a – distribuição geral das ocorrências de orações principais segundo a agentividade.

Podem-se resolver as distribuições de ocorrências ilustradas no Gráfico E12a, portanto, segundo retas cuja inclinação é invariavelmente positiva. Destacam-se, naturalmente e como já salientado, as classes de agentividade Ah e Pn – não apenas em termos numéricos, mas de inclinação das retas de regressão linear: no século XVII, há um salto proporcional nas frequências absolutas, fazendo com que Ah (iRRL = 3,48) e Pn (iRRL = 3,48) se sobressaíam.

São exemplos de ocorrências encontradas no *corpus*:

Ah – Thales, Pitágoras ... afirman que la tierra en ninguna manera se puede toda morar

[IND]

An – La mar ... guarda su redondez en medio y sobre la tierra [IND]

Ph – Ell abadesa que la amaba rogó é mandó á todas las monias [ENX]

Pn – Alentaba estas voces el duque de Norgoña con gran número [FEL]

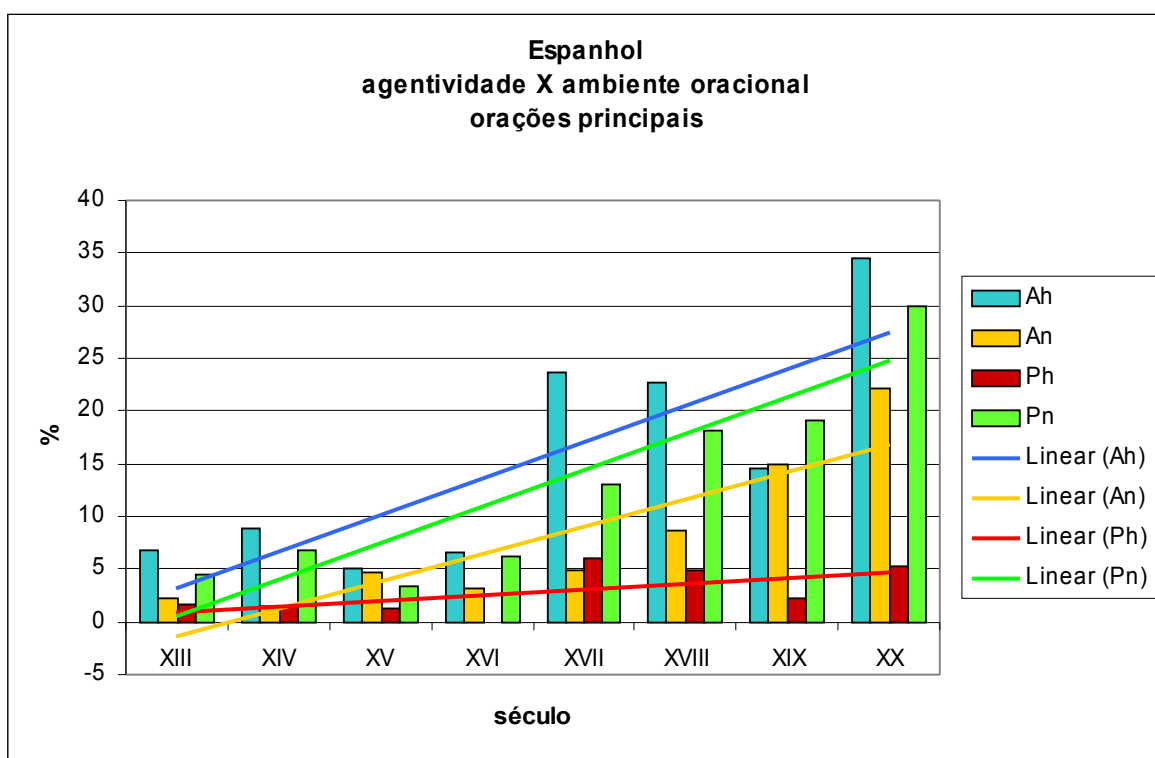


Gráfico E12b – distribuição geral das ocorrências de orações principais segundo a agentividade.

Acompanham o perfil decrescente das coordenadas (cf. Gráf. E7) as classes de *agentividade*, observando-se a diminuição generalizada das frequências das classes dominantes, Ah (iRRL = -3,08) e Pn (iRRL = -1,92). Entretanto, a despeito da escala mais reduzida, as proporções de An parecem apresentar ligeiro aumento.

São exemplos de ocorrências do *corpus* espanhol:

Ah – pero llevaba en la mano una caña rematada por una calavera [PROD]

An – ca este present reino lla muerte lo destaiará [ENX]

Ph – Bautizáronla en la iglesia de Santa María de Magaña [AGR]

Pn – alzó un chapín para darme, y alzando yo la cantarilla se la tiré [CON]

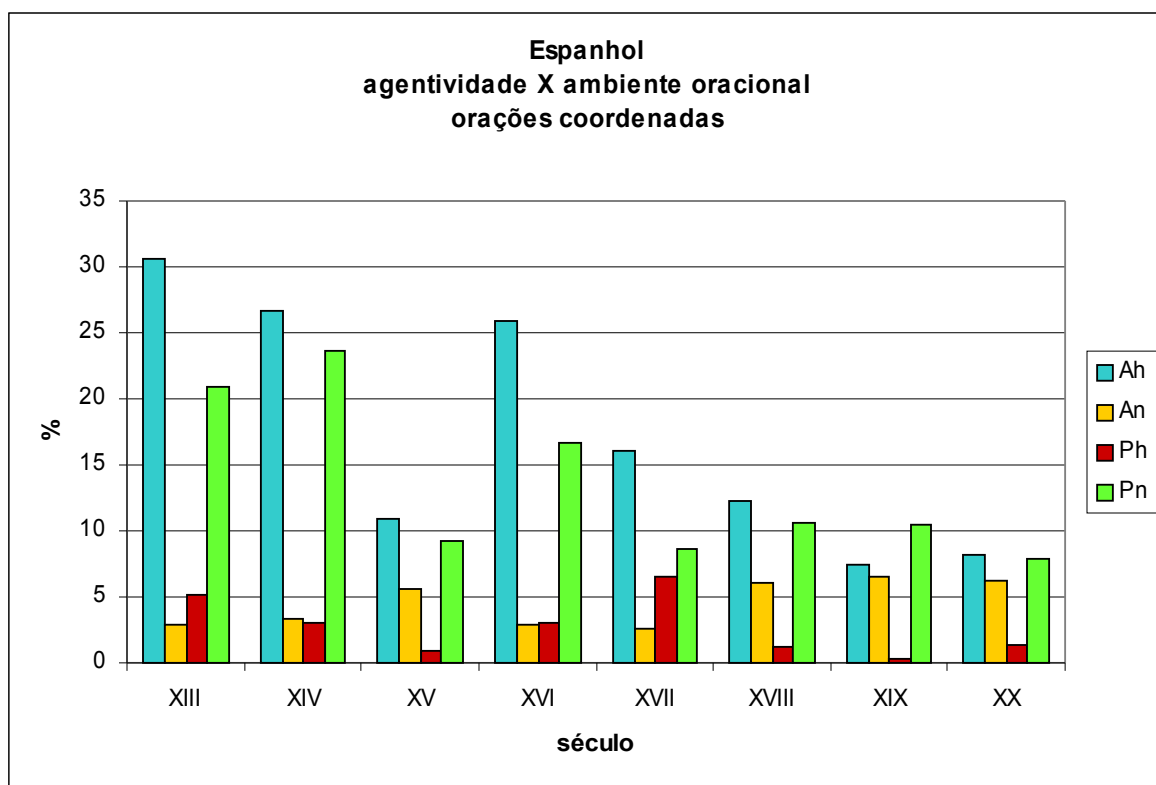


Gráfico E13a – distribuição geral das ocorrências de orações coordenadas segundo a agentividade.

A regressão linear dos perfis apresentados no gráfico anterior ilustra mais claramente as tendências gerais das classes de ocorrências: veja-se que a taxa de diminuição de Ah é, com efeito, a mais alta do grupo. Além disso, as frequências de An sofrem, de fato, pequeno aumento (iRRL = 0,48) no longo intervalo estudado; esse subconjunto de dados se comporta de maneira oposta ao padrão das outras classes de coordenadas, e cruza as retas de Ah e de Pn no último ponto da amostragem.

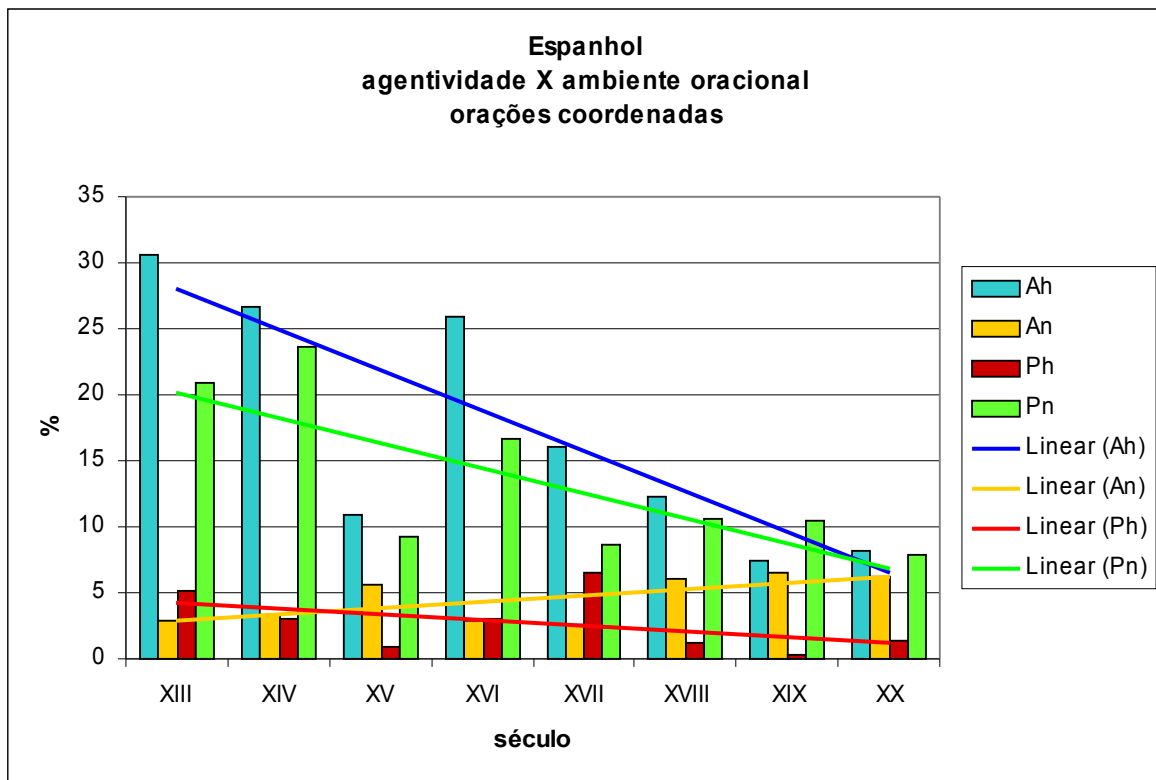


Gráfico E13b – distribuição geral das ocorrências de orações coordenadas segundo a agentividade.

A situação das orações subordinadas em termos do fator *agentividade* não é distinta da que se apresenta no Gráfico E7: há um pequeno aumento das frequências de subordinadas no século XV em relação aos anteriores, mas as médias retomam seu padrão descendente, tornando-se até mesmo menos expressivas para Ah que para Pn. Também no caso das subordinadas, as taxas de An apresentam ligeiro aumento, enquanto as de Ph são relativamente baixas.

Exemplos de ocorrências associadas a orações subordinadas são as seguintes:

Ah, Pn – porque puede hacer obras de que haya mayor merecimiento [CAB]

An – la naturaleza transmitida ... los inclinaba al desorden [PROD]

Ph – las bañistas en Castro eran las que el viajero encuentra en el Lido [CAS]

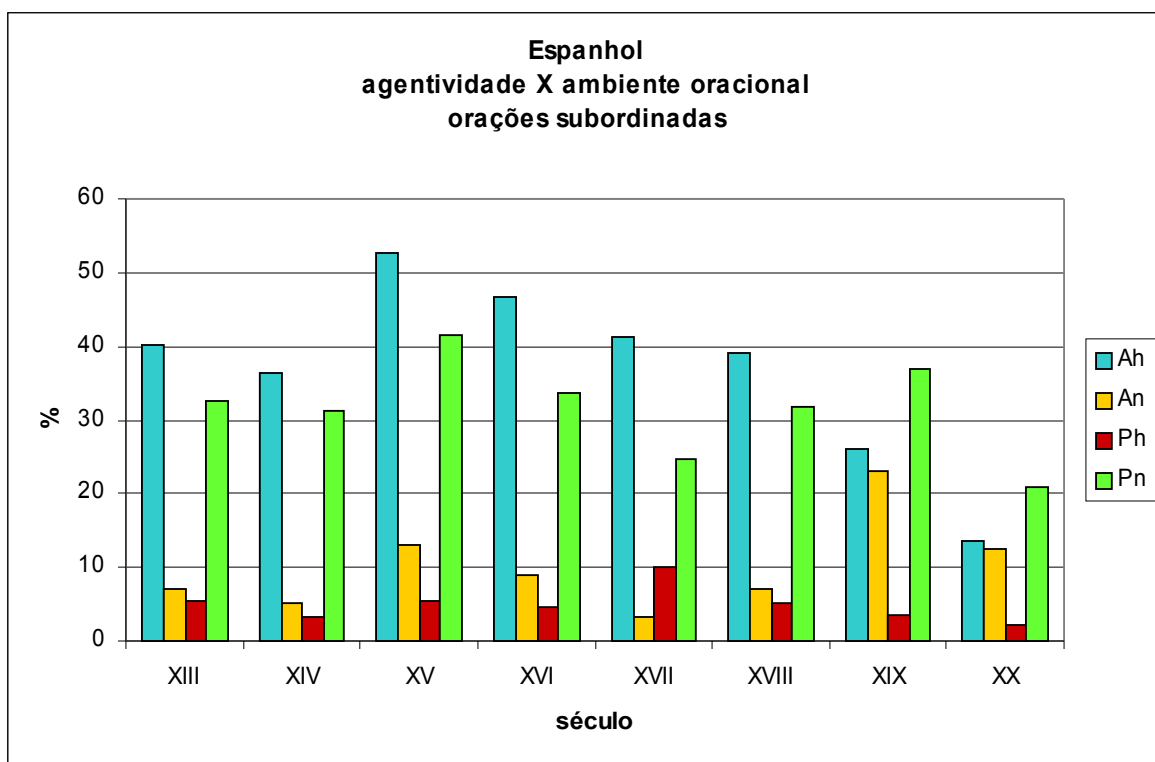


Gráfico E14a – distribuição geral das ocorrências de orações subordinadas segundo a agentividade.

As retas de regressão linear dos conjuntos de ocorrências ligadas às orações subordinadas (classe 3 de *ambiente oracional*) mostram uma semelhança com os resultados de mesma natureza obtidos para as coordenadas (q.v. Gráfico E13b): há convergência entre as retas de Ah (iRRL = -3,35) e de An (iRRL = 1,24). Por outro lado, cruzam-se as retas de Ah e de Pn, uma vez que a inclinação desta é mais baixa (iRRL = -1,09) que a da primeira; em outras palavras, as freqüências relativas de Ah diminuem mais rapidamente que as de Pn em orações subordinadas, ao passo que as de An aumentam.

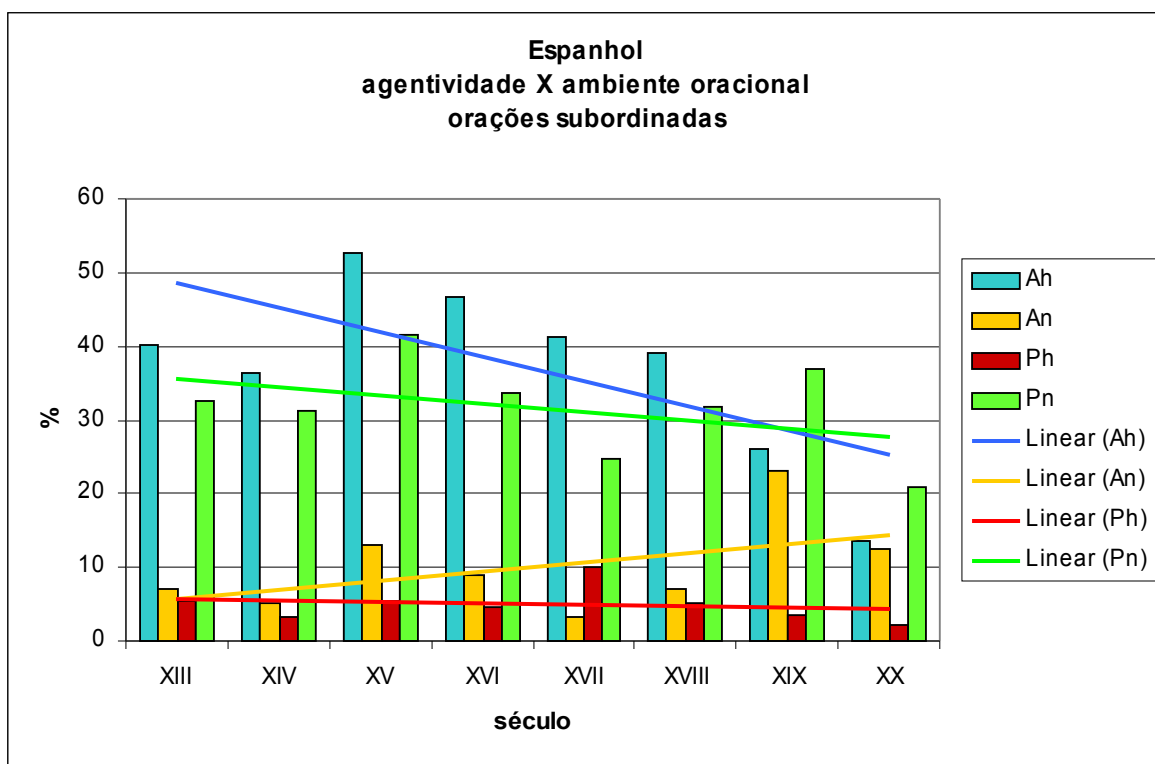


Gráfico E14b – tendências da distribuição das ocorrências de orações subordinadas segundo a agentividade.

As orações coordenadas em subordinação (classe 4 de *ambiente oracional*), por fim, ocorrem em proporções bastante inferiores às dos outros tipos oracionais, e ficam sempre abaixo dos 10% absolutos. Não obstante, é possível divisar-se uma tendência decrescente para os subconjuntos mais comuns – ainda uma vez, Ah e Pn.

Compreendem exemplos de ocorrências desse tipo:

Ah – que por sí sola no podía alcanzar la gracia : perderla y cometer pecado sí [AGR]

An – si yo voy a fazer la barva e se me cae un pelo de la cabeza [FOR]

Ph – no fué menester poco para apaciguarnos y echarnos de la sala [CON]

Pn – y llama a sí y absorbe y emplea en provecho propio los elementos vitales [CAS]

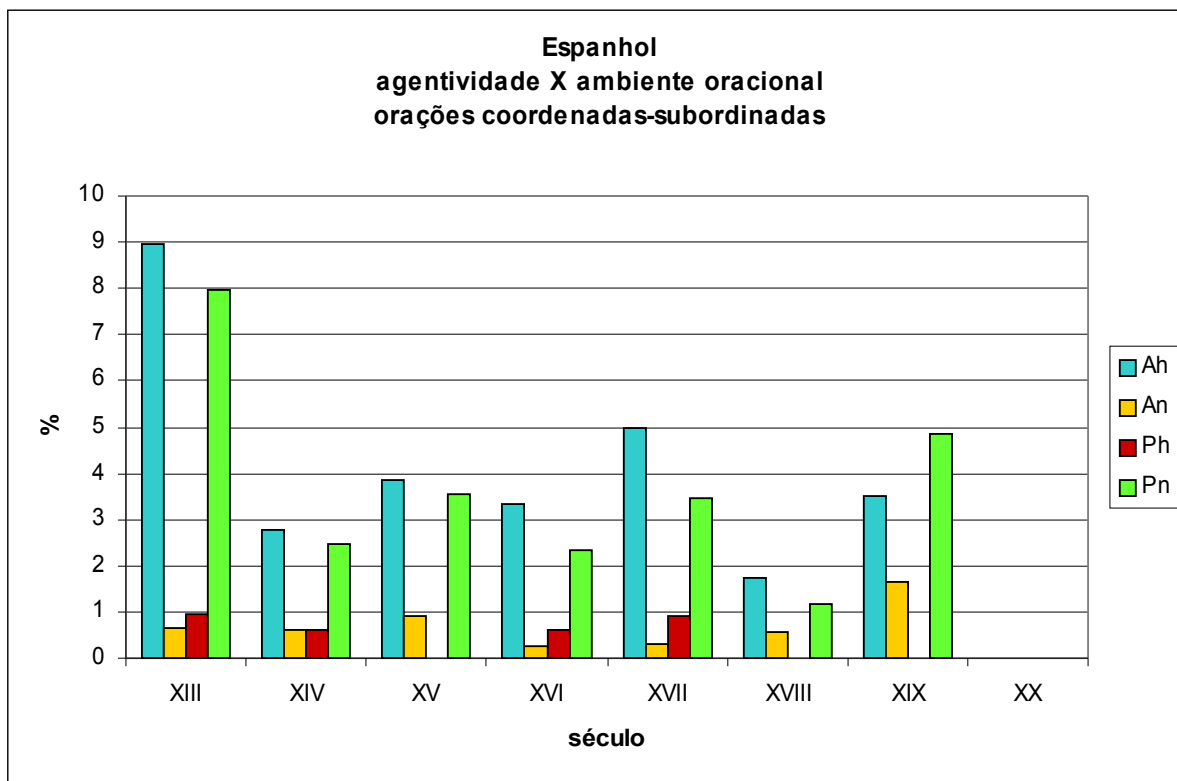


Gráfico E15a – distribuição geral das ocorrências de orações coordenadas em subordinação segundo a agentividade.

Em termos numéricos, os comportamentos das retas de regressão linear para essas categorias são representados por inclinações baixas, em decorrência das pequenas frequências segundo as quais essa classe ocorre. Mesmo assim, as inclinações de Ah (iRRL = -0,76) e de Pn (iRRL = -0,60) não são desprezíveis, e predominam no Gráfico abaixo.

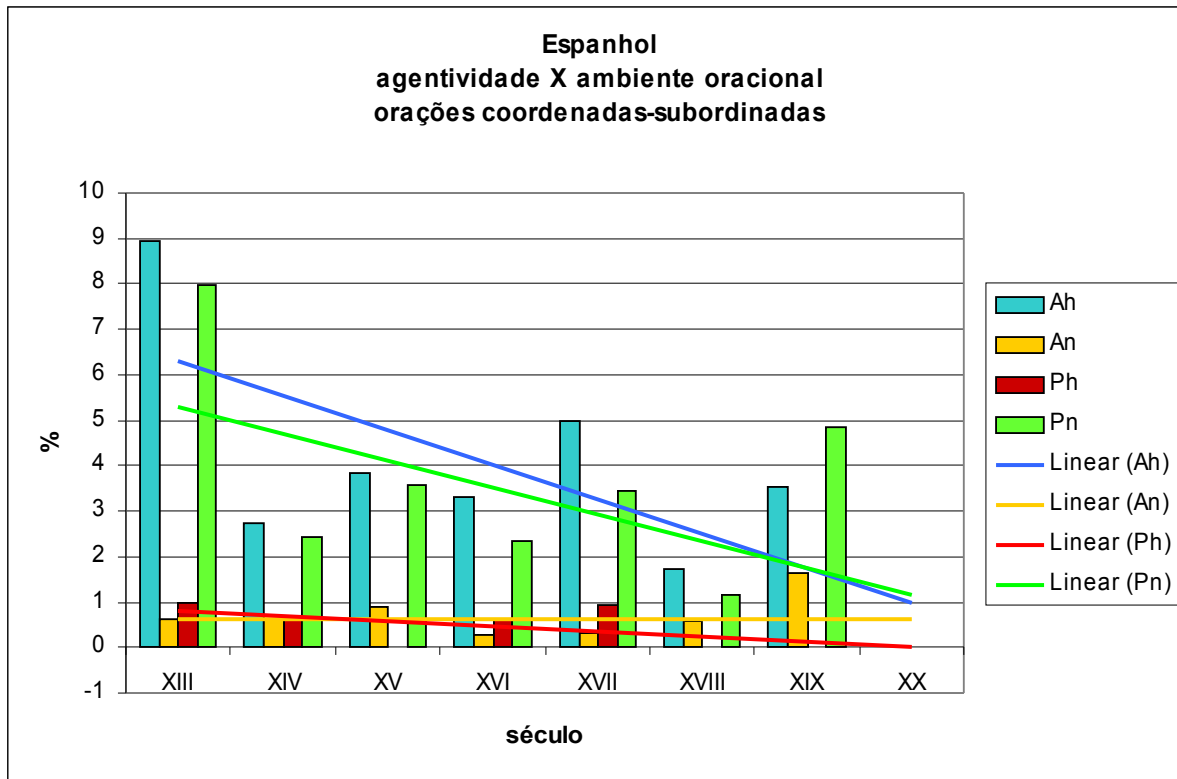


Gráfico E15b – tendências da distribuição das ocorrências de orações coordenadas em subordinação segundo a agentividade.

Apresenta-se, por fim, uma tabela que resume as principais tendências das classes do fator *agentividade*.

<i>agentividad</i>	1	2	3	4
<i>e</i>				
<i>Ah</i>	↗	↘	↘	↘
<i>An</i>	↗	→	↗	→
<i>Ph</i>	↗	→	→	→
<i>Pn</i>	↗	↘	↘	↘

Tabela E3 – comparação entre os comportamentos das classes de *agentividade* em função do ambiente oracional.

O agrupamento dos resultados de *agentividade* para cada ambiente oracional leva à comparação entre os perfis para ambientes diferentes. Uma vez que esses gráficos são confeccionados lançando-se o fator *tempo* no eixo das abscissas, pode-se observar se há ambientes oracionais em que determinada mudança se deu mais rapidamente ou se apresenta defasagem cronológica em relação aos demais. A fim de ampliar a excentricidade das curvas menos proeminentes, alguns gráficos foram normalizados; todos foram construídos a partir de matrizes transpostas (cf. exposto na seção que trata da Metodologia utilizada neste trabalho).

Tome-se, em primeiro lugar, o traço Ah, representado nos Gráficos E16a, E16b e E16c. Nota-se a predominância das proporções relacionadas a orações subordinadas até o século XIX, a partir de quando a tendência crescente apresentada pelas orações principais produz o valor mais elevado no último século da amostragem. É também digno de nota o fato de haver proporcionalmente sempre menos ocorrências de Ah em coordenadas do que em subordinadas.

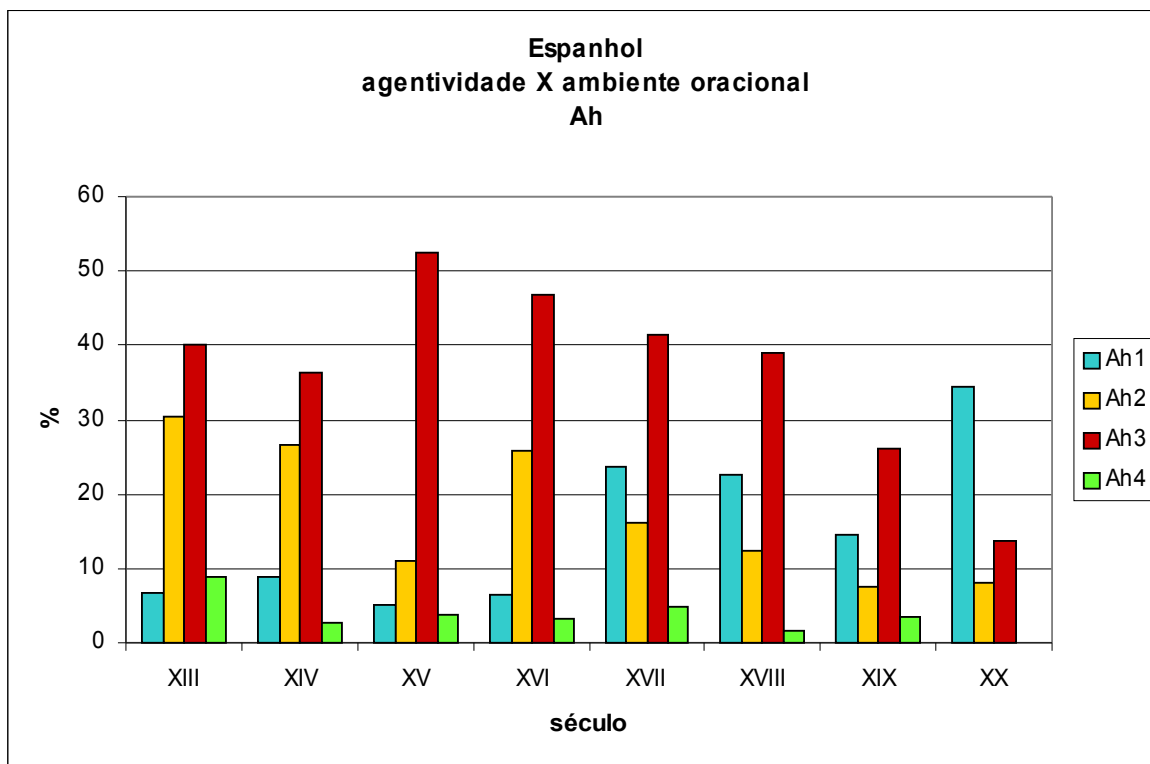


Gráfico E16a(T) – comparação entre os perfis de agentividade **Ah** nos quatro ambientes oracionais propostos.

A resolução desses resultados segundo curvas de ordem 3 ou 4¹ leva à distinção de tendências gerais; entretanto, oferece, além do que o fazem as retas de regressão linear, a possibilidade de se acompanharem, para o eixo horizontal, coincidências ou defasagens de máximos e de mínimos locais. Em outras palavras e no caso específico aqui tratado, torna-se possível verificar se algum dos ambientes oracionais apresenta movimentos de ascensão ou queda antes ou depois dos demais.

Com respeito ao Gráfico E16b, cabe observar que o mínimo relativo às coordenadas em Ah2 no séc. XV e um valor 'compensado' para o século seguinte não comprometem a definição da curva francamente decrescente para a combinação dos traços Ah e *ambiente oracional*. Comentários similares podem ser feitos acerca do que se passa com a curva relativa às orações principais (no gráfico, assinalada como Ah1), em ascensão já mencionada anteriormente.

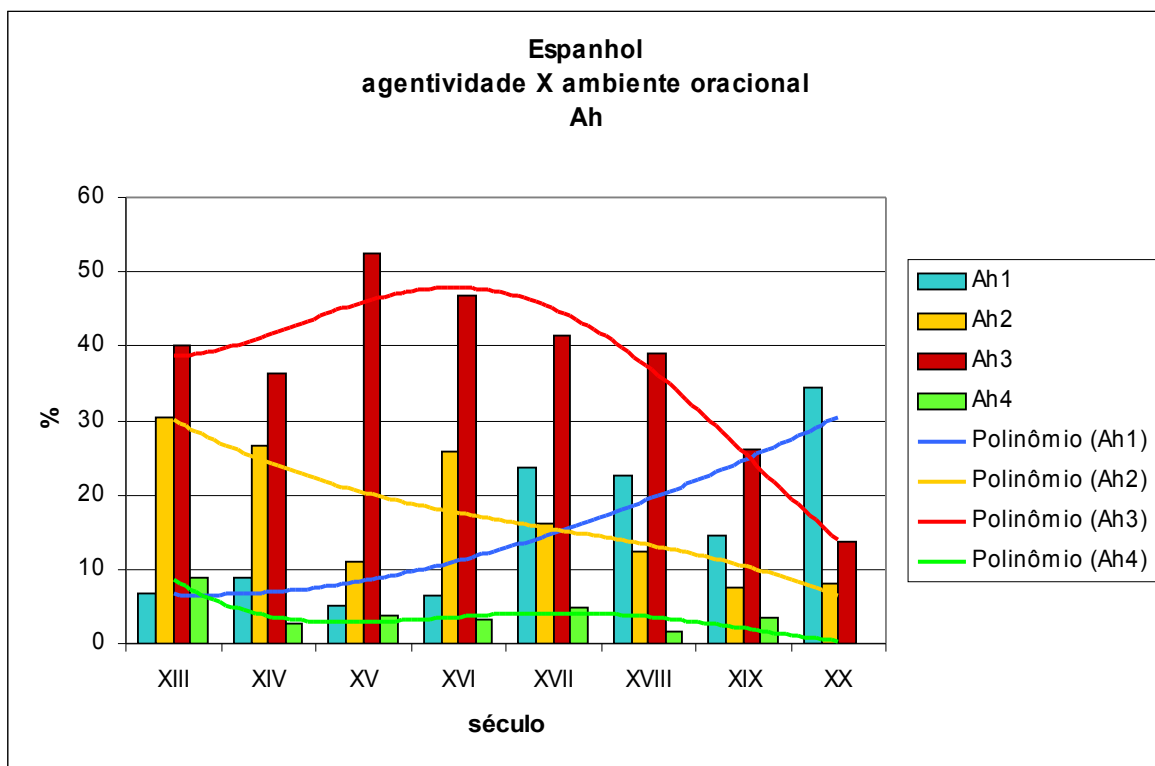


Gráfico E16b(T) – comparação entre os perfis de agentividade **Ah** nos quatro ambientes oracionais propostos.

¹ Polinômios de ordem 4 ou de quarto grau têm a forma geral $y = ax^4 + bx^3 + cx^2 + dx + e$; aos de ordem 3 falta, naturalmente, o termo de grau 4. A curva Ah1 do gráfico E16b, por exemplo, é resolvida pela equação $y = -0,0226x^3 + 0,7684x^2 - 1,8541x + 7,7175$, e a curva Ah2 do mesmo gráfico, por $y = -0,1024x^3 + 1,5325x^2 - 9,7036x + 38,425$.

A normalização das curvas produz resultados interessantes, como é possível observar no Gráfico E16c. Nota-se uma defasagem entre os máximos e dos mínimos locais das curvas correspondentes aos ambientes oracionais mais comuns – 1, 2 e 3 –, como ressaltado para os séculos XVI e XVII: em primeiro lugar, tem-se um pico para as coordenadas (Ah2, em amarelo), seguido das subordinadas (Ah3, em vermelho) e, enfim, das principais (Ah1, em azul). Esse período foi escolhido para comparar os comportamentos das curvas em razão de preceder imediatamente a fase de mudança de AN para NA em espanhol, mas os mínimos locais também se projetam sobre valores levemente distintos do eixo das abscissas.

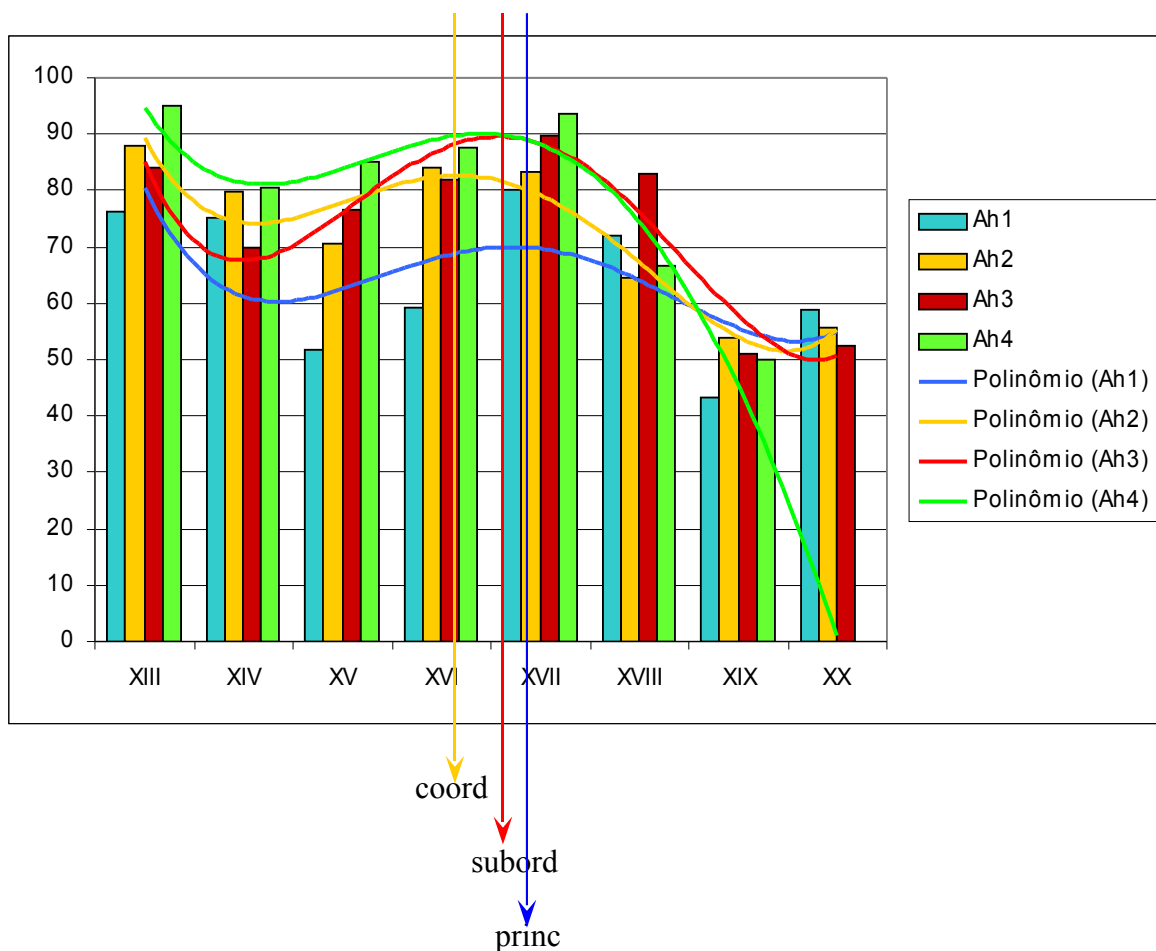


Gráfico E16c(T, normalizado) – comparação entre os perfis de agentividade **Ah** nos quatro ambientes oracionais propostos.

No que toca ao traço An (Gráficos E17a-c), nota-se, inicialmente, que as frequências são mais modestas que as relativas a Ah; também neste caso tem-se aumento das frequências relativas de An em orações principais (cf. Gráfico E7 acima). Entretanto, diferentemente do que se passa com Ah, as frequências de coordenadas e as de subordinadas aumentam leve ($iRRL(An2) = 0,48$) a moderadamente ($iRRL(An3) = 1,24$), acompanhando as principais. As coordenadas em subordinação, An4, não ocorrem em frequência apreciável, e, em decorrência disso, seu perfil não é conspícuo nessa escala.

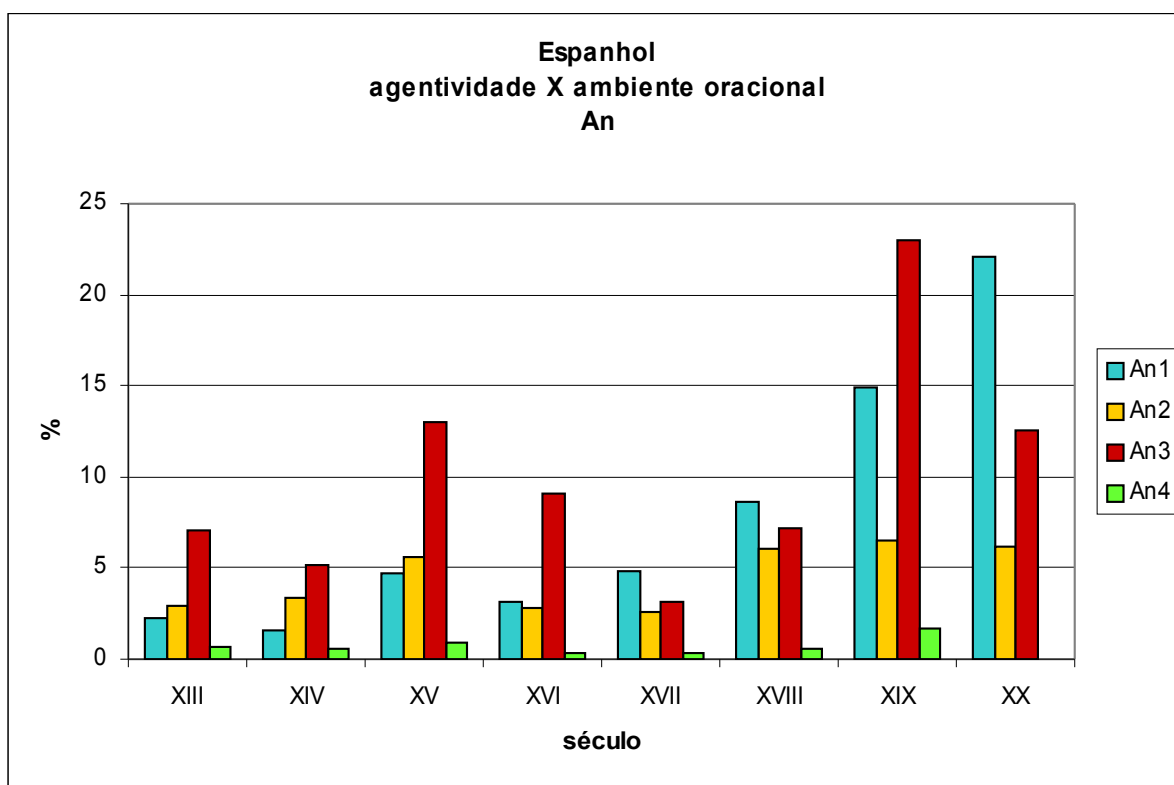


Gráfico E17a(T) – comparação entre os perfis de agentividade **An** nos quatro ambientes oracionais propostos.

As curvas que resolvem os perfis dos grupos An no Gráfico E17b refletem as tendências desses conjuntos de dados: da mesma forma que o observado para E16 e o exposto para as orações principais, o polinômio relativo a An1 descreve uma curva ascendente, ultrapassando o máximo absoluto do gráfico, An3(XIX). Observe-se o perfil levemente crescente de An2, em contraste com o sensível decréscimo ($iRRL = -3,08$) obtido para essa mesma ordem em Ah. De maneira geral, An4 também apresenta valores extremamente baixos, e o perfil dessa classe deve ser apreciado a partir da normalização das proporções relativas.

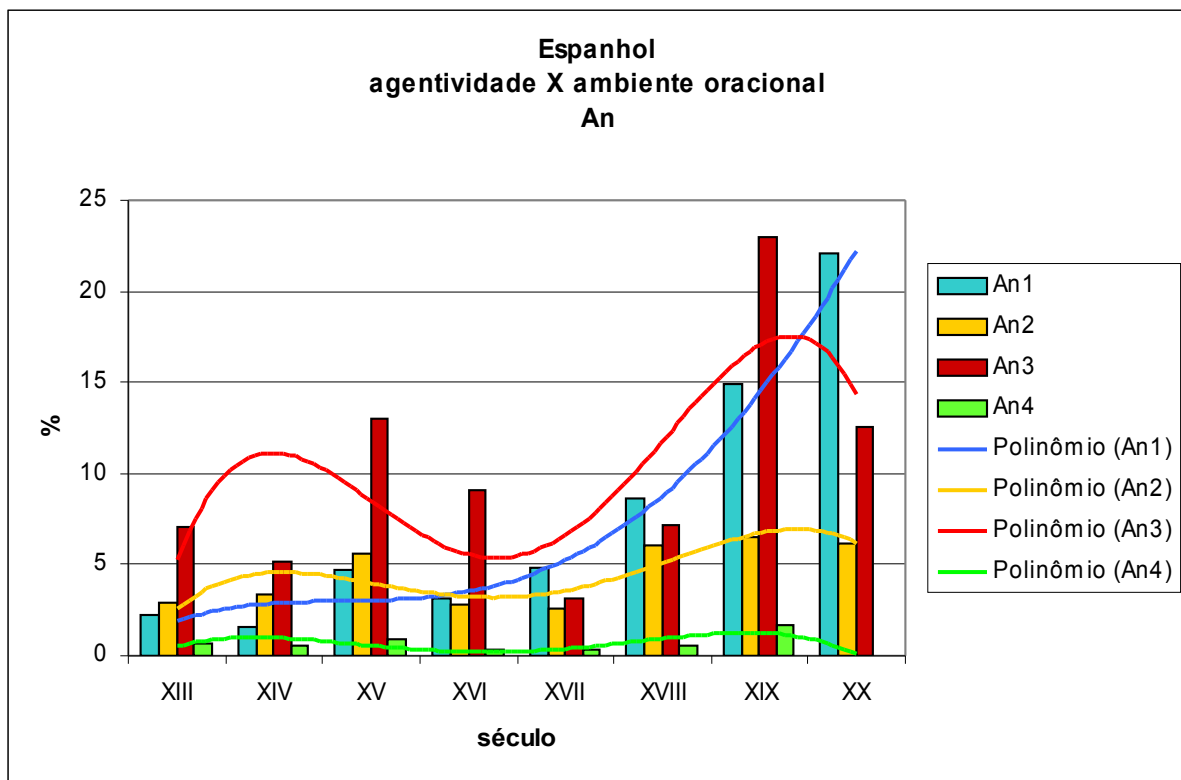


Gráfico E17b(T) – comparação entre os perfis de agentividade An nos quatro ambientes oracionais propostos.

Após a normalização dos perfis, percebe-se a mesma relação temporal identificada para Ah: a curva das coordenadas apresenta mínimo local antes das subordinadas (ambas no século XVI), e estas antes das principais (no século XVII). Os pontos referentes aos máximos absolutos, à direita do gráfico, devem ser adotados com cautela, já que os perfis nas regiões limítrofes podem-se alterar.

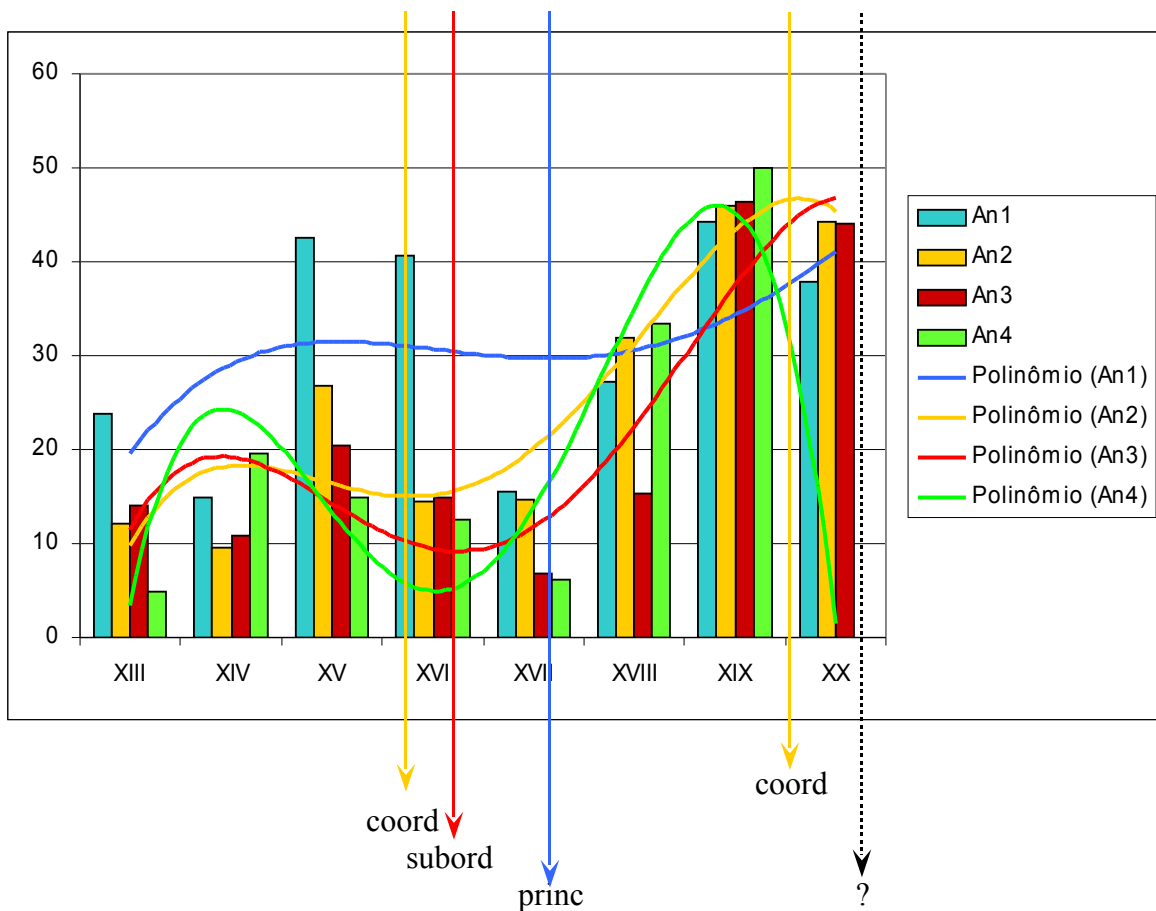


Gráfico E17c(T, normalizado) – comparação entre os perfis de agentividade **An** nos quatro ambientes oracionais propostos.

O traço [-humano] para o paciente da ação foi atribuído a um número menor de ocorrências do que se observa para as classes anteriores. Não obstante, é possível divisar um aumento discreto das freqüências absolutas de Ph em orações principais, a exemplo do que se viu, em maior escala, para Ah (Gráficos E16) e An (Gráficos E17). É digna de nota outra semelhança entre as três classes: o fato de a freqüência de Ph3 predominar em quase todo o intervalo de amostragem, mas ser ultrapassada por Ph1 no século XX. Por fim, observe-se que os máximos absolutos para as três categorias mais freqüentes estão no século XVII – enquanto, nas classes Ah e An, esses máximos estão desencontrados.

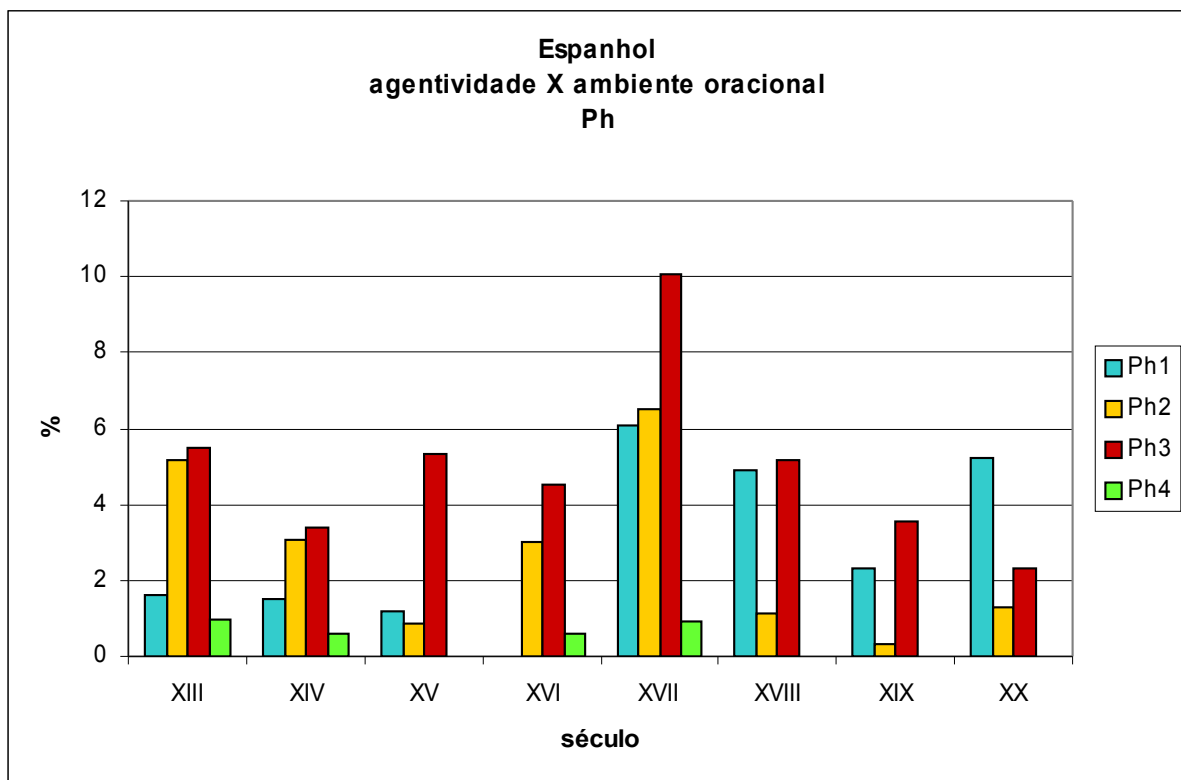


Gráfico E18a(T) – comparação entre os perfis de agentividade **Ph** nos quatro ambientes oracionais propostos.

Os polinômios que resolvem a distribuição das freqüências de Ph corroboram as observações feitas ao Gráfico E18a e acrescentam novas informações. Veja-se que a curva de Ph1, de fato, intercepta e ultrapassa a de Ph2 muito antes de chegar à de Ph3, e que esta última se mantém acima das demais até o início do século XX. Os máximos para as três curvas mais expressivas estão entre os séculos XVII e XVIII.

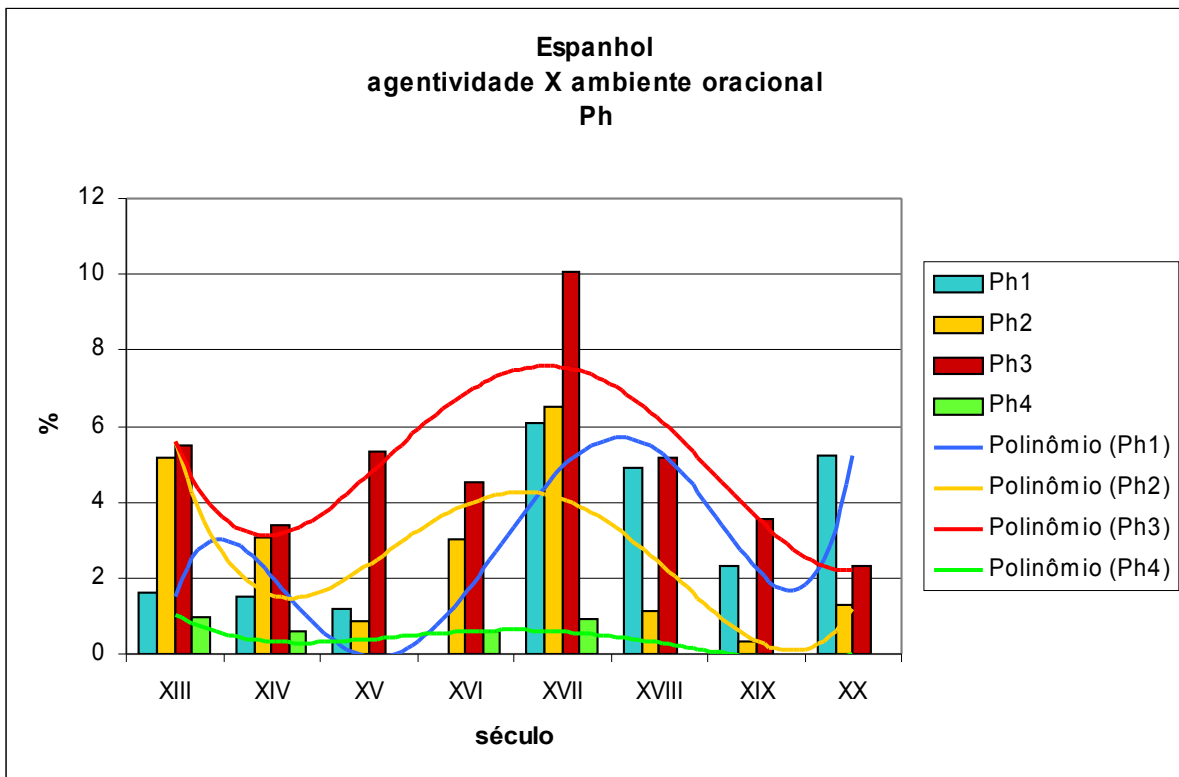


Gráfico E18b(T) – comparação entre os perfis de agentividade **Ph** nos quatro ambientes oracionais propostos.

A normalização das curvas apenas magnifica os perfis já observados no Gráfico anterior, E18b. Entretanto, mais uma vez, vê-se que o máximo da curva correspondente às coordenadas precede o das subordinadas, e que este guarda um intervalo de quase um século para o máximo de Ph1; repete-se aqui, portanto, a ordem em que os pontos mediais das curvas se apresentam.

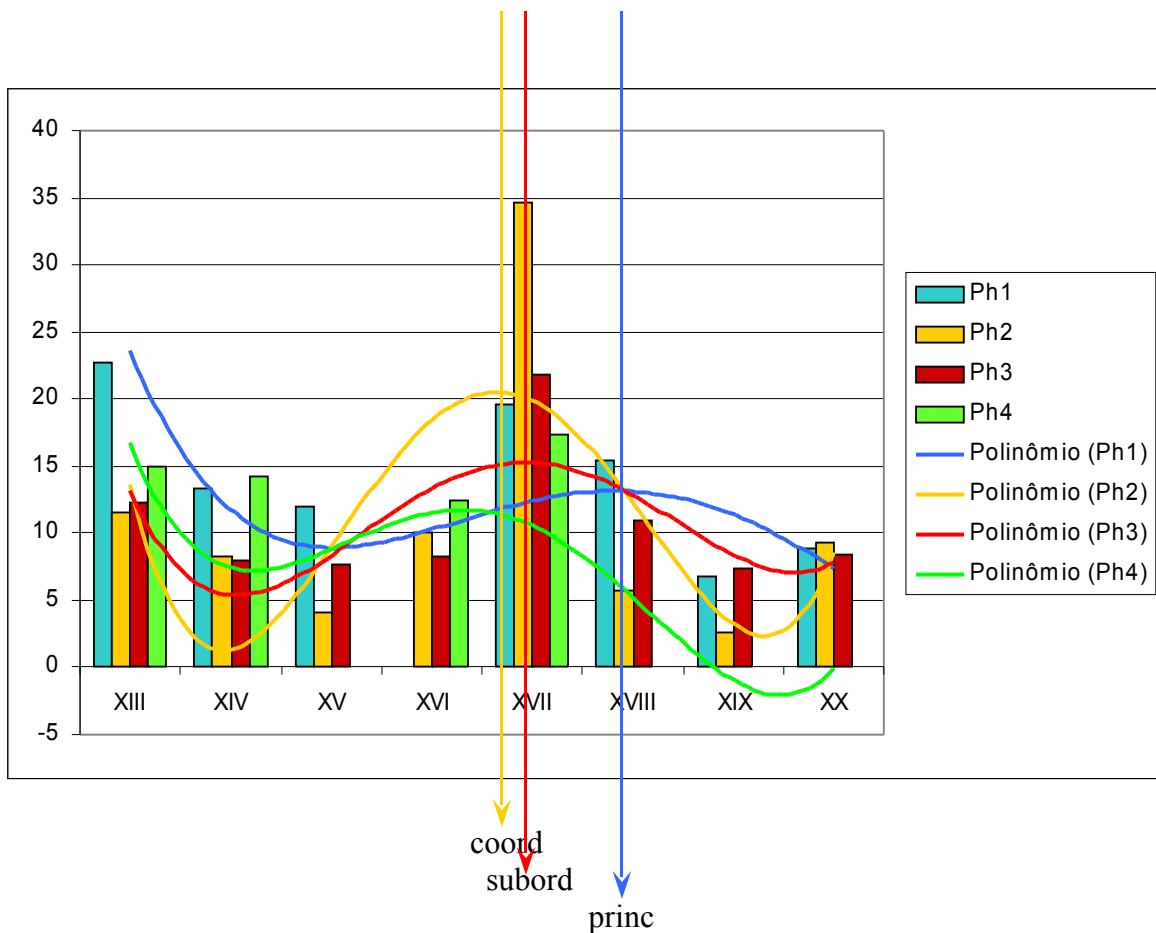


Gráfico E18c(T, normalizado) – comparação entre os perfis de agentividade **Ph** nos quatro ambientes oracionais propostos.

O traço Pn, por fim, é um dos mais comuns no *corpus* analisado – daí a escala ter uma amplitude maior do que a que cabe às duas classes anteriores. A tendência crescente observada para o grupo Pn1 (iRRL = 3,48), diante da inclinação negativa da RRL(Pn3) (1,09) contribui para que ambas se cruzem, como nas classes anteriores, no século XX. Ainda uma vez, Pn3 é o grupo mais freqüente até o século XIX, e as construções paratáticas associadas a Pn (iRRL = -1,92) tendem a ser menos comuns do que as orações principais já no século XVII.

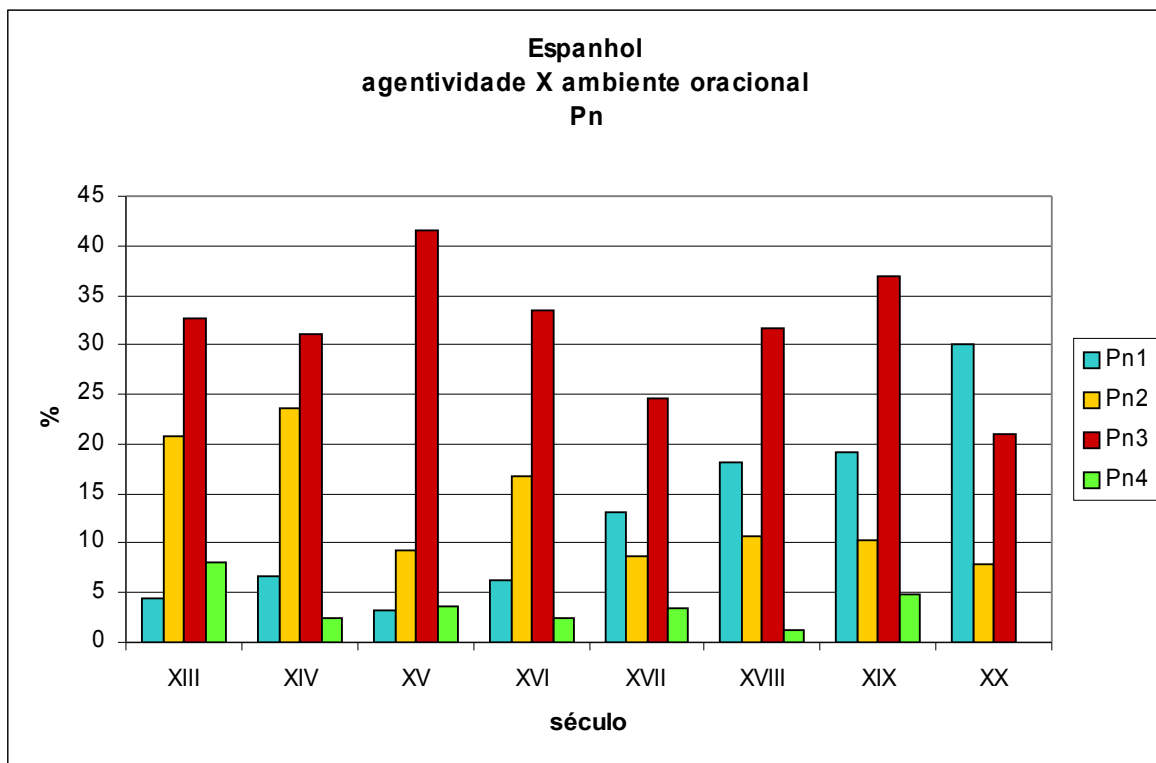


Gráfico E19a(T) – comparação entre os perfis de agentividade **Pn** nos quatro ambientes oracionais propostos.

A inclusão das curvas que expressam o comportamento dos conjuntos de ocorrências sugere um padrão muito semelhante aos das classes anteriores, em que a) a curva de Pn3 domina o gráfico até o século XIX, mas é suplantada pela curva de Pn1 no século XX; b) a curva de Pn2 confirma o exposto no Gráfico E19a, segundo o qual, apesar da flutuação dos valores, já se divisa a tendência decrescente desse conjunto; a curva crescente de Pn1 cruza a de Pn2 no século XVI; e c) a magnitude das frequências de Pn4 é invariavelmente baixa. Esses resultados mostram que o fator *ambiente oracional* (cf. Gráfico E7 acima) é determinante dos perfis construídos a partir de outros critérios de classificação das ocorrências – ou seja, fatores como *agentividade* são dependentes de *ambiente oracional*.

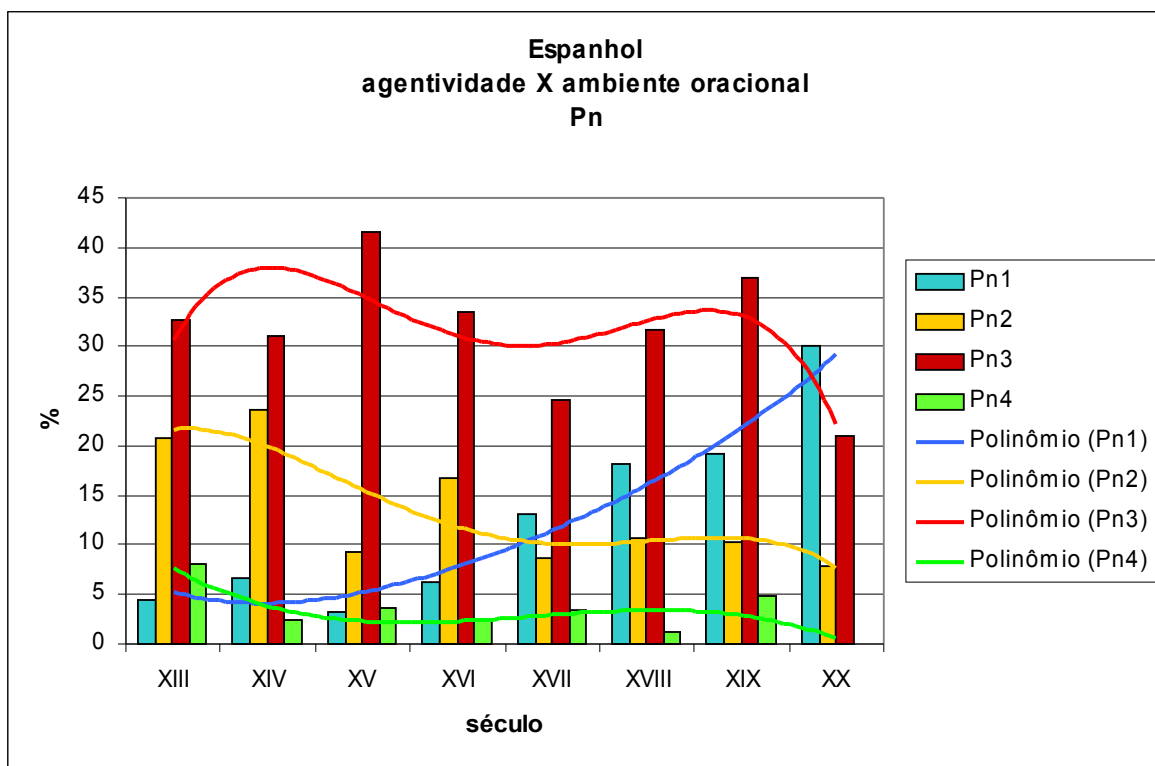


Gráfico E19b(T) – comparação entre os perfis de agentividade **Pn** nos quatro ambientes oracionais propostos.

O procedimento de normalização, para o caso das ocorrências de Pn, leva a resultados instigantes. Observe-se, em primeiro lugar, o fato de que as curvas relativas aos três conjuntos mais freqüentes apresentam mínimos locais compreendidos no século XVI; contudo, a seqüência em que se ordenam é exatamente oposta à das classes anteriores – i.e., para Pn, os mínimos mediais seguem a ordem *principais* (Pn1) > *subordinadas* (Pn3) > *coordenadas* (Pn2). Isso mostra que a classe Pn tem um comportamento distinto das demais, segundo o qual o início do avanço dessa classe ocorreu, aqui, nas orações principais, ao passo que, nas outras, começou pelas coordenadas.

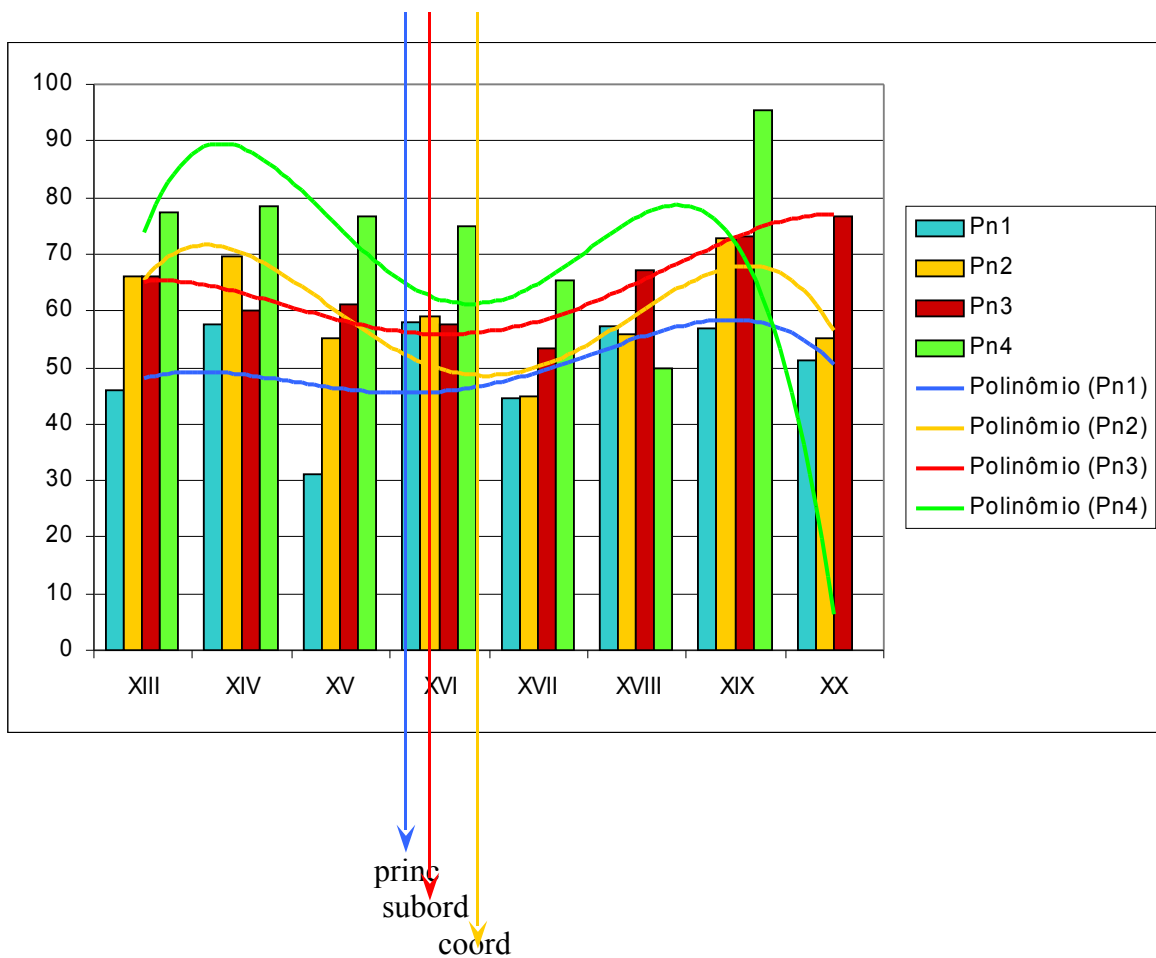


Gráfico E19c(T, normalizado) – comparação entre os perfis de agentividade **Pn** nos quatro ambientes oracionais propostos.

O fator *peso de S e de C* indica a predominância das classes **k** e **n** (respectivamente, sujeito pleno e complemento verbal pleno), que apresentam tendências distintas: enquanto a primeira aumenta em frequência, a reta de regressão linear da segunda tem inclinação levemente negativa. Diminuem também as frequências das classes de sujeito pronominal, **l**, e de complemento verbal pronominal, **o**. As classes **m** e **p**, que correspondem aos argumentos oracionais, serão esparsamente tratadas, em decorrência de suas frequências serem muito baixas (de ordinário, menores que 5%).

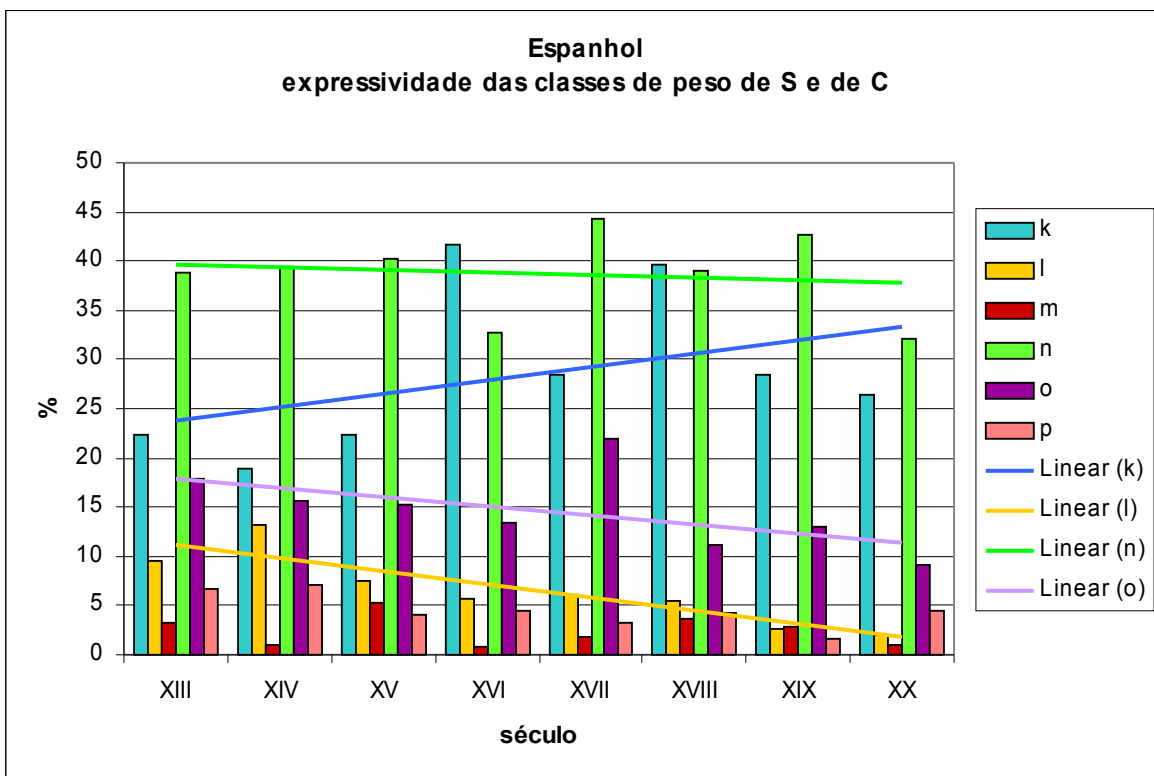


Gráfico E20 - perspectiva geral da distribuição relativa das classes de *peso de S e de C* no intervalo de amostragem.

A classe **k** – sujeito pleno – mostra especialização nas ordens a (SVC), g (SV) e h (VS), como se pode observar no Gráfico E21a. Das três, a única que tem suas freqüências reduzidas no intervalo de amostragem é SV (iRRL (SV) = -0,55), enquanto as freqüências das outras duas aumentam (iRRL (SVC) = 0,44; iRRL (VS) = 1,92). As freqüências das demais ordens flutuam em torno dos 5% ou menos.

São exemplos de ocorrências classificadas nesse grupo:

que los fijos de Cam e de Japhet ganaron alguna cosa en Asia por fuerça [CGE]

Et las gentes extrañas deseaban que diesse Dios razon porque [CAB]

Mi madre había hecho particiones de la hacienda y sacado su dote [CON]

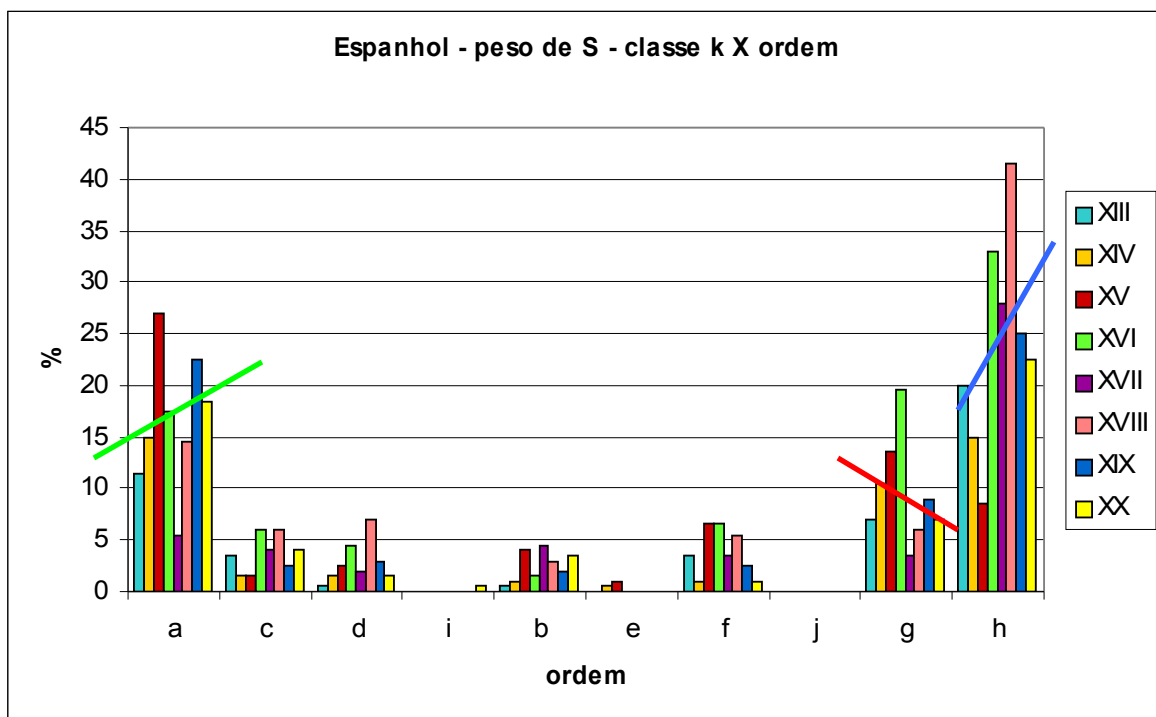


Gráfico E21a - perspectiva geral da distribuição relativa das ocorrências de sujeito pleno em função da ordem de constituintes.

Apesar de ocorrer em frequência absoluta mais modesta do que a classe **k**, o sujeito pronominal, **I**, mostra a mesma preferência distribucional, concentrando-se nas ordens SVC, SV e VS. Devido à natureza distinta dos argumentos **k** e **I**, as retas de regressão linear desse gráfico têm comportamentos diferentes ($iRRL(SVC) = -0,41$; $iRRL(SV) = -0,83$; e $iRRL(VS) = -0,70$) dos encontrados no anterior, como seria de se esperar.

Correspondem a exemplos associados a essa classe:

Yo nació entre las cortaduras de papel y los rollos del pergamino [TOR]
cuando fundaba yo mi voto y explicaba las intenciones del Gobierno [PAZ]
como si yo voy al mercado non con intención de ronper çapatos [FOR]

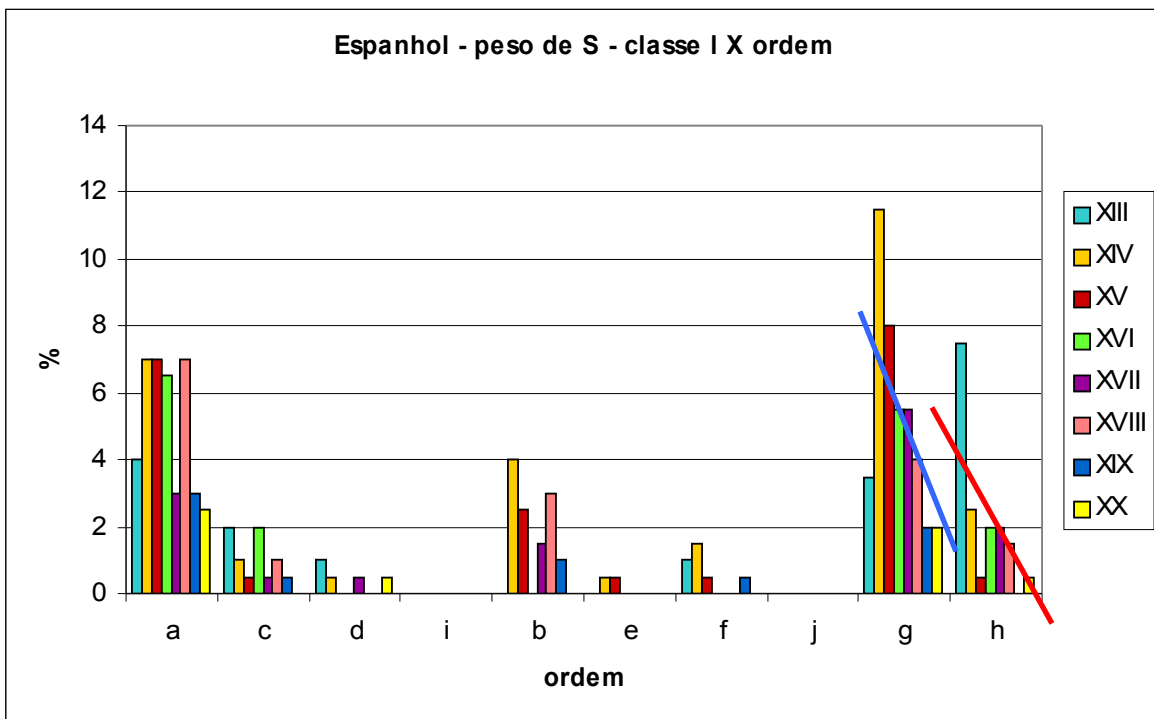


Gráfico E21b - perspectiva geral da distribuição relativa das ocorrências de sujeito pronominal em função da ordem de constituintes.

No caso dos complementos verbais plenos (classe **n**), as ocorrências tendem a se concentrar preferencialmente na ordem **i** (VC), e, em menor escala, em **a** (SVC). Contudo, apesar de o subconjunto **a** (SVC) ser representado por reta de regressão linear de inclinação positiva ($iRRL = 0,76$), a flutuação² de **i** (VC) retorna um coeficiente bastante baixo para a inclinação dessa RRL (-0,11).

Exemplos dessa classe podem ser encontrados abaixo:

despiértase en el espíritu deseo ardiente de penetrar su vida entera [CAS]
 é demas dióle Sancta María del Puerto por heredad con tal condicion [SAN]
 para obligar al Rey Católico a dejar la España y contentarse [FEL]

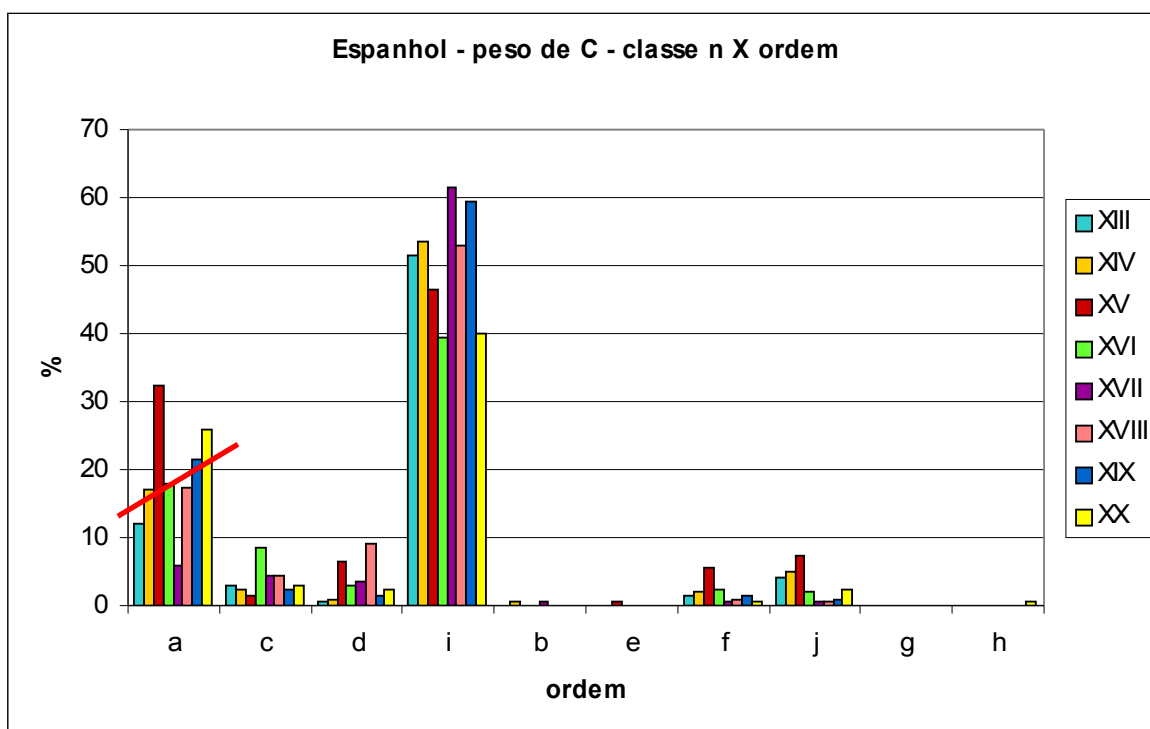


Gráfico E21c - perspectiva geral da distribuição relativa das ocorrências de complemento verbal pleno em função da ordem de constituintes.

² É possível que a expansão do *corpus* leve a uma definição mais clara da tendência desse subgrupo de ocorrências.

É interessante notar, para os complementos verbais pronominais (classe **o**), uma orientação que privilegia francamente as ordens de sujeito não-expresso, **i** (VC) e **j** (CV); proporções mais reduzidas de ocorrências com esse tipo de complemento verbal são representadas pelas ordens **b** (SCV) e **f** (VCS) – ambas na faixa dos 5%. Para as ordens mais comuns, tem-se $iRRL(VC) = -0,68$ e $iRRL(CV) = -0,88$; em outras palavras, o complemento verbal pronominal tende a ser cada vez menos utilizado nas ordens em que até o momento é mais freqüente.

Encontram-se, abaixo, exemplos dessa classe:

segun lo certifican sant Juan Evangelista, y mas largamente Moisen [IND]
arrullan el oído y parece que convidan al espíritu a seguirlas en su ... vuelo [CAS]
la mayoría de los casos los ganaba contra todo pronóstico [PROD]

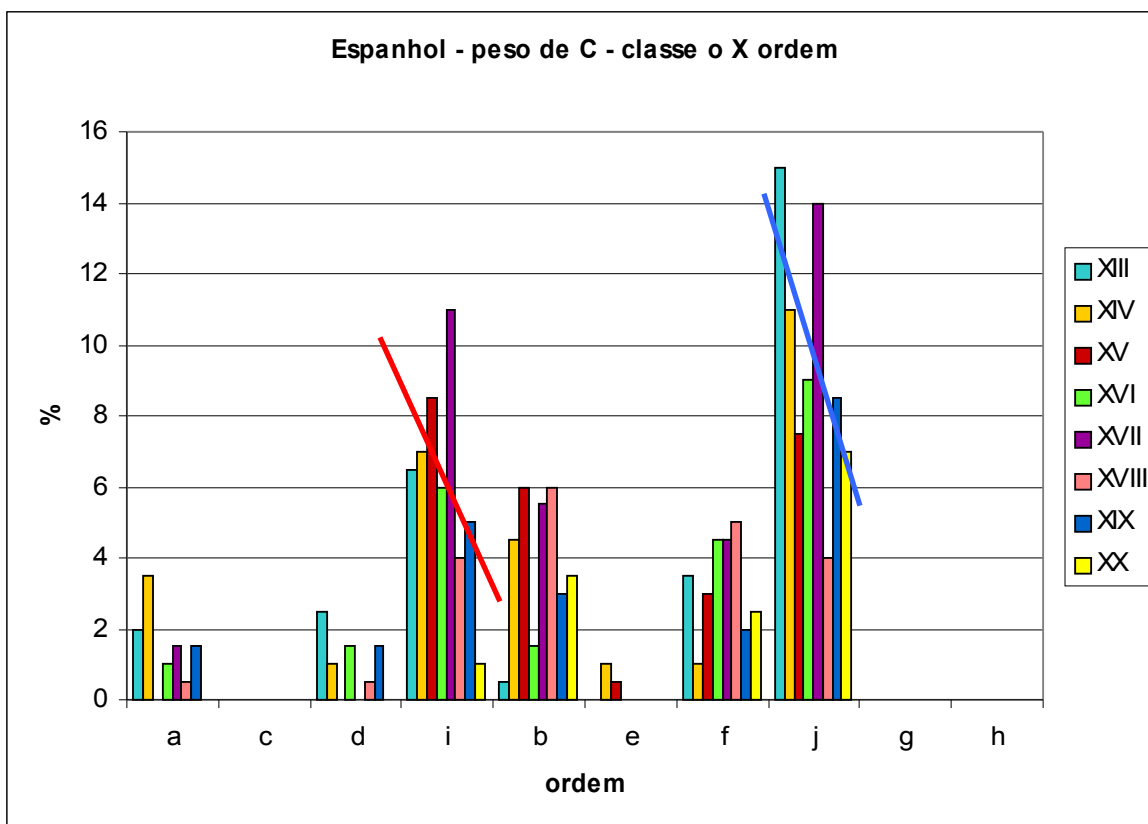


Gráfico E21d - perspectiva geral da distribuição relativa das ocorrências de complemento verbal pronominal em função da ordem de constituintes.

Se se fixa outro dos graus de liberdade previstos na classificação das ocorrências – especificamente, o *ambiente oracional* –, observam-se padrões semelhantes aos já discutidos para os critérios anteriores. Veja-se, inicialmente, o caso das orações principais, em que há aumento das frequências das classes mais comuns, **k** (iRRL = 3,44) e **n** (iRRL = 2,67). Isso significa que, nas orações principais, as ocorrências com sujeito e com complemento verbal plenos tendem a ser cada vez mais frequentes.

Tomem-se as ocorrências abaixo como exemplos de cada classe:

k	Mi madre cuenta todavía algunas niñadas de aquel tiempo [TOR]
	Pusiéronla sus fundadores sobre las rocas [CAS]
n	Hizo el Altísimo este siglo para muchos; y el otro, que es la gloria, para pocos [IND]
	No pierde tiempo el que gasta tres ó cuatro años entre los Horacios [TOR]

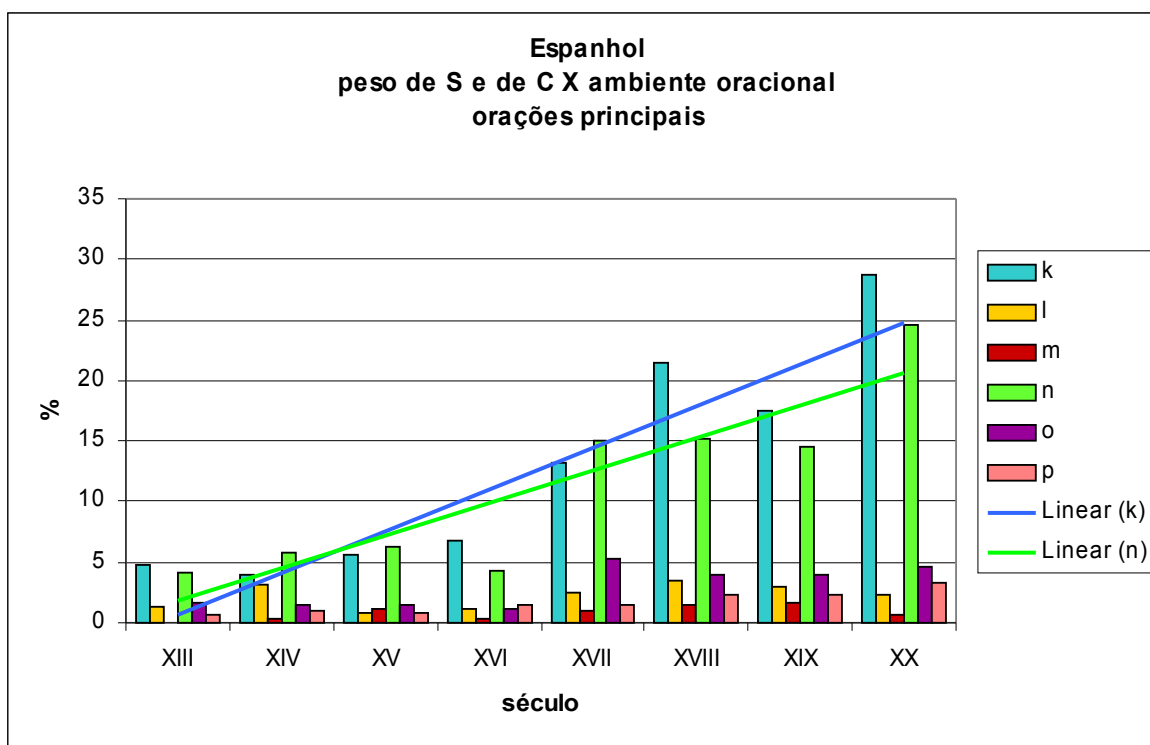


Gráfico E22a - distribuição relativa das classes de *peso de S e de C X* em orações principais.

Para as orações coordenadas, como se pode prever a partir do exame do Gráfico E7 e do comportamento das categorias anteriores, as RRLs das classes mais comuns têm inclinações de valor negativo: $iRRL(k) = -0,61$; $iRRL(n) = -1,25$; e $iRRL(o) = -0,70$. Portanto, ainda que a escala das coordenadas seja menor que a das orações principais, é possível identificar uma tendência geral decrescente para o grupo.

Tomem-se as ocorrências abaixo como representativas dessa classe:

k	y por esto la venerable matrona la crió con más afectuoso cuidado [AGR]
	y enfrente duerme tendida la inmensidad del Océano [CAS]
n	les sufoca las inconsideraciones y modera aun las inculpables altanerías [TOR]
	los fiscales tiraban al suelo los códigos y las mucetas [PROD]
o	pero los franceses lo gobernaban de forma que se hallase obligado a dejar [FEL]
	e besauanlo todos e abraçauanlo con muy grand alegría [TRO]

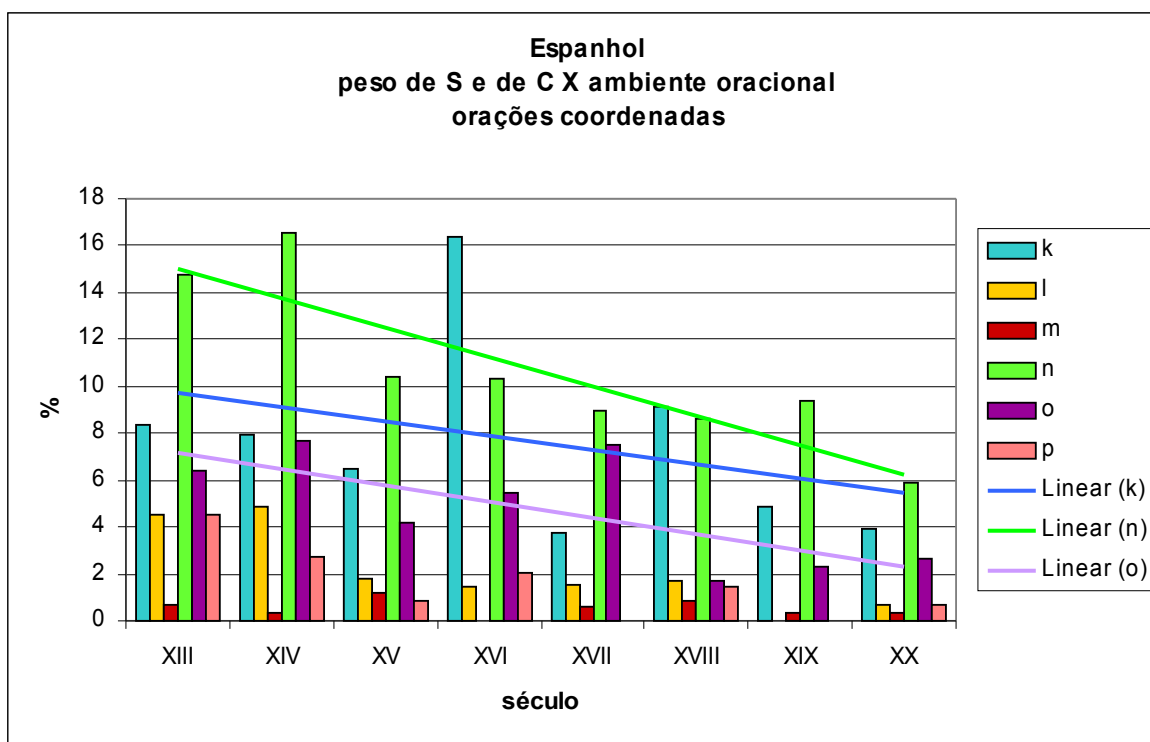


Gráfico E22b - distribuição relativa das classes de *peso de S e de C* em orações coordenadas.

As orações subordinadas comportam-se como as demais classes analisadas diante do fator *ambiente oracional*: há um máximo local em torno do século XV, mas a tendência do conjunto é decrescente ($iRRL(\mathbf{k}) = -1,11$; $iRRL(\mathbf{l}) = -1,02$; $iRRL(\mathbf{n}) = -0,72$; $iRRL(\mathbf{o}) = -0,64$). Observe-se que as frequências gerais também são mais elevadas do que as das classes anteriores.

Exemplos desse subconjunto de ocorrências são encontrados na Tabela abaixo:

k	que Su Divina Providencia tenía determinado hacer en esta casa [AGR] en la cual las lluvias han señalado bucles esfumados de oro [SIL]
l	los sucesos que ellos solos conocen [PAZ] del siglo XIX nadie discutía la igualdad de todos los hombres [PROD]
n	recibiendo en sus calidades muchas advertencias para la dirección [AGR] en los que el sueño va tendiendo sus cadenas melosas e invisibles [SIL]
o	mas vengarla en tal guisa aquel a quien la faz [TRO] ca por generación la ovieron, e generación, segund Damasceno, obra es natural [FOR]

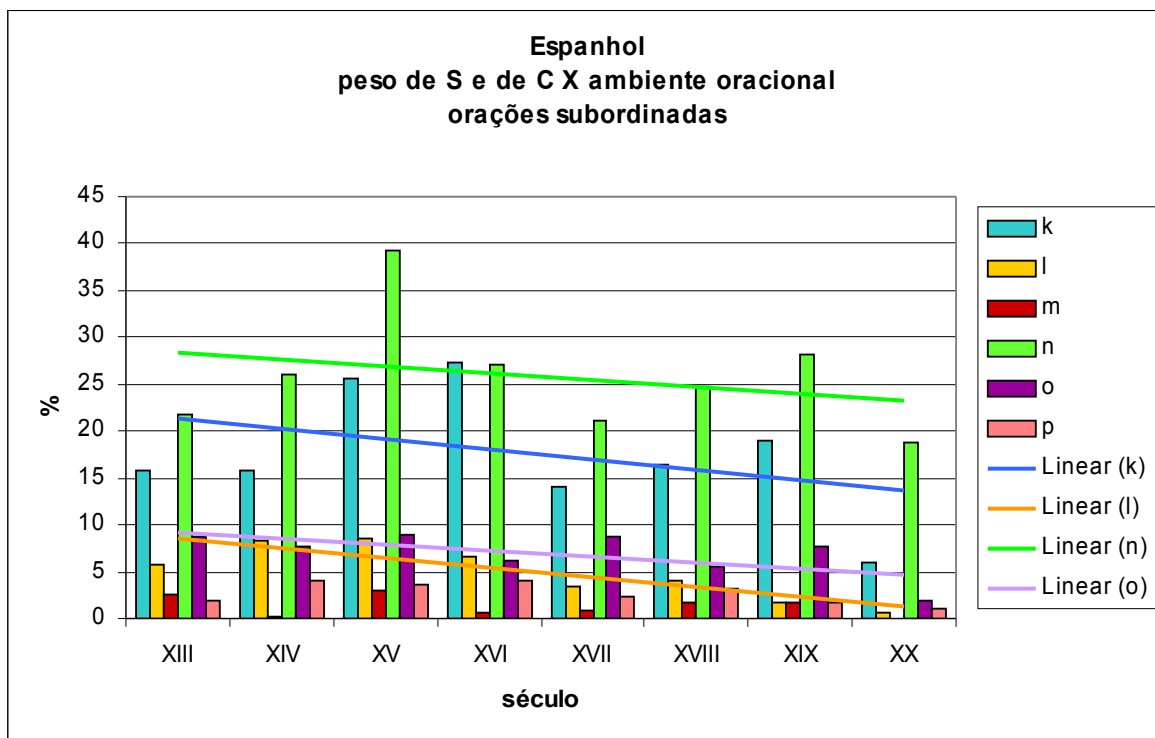


Gráfico E22c - distribuição relativa das classes de *peso de S e de C X* em orações subordinadas.

Nas Tabelas E4 e E5, sumarizam-se os comportamentos delineados até agora para o cruzamento do fator *peso de S e de C* com os critérios de análise *ordem de constituintes e ambiente oracional*. Nota-se claramente que, apesar de haver algumas tendências definidas na Tabela E4, os padrões de comportamento das ocorrências são muito mais sensíveis ao *ambiente oracional* do que à *ordem*.

<i>peso</i>	<i>a (SVC)</i>	<i>i (VC)</i>	<i>j (CV)</i>	<i>g (SV)</i>	<i>h (VS)</i>
<i>k (sujeito pleno)</i>	↗	.	.	↘	↗
<i>l (sujeito pronominal)</i>	→	.	.	↘	↘
<i>n (complemento pleno)</i>	↗
<i>o (complemento pronominal)</i>	.	↘	↘	.	.
<i>p (complemento oracional)</i>	.	↘	.	.	.

Tabela E4 – resumo das tendências de distribuição das ocorrências segundo os critérios *ordem de constituintes* e *peso de S e de C*.

<i>peso</i>	<i>1 or. principais</i>	<i>2 or. coordenadas</i>	<i>3 or. subordinadas</i>	<i>4 or. coord.-subord</i>
<i>k (sujeito pleno)</i>	↗	↘	↘	.
<i>n (complemento pleno)</i>	↗	↘	↘	↘
<i>o (complemento pronominal)</i>	.	↘	↘	.

Tabela E5 – resumo das tendências de distribuição das ocorrências segundo os critérios *ambiente oracional* e *peso de S e de C*.

Da mesma forma que se fez para os traços de *agentividade*, propõe-se, adiante, a unificação dos resultados de *peso de S e de C*, reduzindo-se-os à mesma escala; esse procedimento permitirá observar as relações entre os comportamentos das ocorrências nos quatro ambientes oracionais aqui estudados. Fazem-se acompanhar esses gráficos de um terceiro para cada traço, em que se normalizam as proporções em relação ao total de ocorrências por texto para cada ambiente – e não para o total de ocorrências de cada texto.

Observemos, em primeiro lugar, o que se passa com o traço **k**. O Gráfico E23a mostra que os sujeitos plenos sempre passam a ser a categoria mais comum em orações principais a partir do século XVII – portanto, antes da mudança observada para a maioria dos gráficos de *agentividade X ambiente oracional*, em que as subordinadas predominam na primeira metade do gráfico, e as principais ganham mais espaço entre os séculos XIX e XX.

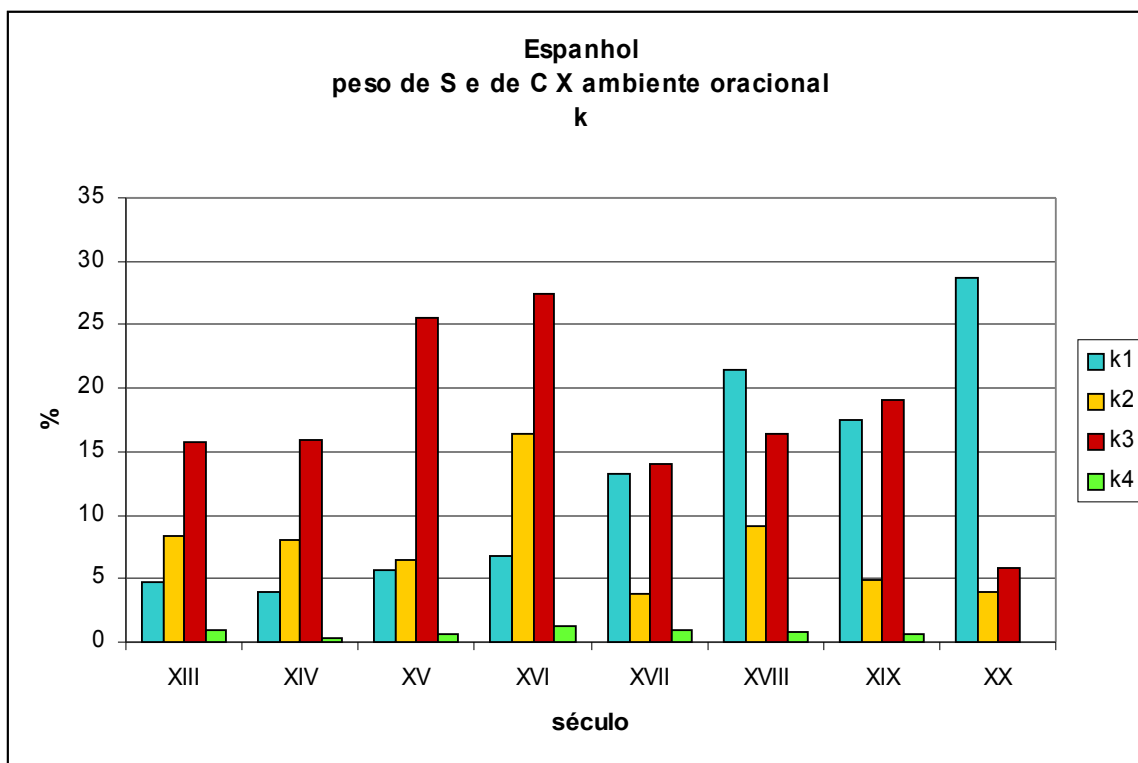


Gráfico E23a(T) – comparação entre os perfis de sujeito pleno nos quatro ambientes oracionais propostos.

No Gráfico E23b, observa-se mais nitidamente o franco e seguro progresso das frequências de sujeito pleno, **k**, nas orações principais (curva azul, **k1**), o que se faz acompanhar de um decréscimo nas proporções relativas de sujeito pleno em subordinadas (curva vermelha, **k3**); a curva descrita pelas coordenadas (**k2**) mostra pequena flutuação, com máximo local no século XVI. Nessa escala, é praticamente inconspícuo o comportamento de **k4** (sujeito pleno em orações coordenadas-sunordinadas).

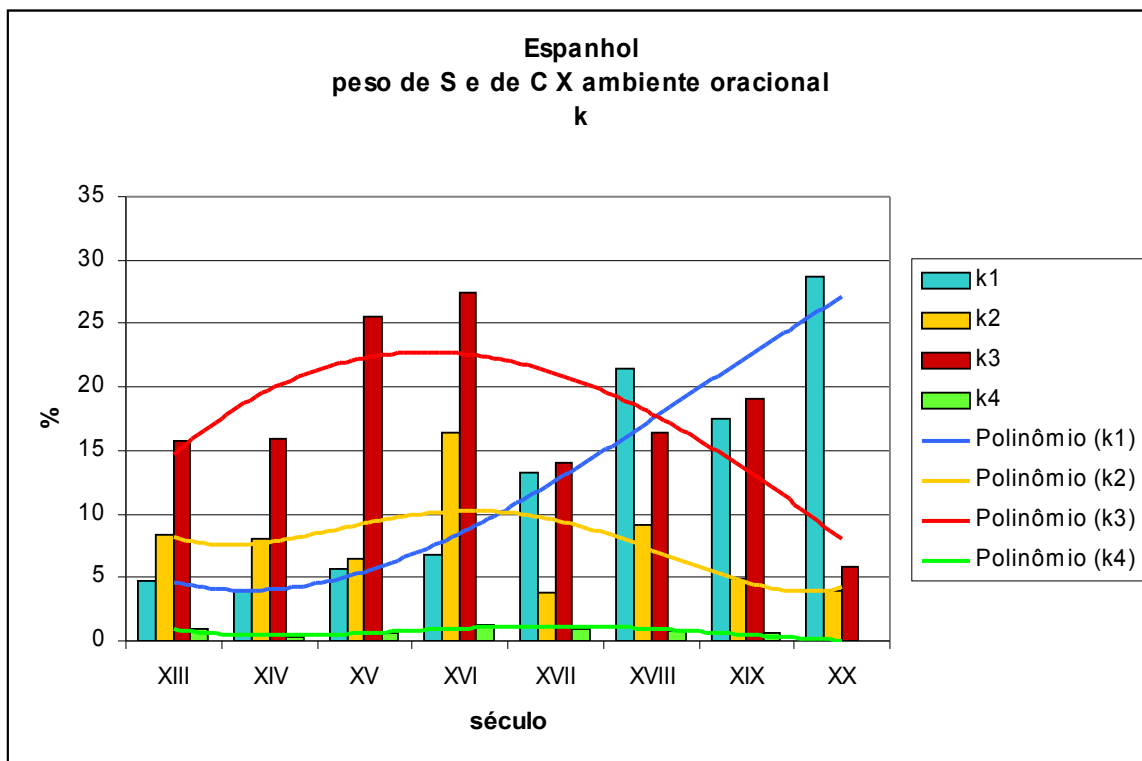


Gráfico E23b(T) – comparação entre os perfis de sujeito pleno nos quatro ambientes oracionais propostos.

É útil compararem-se os gráficos anteriores com o Gráfico E23c, que mostra as proporções de **k** nos diversos ambientes oracionais em relação à frequência desse traço em cada texto. É interessante notar que os máximos mediais das curvas correspondentes aos três grupos mais comuns são praticamente coincidentes – todos com abscissa no século XVII. Entretanto, se, por um lado, a distância entre a reta amarela e as outras duas é pequena, a que se guarda entre as retas vermelha (**k3**) e azul (**k1**) já é da mesma ordem que a observada nos gráficos para *agentividade X ambiente oracional* – i.e., há mais segurança na afirmação de que o máximo da curva das subordinadas precede o das principais para o traço **k**.

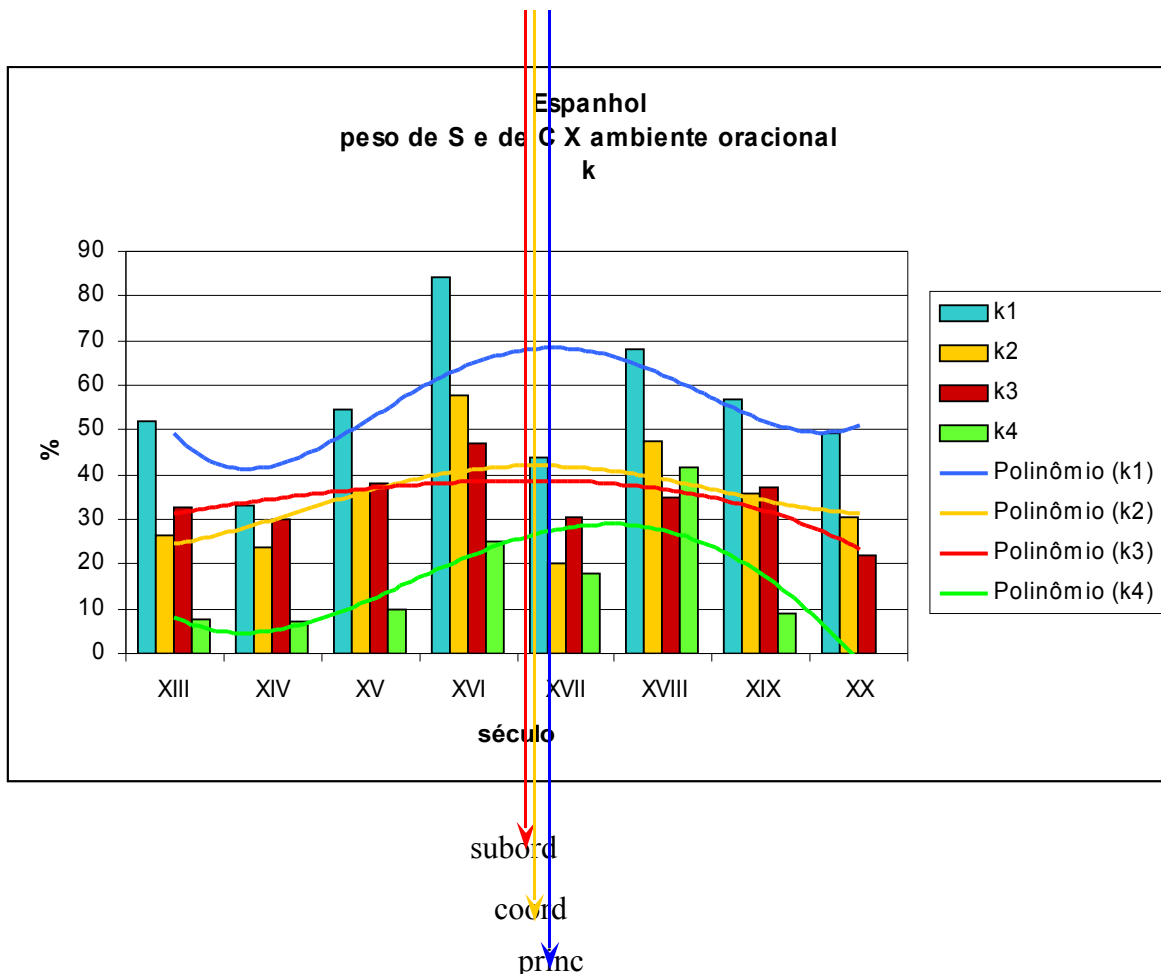


Gráfico E23c(T, normalizado) – comparação entre os perfis de sujeito pleno nos quatro ambientes oracionais propostos.

O traço seguinte, dentre os mais comuns, é **I** – sujeito pronominal –, cujas freqüências são mais baixas que as de **k**, mas correspondem a valores diferentes de zero para as ordens mais expressivas; essa característica contribui para a definição dos perfis observados nos Gráficos E24, e que são semelhantes aos dos Gráficos precedentes de mesma natureza. Em outras palavras, ainda que as médias de **I** sejam mais reduzidas, pode-se perceber a diminuição das freqüências absolutas de coordenadas e de subordinadas, bem como o aumento discreto das proporções de orações principais.

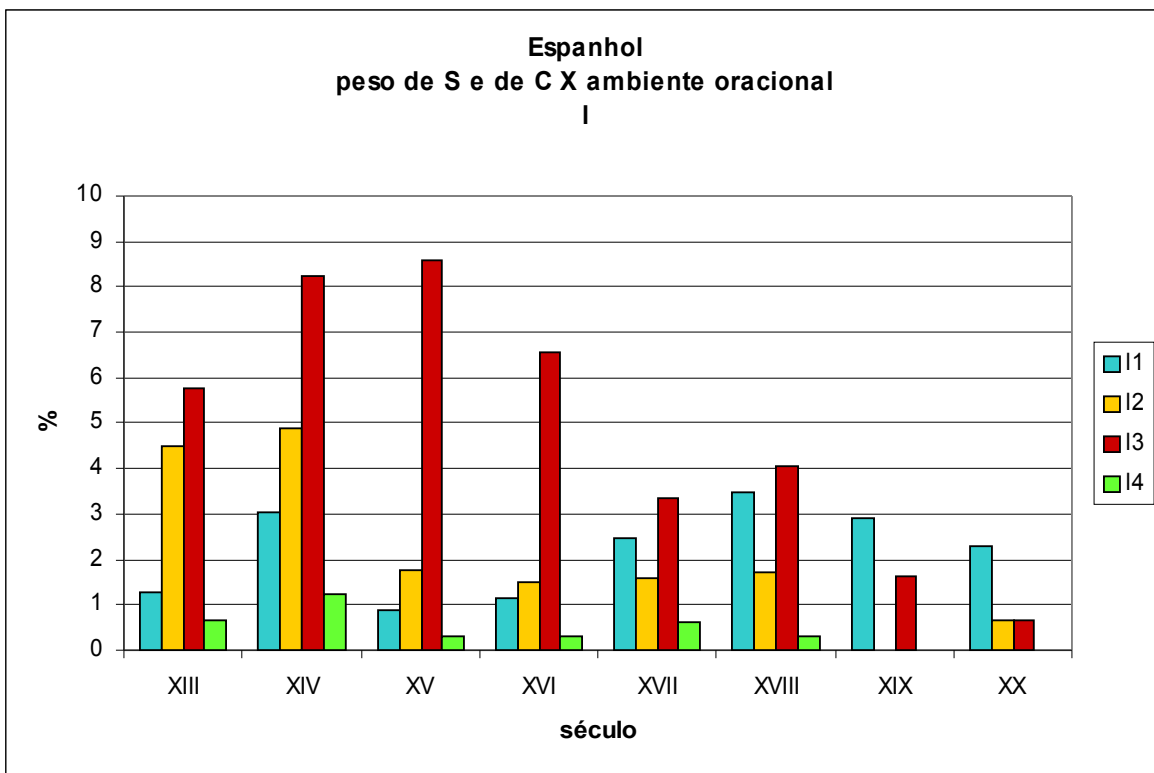


Gráfico E24a(T) – comparação entre os perfis de sujeito pronominal nos quatro ambientes oracionais propostos.

Os perfis da figura anterior podem ser resolvidos por polinômios cujas expressões gráficas são as que constam da imagem abaixo. Também o sujeito pronominal mostra um comportamento semelhante aos das classes precedentes, notando-se que as orações principais associadas ao traço I tornam-se ligeiramente mais freqüentes que as subordinadas no século XVIII. Aliás, dada a reduzida escala em que essas curvas se apresentam, o único perfil de mudança nítida – da ordem dos 10% de amplitude – é o das subordinadas, **I3** (curva vermelha).

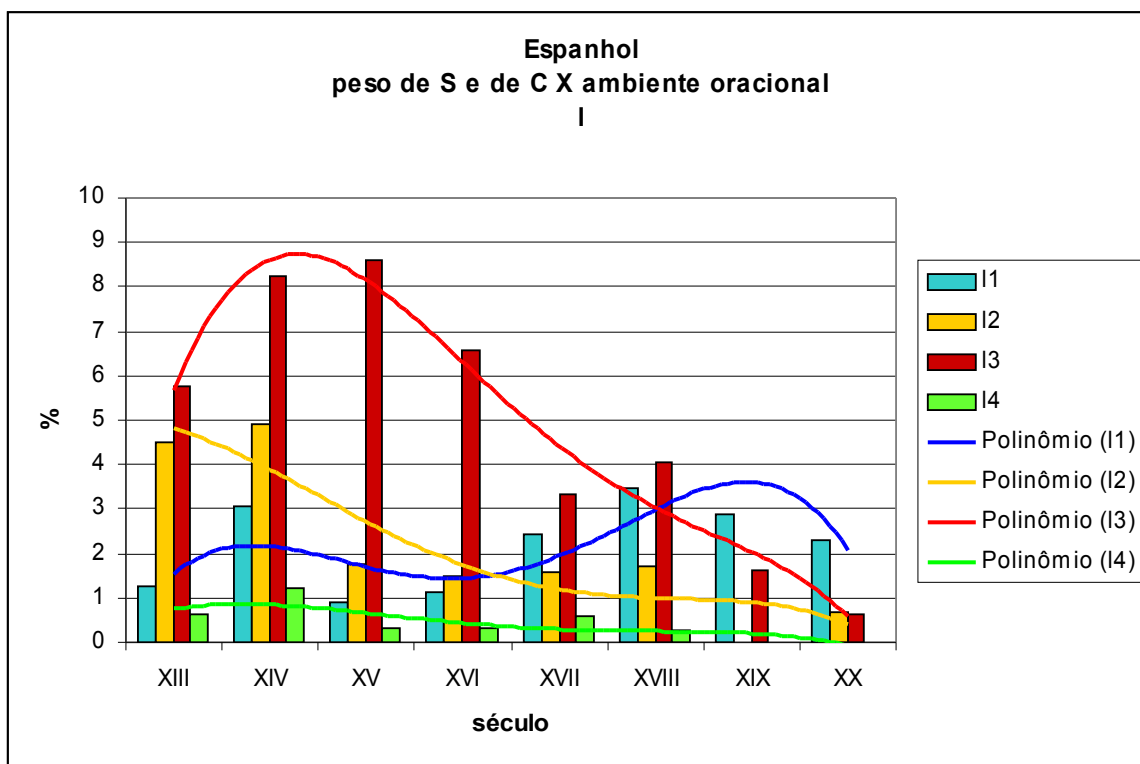


Gráfico E24b(T) – comparação entre os perfis de sujeito pronominal nos quatro ambientes oracionais propostos.

A normalização das frequências para o total de ocorrências por texto amplifica diferencialmente os padrões encontrados no Gráfico E24b, de maneira a aproximar as curvas. Entretanto, as duas únicas que mostram mínimos locais de maneira mais clara são **II** (azul), no início do século XVII, e **I4**(verde) – esta última, associada a frequências absolutas muito baixas.

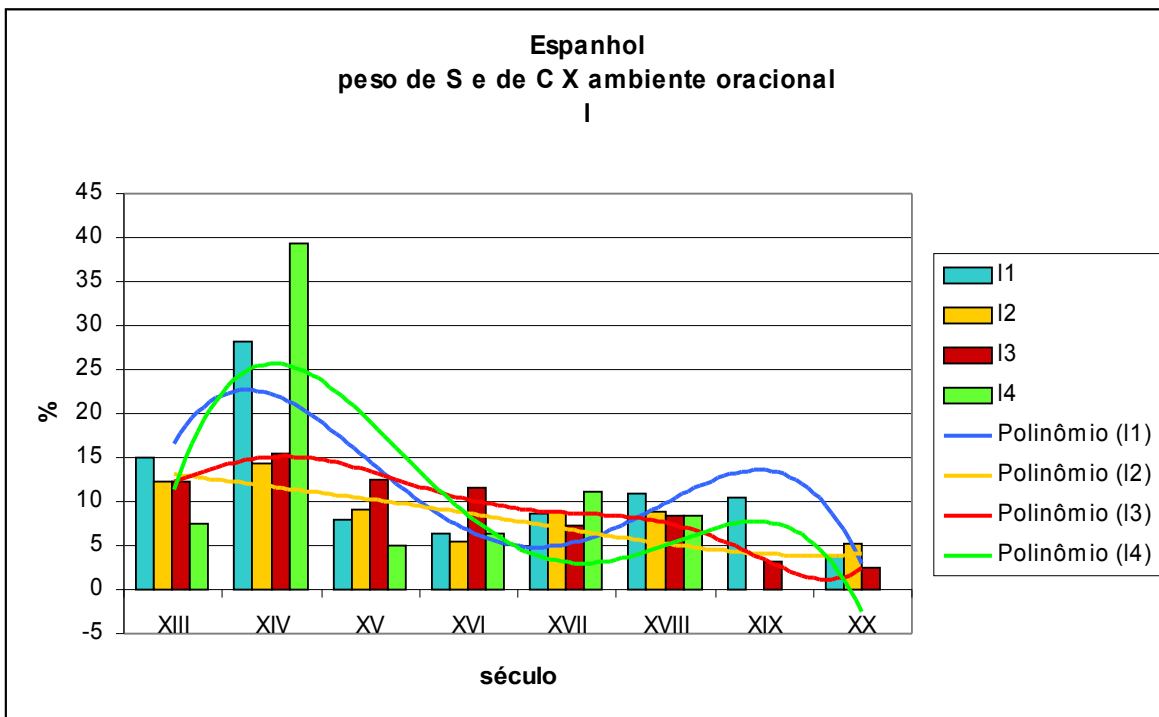


Gráfico E24c(T, normalizado) – comparação entre os perfis de sujeito pronominal nos quatro ambientes oracionais propostos.

Diante das reduzidíssimas freqüências absolutas associadas à classe **m** (sujeito oracional) e a subsequente existência de muitas lacunas no gráficos produzidos a partir desses dados, os perfis que representam essas freqüências absolutas não serão apresentados. Trataremos unicamente do Gráfico E25a, que mostra as curvas normalizadas para o traço **m** em todos os ambientes oracionais. Mais uma vez, o mínimo local da curva relativa às orações principais ocorre no século XVII.

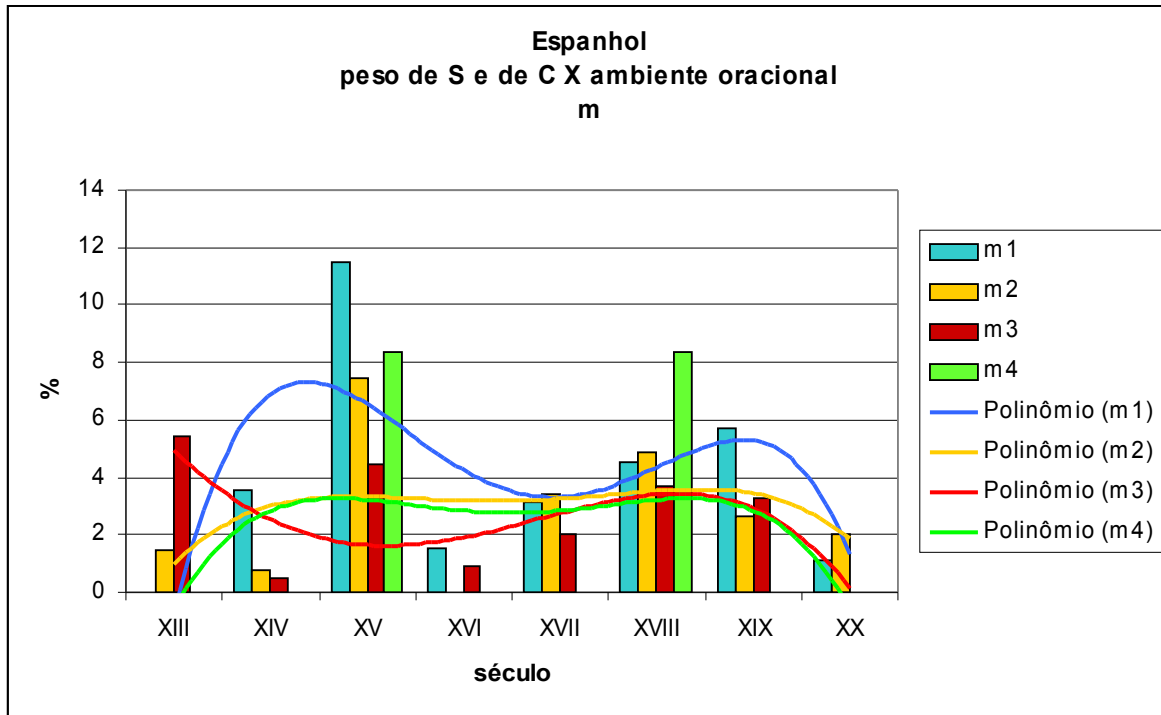


Gráfico E25a(T) – comparação entre os perfis de sujeito oracional nos quatro ambientes oracionais propostos.

O perfil de distribuição das ocorrências de complemento pleno, **n**, segundo o ambiente oracional segue, como observado para as categorias anteriores, o mesmo padrão geral das curvas de frequência dos traços de *ambiente oracional* (q.v. Gráfico E7). Notam-se, na figura abaixo, a) as taxas crescentes de **n** atribuídas a orações principais (no gráfico, **n1**, em azul), e que passam a ser dominantes no século XX, b) o decréscimo das proporções absolutas de **n** relacionado a orações coordenadas (**n2**, em amarelo), c) a flutuação de **n3** (em vermelho), que é suplantada por **n1** no século XX, e d) as frequências relativamente baixas de complemento pleno associado a orações coordenadas em subordinação, **n4** (em verde). A diferença óbvia entre os gráficos das duas classes anteriores e E23a está na escala das ordenadas, que assinala a frequência mais expressiva da classe **n** (ligeiramente mais alta que a de **k** – Gráfico E23a – devido ao máximo local e absoluto de **n** no século XV).

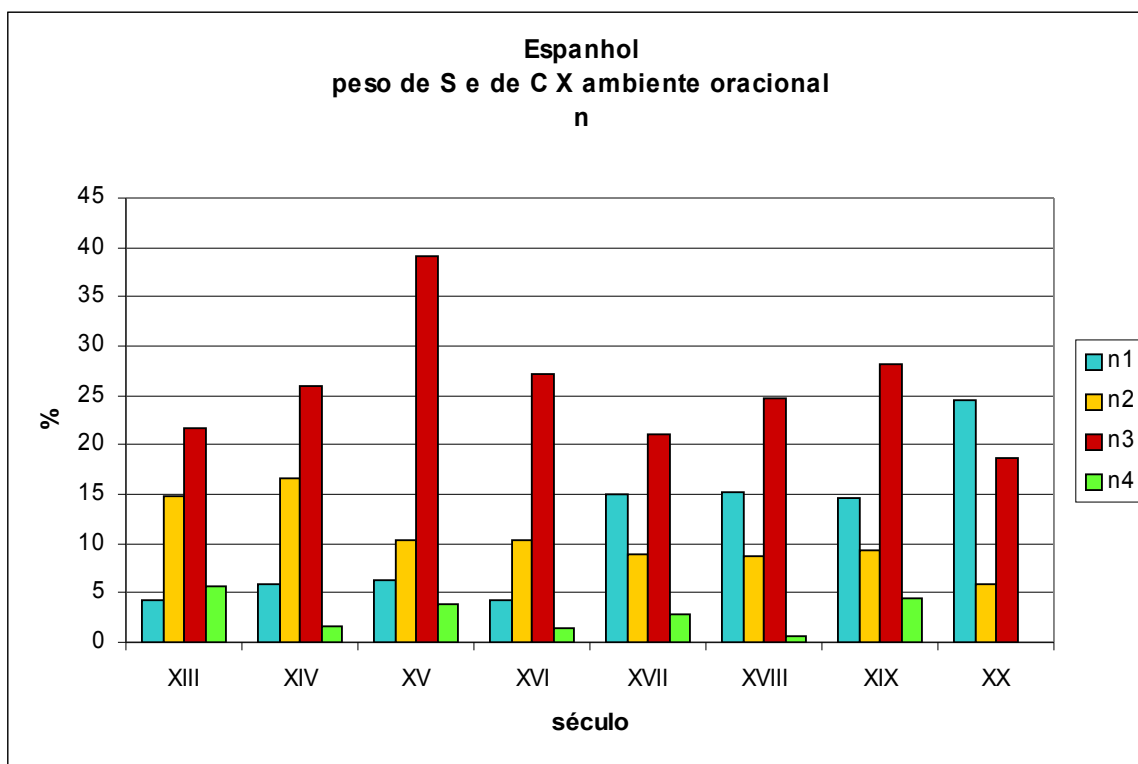


Gráfico E26a(T) – comparação entre os perfis de complemento verbal pleno nos quatro ambientes oracionais propostos.

Os polinômios que encompassam as distribuições das ocorrências da classe **n** repetem o mesmo padrão já identificado para outras categorias que envolvem o fator *ambiente oracional* – aumento das freqüências associadas a orações principais (**n1**), diminuição menos acentuada das freqüências ligadas a orações coordenadas (**n2**) e flutuação com decréscimo das freqüências de subordinadas (**n3**). O nível de 5% de freqüência absoluta raramente é ultrapassado pelas orações coordenadas em subordinação, **n4**.

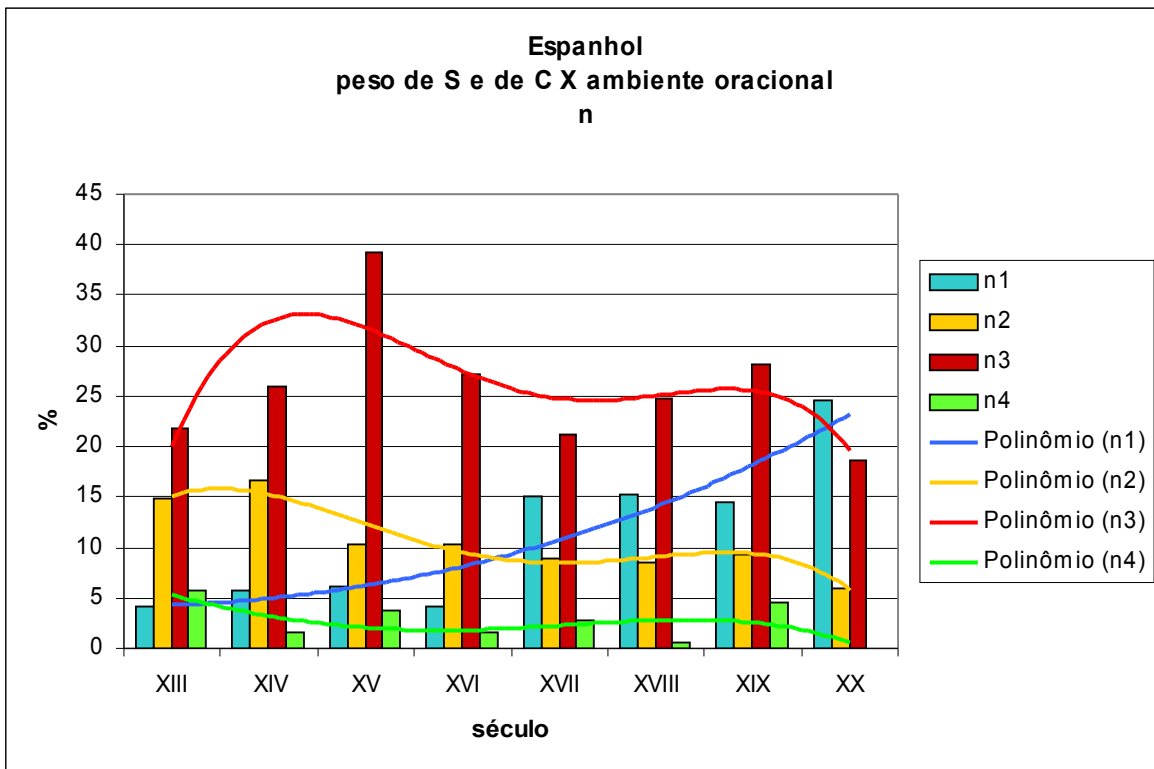


Gráfico E26b(T) – comparação entre os perfis de complemento verbal pleno nos quatro ambientes oracionais propostos.

Graças à normalização das freqüências, podem-se comparar os padrões individuais de comportamento das curvas apresentadas no gráfico anterior. Observam-se, aqui, as relações temporais entre aspectos semelhantes dos diferentes subconjuntos de ocorrências, como a seqüência dos mínimos locais nos séculos XVI e XVII: a curva que representa os complementos verbais plenos em orações coordenadas, **n2**, tem mínimo no final do século XVI, o qual é seguido pelos mínimos de orações principais, **n1**, e de subordinadas, **n3**. Essa seqüência é diferente da obtida para a classe **k** (sujeito pleno, Gráf. E23c) – a outra classe dominante no fator *peso de S e de C* –, a qual mostra quase uma sobreposição dos mínimos dessas curvas. Portanto, as duas categorias mais freqüentes do traço *peso de S e de C*, **k** (sujeito pleno) e **n** (complemento verbal pleno) têm comportamentos distintos diante do ambiente oracional – apesar de as curvas de ambas mostrarem mínimos entre o final do século XVI e a metade do século XVII.

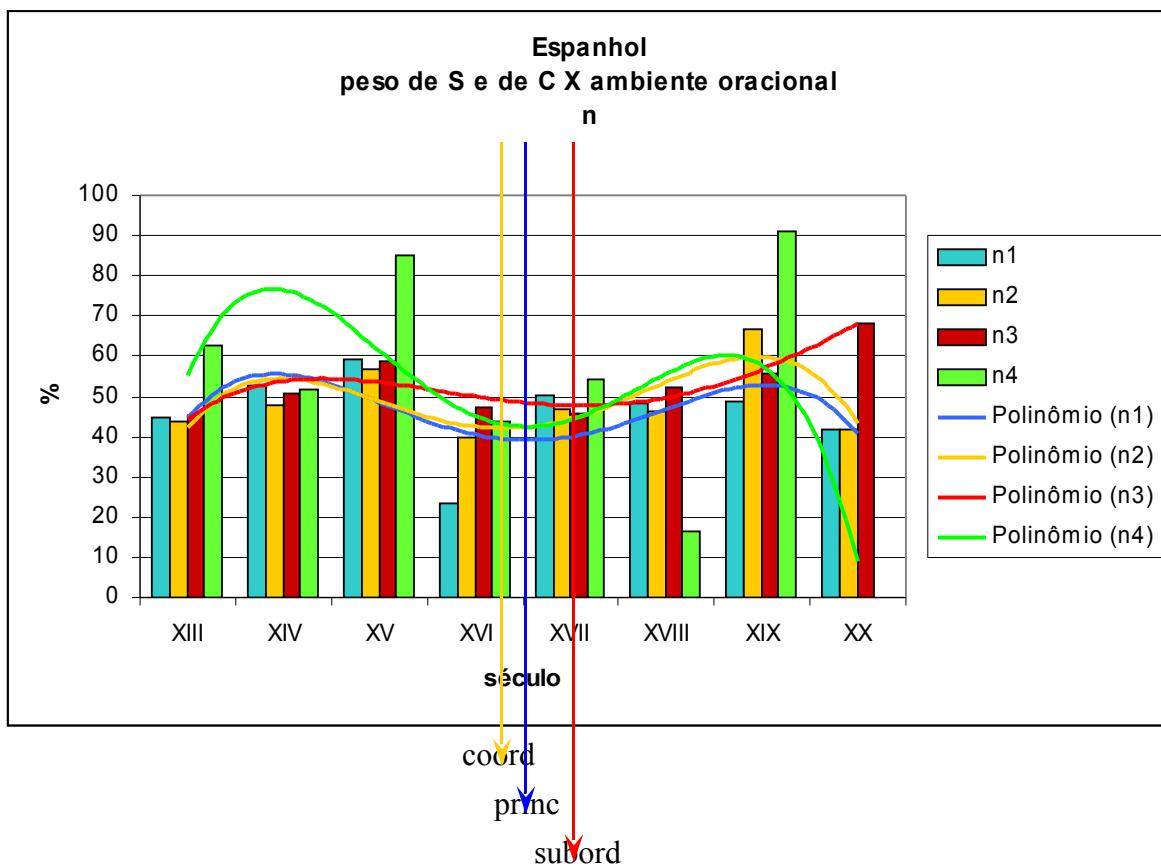


Gráfico E26c(T, normalizado) – comparação entre os perfis de complemento verbal pleno nos quatro ambientes oracionais propostos.

A classe **o** (complemento verbal pronominal) apresenta, em linhas gerais, um padrão idêntico de distribuição das ocorrências em relação ao Gráfico E7 *et similia*: à predominância das frequências associadas às orações coordenadas (**o2**) e subordinadas (**o3**), nos dois terços iniciais do gráfico, segue-se o aumento das proporções de orações principais (**o1**), que passam a constituir a classe mais comum no trecho final do gráfico. Observe-se, ainda, a escala diminuta contra a qual essas ocorrências são computadas – comparável à da classe **I**, também pronominal (sujeito, Gráficos E24).

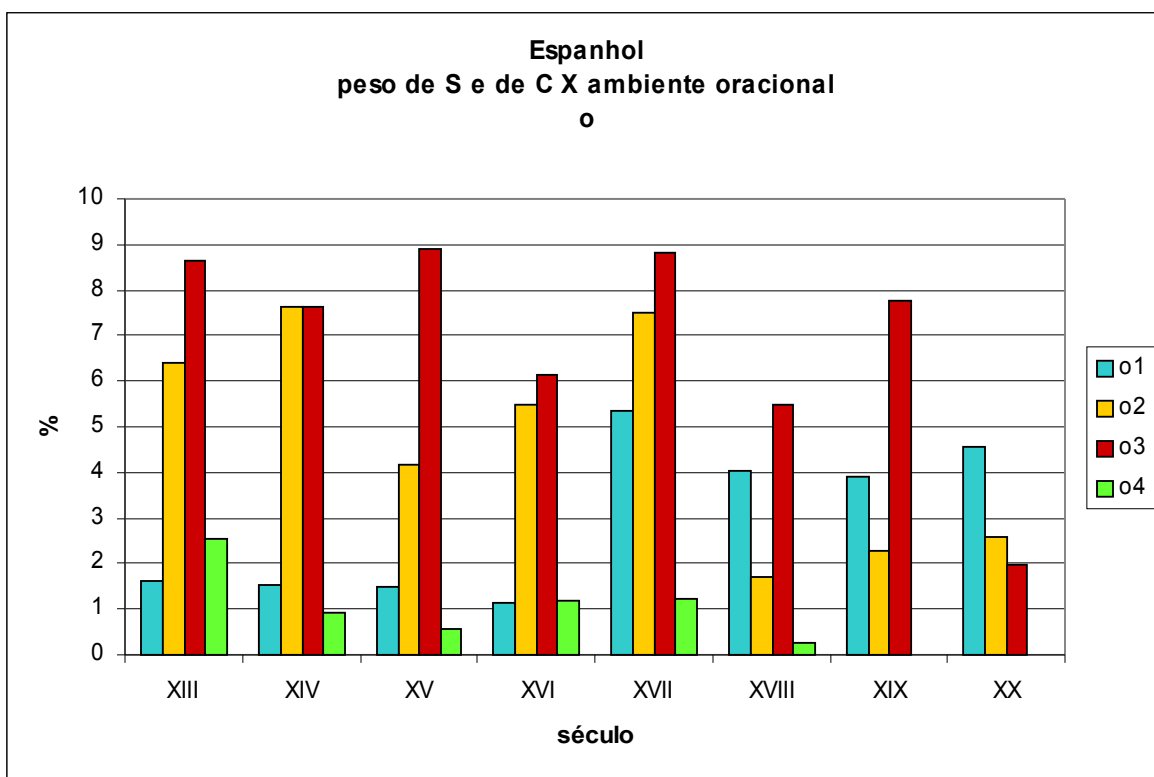


Gráfico E27a(T) – comparação entre os perfis de frequência de complemento verbal pronominal nos quatro ambientes oracionais propostos.

Não há diferenças apreciáveis entre as curvas representativas do comportamento das ocorrências de **o** nos ambientes oracionais em estudo – ou seja, o padrão já delineado pelo fator *ambiente oracional* vai-se repetir mesmo à escala mais modesta de frequência de complementos pronominais. Veja-se também que a classe **o** em orações principais passa a ser mais comum no *corpus* analisado do que coordenadas (**o2**) no século XVIII, mas só no início do século XX suplanta as frequências de subordinadas (**o3**).

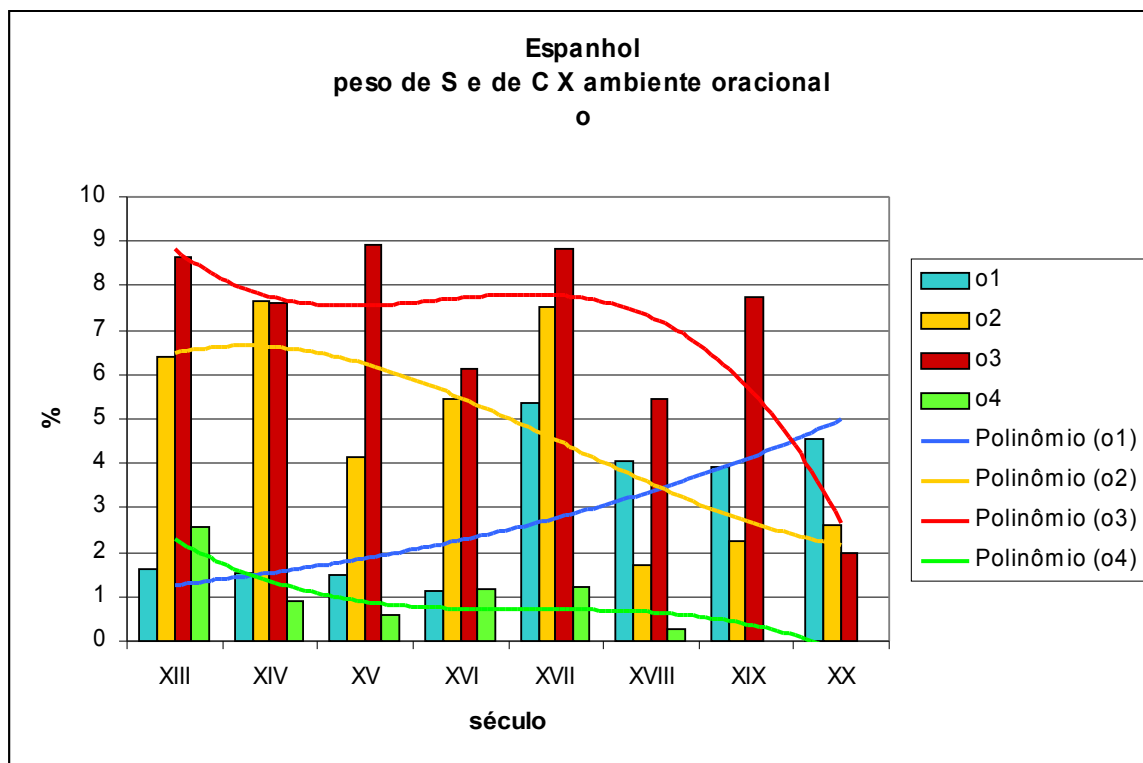


Gráfico E27b(T) – comparação entre os perfis de complemento verbal pronominal nos quatro ambientes oracionais propostos.

Normalizando-se as freqüências, nota-se que as variações relativas são mais relevantes para as ocorrências associadas às orações coordenadas (**o2**) e às coordenadas em subordinação (**o4**); as curvas têm, aos pares, o mesmo padrão gráfico. No intervalo para o qual foram feitas as observações de máximos e de mínimos locais (entre os séculos XV e XVII, principalmente), esses dois pares também se comportam de maneira distinta: [principais+subordinadas] mostram mínimo no século XV e máximo tardio, no século XVIII; o outro par tem máximos no século XVI.

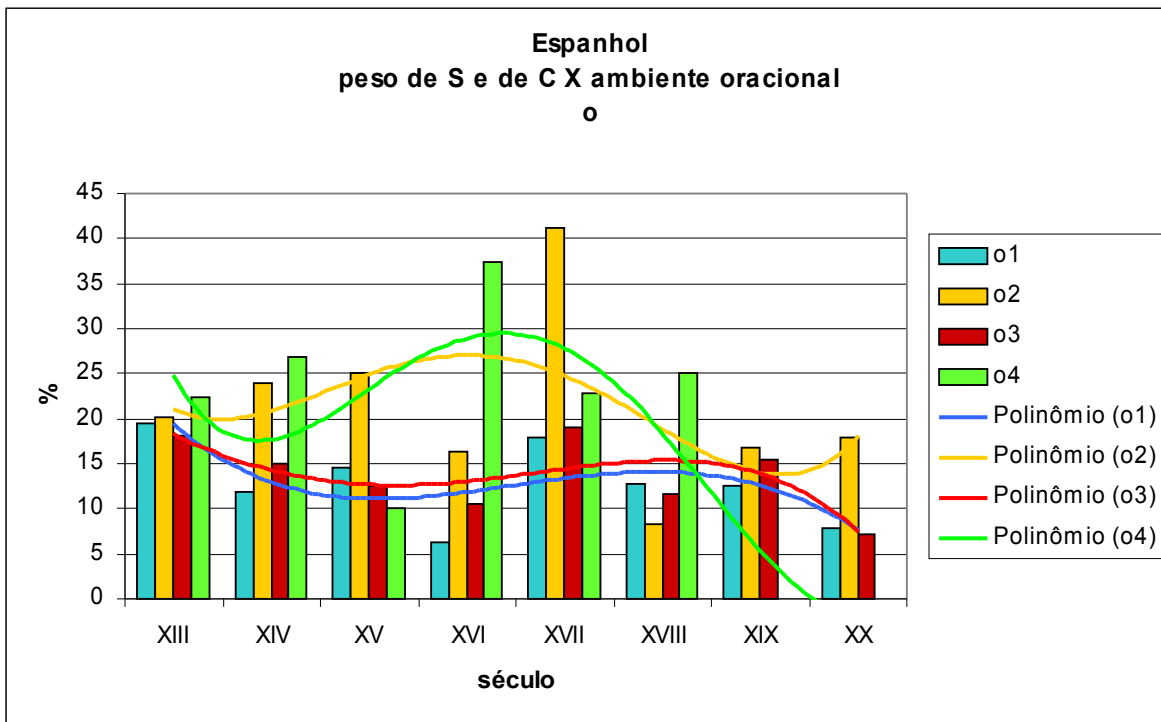


Gráfico E27c(T, normalizado) – comparação entre os perfis de complemento verbal pronominal nos quatro ambientes oracionais propostos.

O fator *regência do verbo* mostra distribuições peculiares do conjunto de dados. O Gráfico E28a mostra a distribuição dos dados com respeito aos desdobramentos desse fator, destacando-se a prevalência do grupo **u** (transitividade direta) em todo o *corpus*.

À exceção do subconjunto **xx** (intransitivos), as classes de regência verbal apresentam freqüências estáveis (**u** (diretos), **v** (indiretos)) ou em leve declínio (**w** (bitransitivos), **x** (ligação)). As retas de regressão linear para essas classes retornam os seguintes valores: $iRRL(u) = -0,27$; $iRRL(v) = -0,10$; $iRRL(w) = -0,65$; $iRRL(x) = -0,64$; e $iRRL(xx) = 1,64$. Mais interessante, entretanto, é plotar as curvas correspondentes à evolução das freqüências de cada classe: vê-se, então, que são praticamente especulares as tendências dos subconjuntos **u** (diretos), o mais freqüente, e **xx** (intransitivos), o único que mostra crescimento efetivo; considerando-se uma excentricidade maior entre os séculos XIV e XV, também são complementares as tendências do mesmo grupo **u** (diretos) e o de **x** (ligação). Em outras palavras, o aumento nas freqüências de **u** (diretos) é acompanhado de uma diminuição nas proporções absolutas de **x** (ligação) e de **xx** (intransitivos).

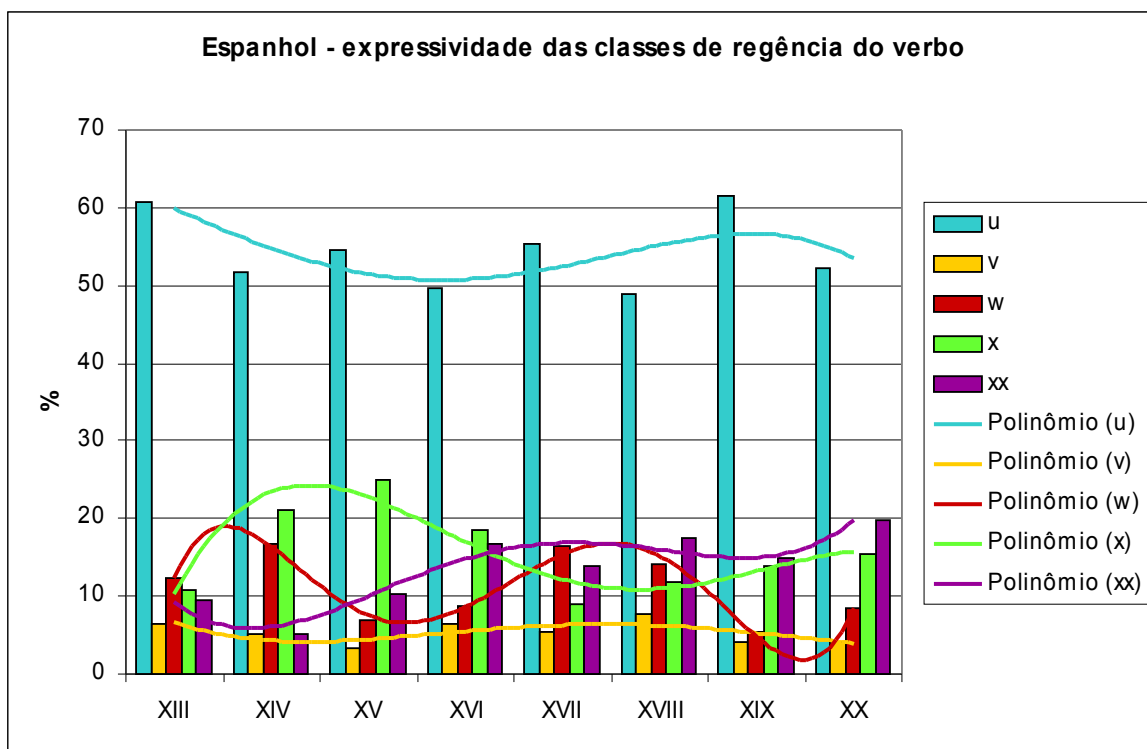


Gráfico E28a – freqüências absolutas das classes de regência verbal no *corpus* espanhol.

Tome-se em primeiro lugar, na análise individual, a classe **u** (diretos), na qual há predominância absoluta da ordem **i** (VC) – que permanece estável no cômputo geral e a despeito da flutuação em alguns pontos. As ordens secundárias, **a** (SVC) e **j** (CV), mostram, por outro lado, tendências opostas: enquanto a primeira dobra em frequência absoluta, a última sofre diminuição em suas proporções. Dentre as demais ordens, apenas **g** (SV, cuja tendência é decrescente) e **h** (VC, de perfil ascendente) mostram comportamentos definidos em todo o intervalo de amostragem. As outras ordens, todas abaixo de 10%, têm expressão demasiado baixa.

São exemplos de ocorrências associadas a essa classe:

E desy venieron ... Eufremes el almirante, e traxo mill caualleros muy ardites [TRO]
 Creer que hay infinitos mundos procedió de querer medir el mundo á piés [IND]
 me dieron sentencia de destierro de la Corte por un año a cinco leguas [CON]

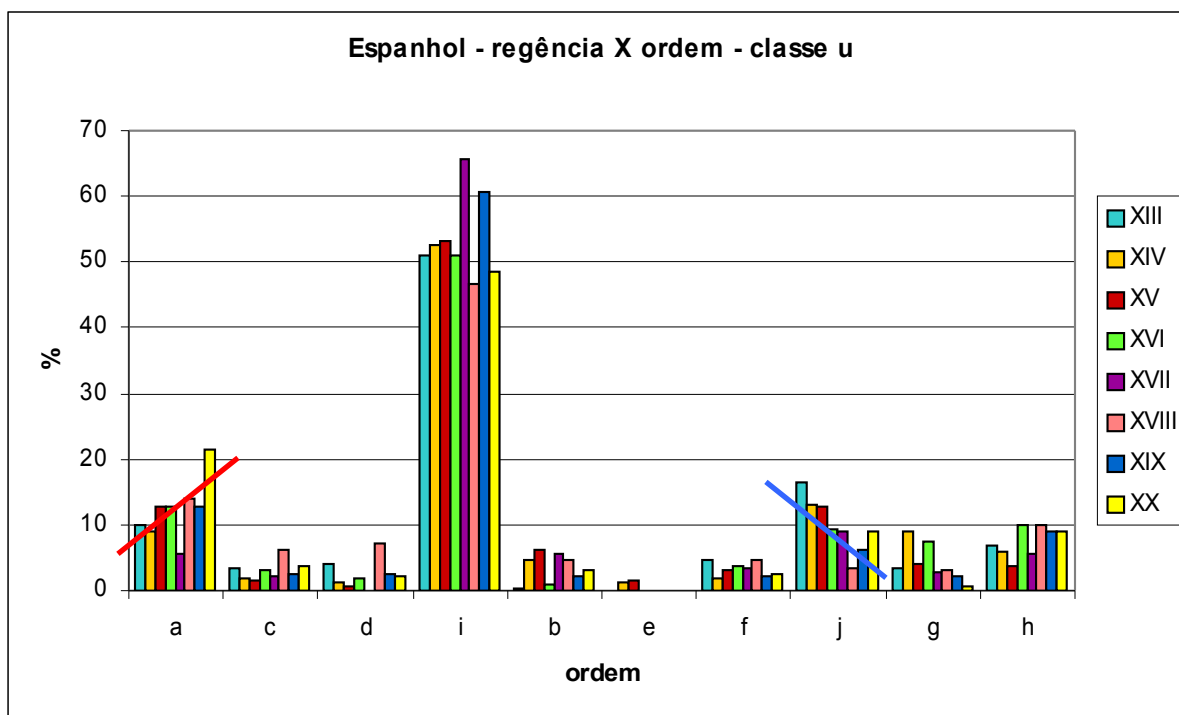


Gráfico E28b – frequências relativas das ocorrências associadas a verbos transitivos diretos no *corpus* espanhol.

Em contraste com a classe **u** (diretos), **v** (indiretos) apresenta decréscimo nas ordens **i** (VC) e **g** (SV), e aumento nas proporções relativas na ordem **j** (CV); contudo, seus índices de ocorrência são marcadamente reduzidos para o intervalo em estudo, e as variações nas proporções relativas dessa classe são, por isso, consideradas residuais ou incipientes; por exemplo, a ordem **a** (SVC), mesmo que ascendente, é por demais descontínua para permitir uma avaliação segura.

São exemplos de ocorrências às quais foi atribuído o traço **v** (indiretos):

e rrogoles que rresçebiesen a Telefo por rrey en su logar [TRO]

et enseñoreó á todos los regnos et tierras de sus comarcas [CAB]

El esplendor de vuestras virtudes claras ... me inclinó e dió afición [FOR]

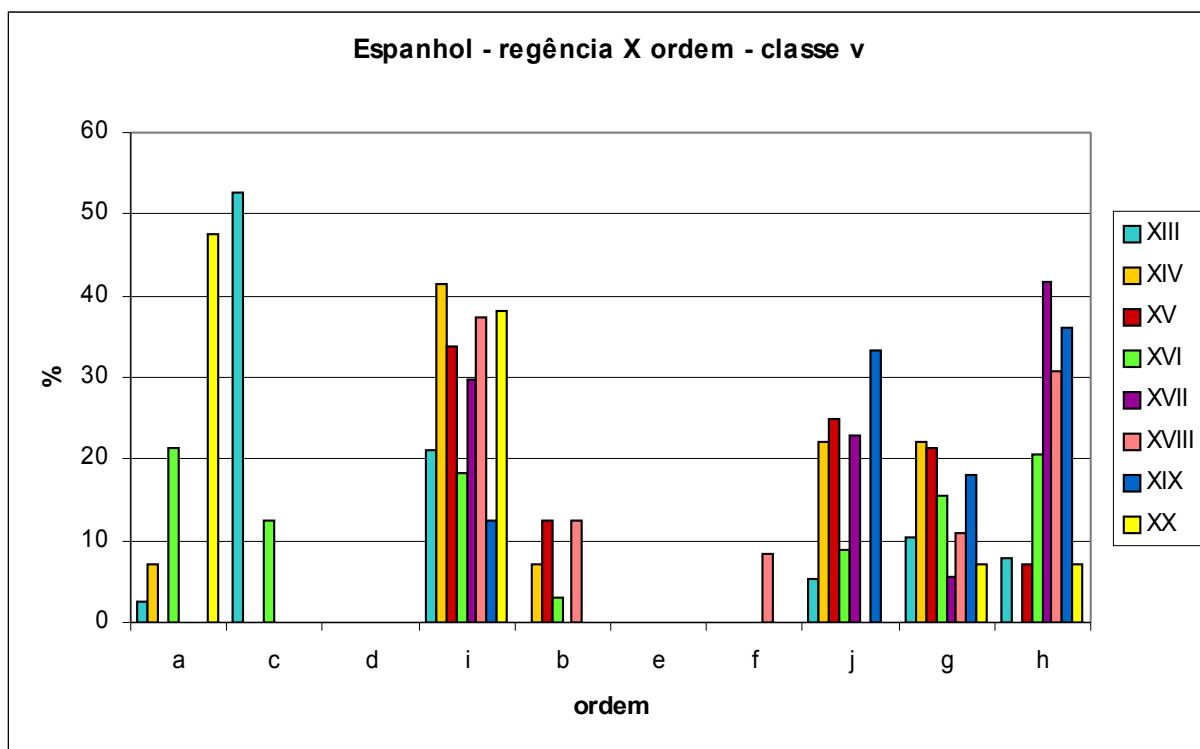


Gráfico E28c – frequências relativas das ocorrências associadas a verbos transitivos indiretos no *corpus* espanhol.

Para a classe dos verbos bitransitivos, **w**, a distribuição das frequências de ocorrência relativa ainda mostra dominância da ordem **i** (VC) – em diminuição, a exemplo

do que também acontece, com menor expressividade, para as ordens **g** (SV) e **h** (VS); a ordem **j** (CV), por sua vez, mostra tendência crescente. Note-se que o sensível aumento dos valores de **c** (VSC) no século XIX pode ser devido à excentricidade de um dos textos: dada a pequena expressividade das proporções (cf. Gráfico E28a) da classe **w** diante do grupo predominante, **u**, um pequeno aumento em uma das ordens produz perfis de acentuada variação.

Compreendem exemplos associados a essa classe:

con sus donas para ofresçer Apolo aquel lugar do estaua Anchiles [TRO]

la mesquindat e malicia nuestra no quita ninguna cosa de su bondat a los buenos [MUG]

No desesperaban los coligados de traer a indecorosos partidos al Rey de Francia [FEL]

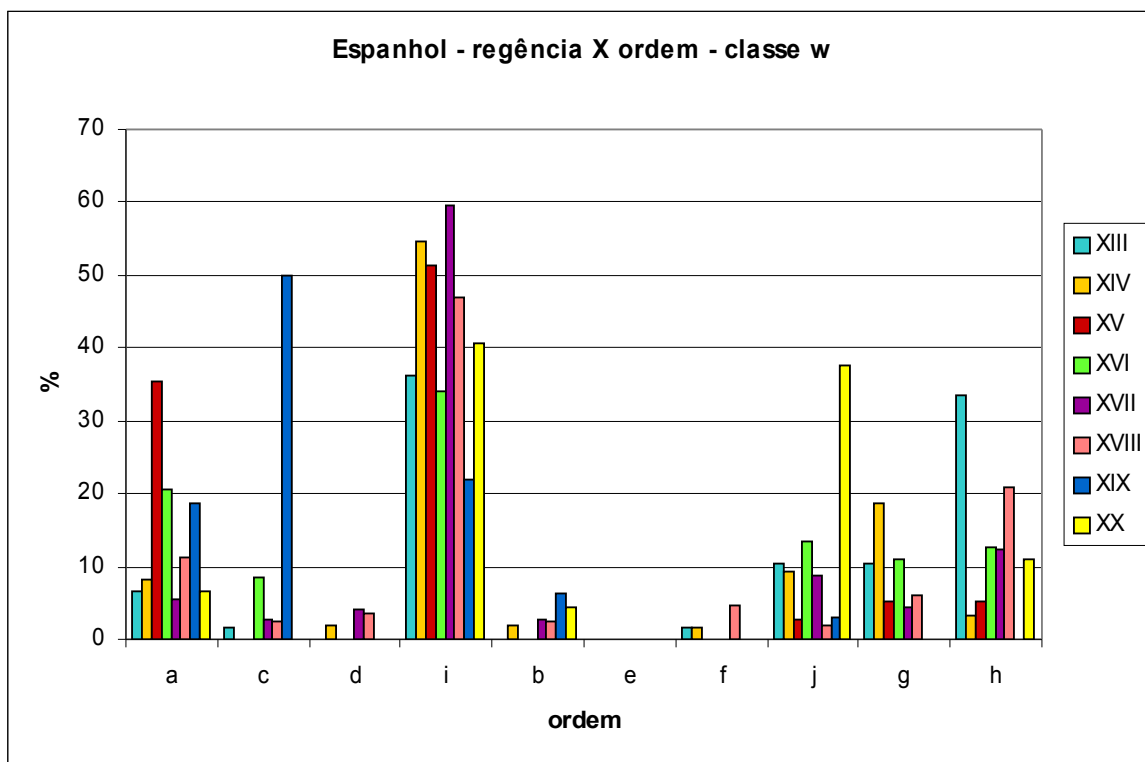


Gráfico E28d – frequências relativas das ocorrências associadas a verbos bitransitivos no *corpus* espanhol.

Os verbos de ligação (**x**) mostram um perfil distribucional diferente dos transitivos: ao contrário destes, a classe **x** predomina na ordem **a** (SVC), além de apresentar uma tendência ao aumento na proporção relativa dessa ordem; também é interessante notar que a ordem **h** (VS) tende a diminuir em termos de frequência relativa de ocorrência.

Exemplos de dados associados a esse traço são oferecidos abaixo:

Con mi padre fuiste terco y atrevido, pero no llegaste hasta [PAZ]

Es la leyenda incompleta, o a mí no me la contaron [SIL]

Fué éste un género de trabajo que dispuso Dios con alta providencia [AGR]

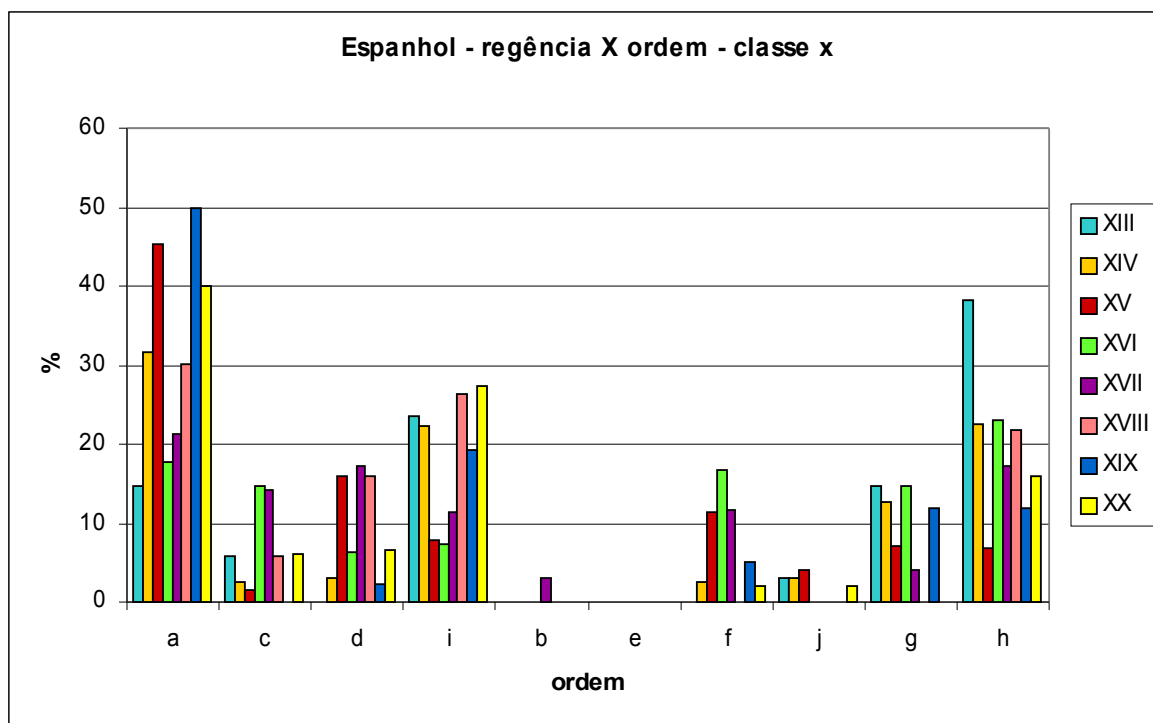


Gráfico E28e – frequências relativas das ocorrências associadas a verbos de ligação no *corpus* espanhol.

Outro perfil bastante distinto dos anteriores é o apresentado pela classe dos intransitivos, **xx**: aqui, a predominância absoluta e indiscutível está relacionada com as ordens **g** (SV) e **h** (VS); contudo, o que mais chama a atenção nesses perfis é sua complementaridade: um aumento em **g** (SV) é acompanhado de uma diminuição em **h** (VS) e vice-versa, como apresentado no gráfico E28g. Resumem-se em seguida, na Tabela E6, as tendências de comportamento da amostra.

Os exemplos apresentados abaixo foram relacionados à classe **xx**:

del entendimiento se siguieron diversos y grandiosos afectos de su voluntad [AGR]
No nacía este rebelión de aquel común alivio que sienten los muchachos con el ocio [TOR]
iba declinando el sol cuando yo llegaba a hacer prueba personal de ello [CAS]

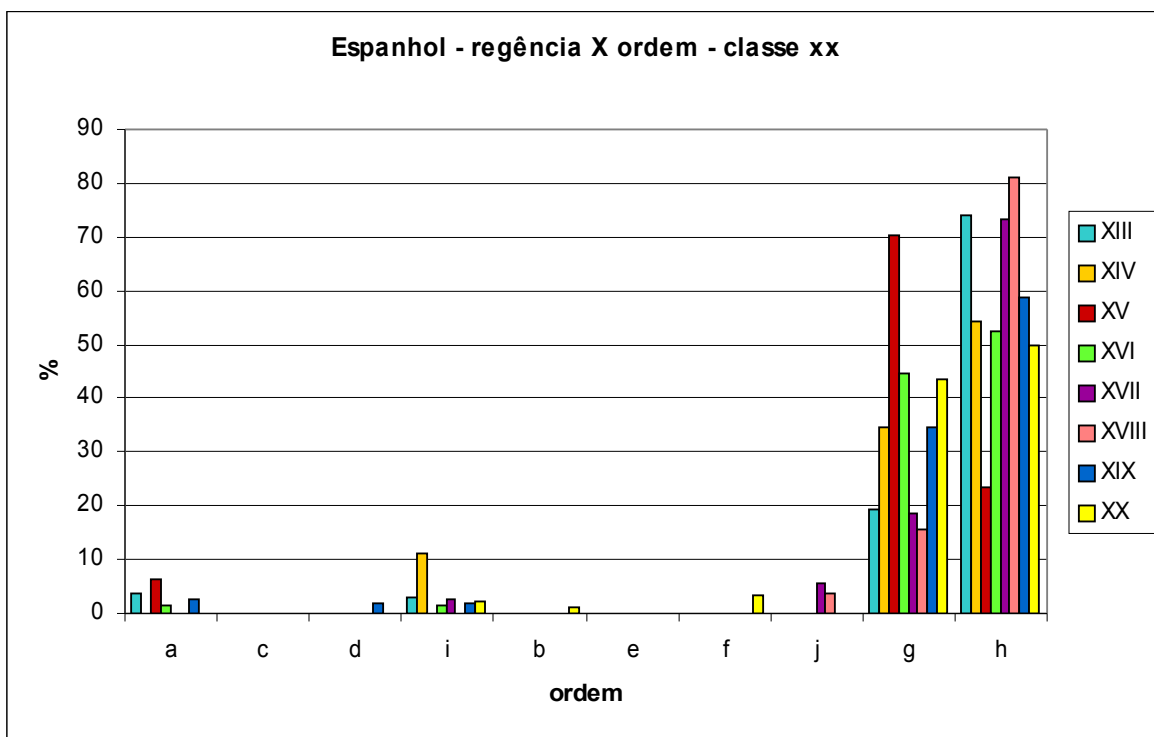


Gráfico E28f – frequências relativas das ocorrências associadas a verbos intransitivos no *corpus* espanhol.

No gráfico abaixo, ilustram-se as vantagens da utilização da regressão linear e das curvas para a extração de informações sobre o comportamento dos dados. As RRLs mostram que, apesar da flutuação dos valores para cada século, os conjuntos de dados tendem à estabilidade – e que, de maneira geral, a classe **g** (SV) varia em torno de um valor médio que é inferior ao da classe **h** (VS). Por outro lado, os perfis das curvas são, de fato, invertidos – como se pode observar a partir do exame das curvas polinomiais.

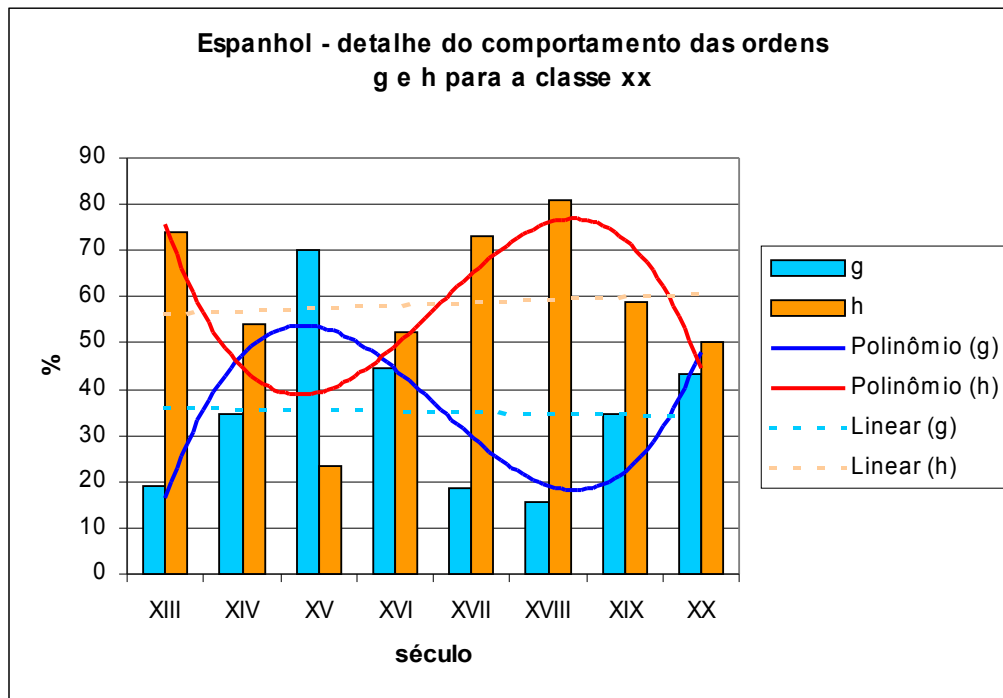


Gráfico E28g – frequências relativas das ocorrências associadas a verbos transitivos diretos no *corpus* espanhol.

regência	a	i	j	g	h
u	↗	→	↘	↘	↗
v	↗	→	↗	↘	→
w	↘	↘	↗	↘	↘
x	↗	↗		↘	↘
xx				→	↗

Tabela E6 – resumo das principais tendências de comportamento do *corpus* espanhol segundo o critério *regência do verbo*. Destacam-se as categorias dirigidas pela ordem **g** (SV) e pela regência **x** (ligação).

Uma primeira aproximação ao fator *tempo do verbo* também acusa uma distribuição desigual das ocorrências. A classe α (tempo verbal finito) predomina largamente sobre as demais, mas mostra leve tendência ao decréscimo (iRRL = -0,35) no intervalo considerado; β (formas reduzidas do verbo), cujos máximos beiram os 25% da expressividade de α , apresenta aumento um pouco mais firme (iRRL = 0,83), e γ e δ ocorrem segundo frequências absolutas baixas – geralmente abaixo dos 10% absolutos.

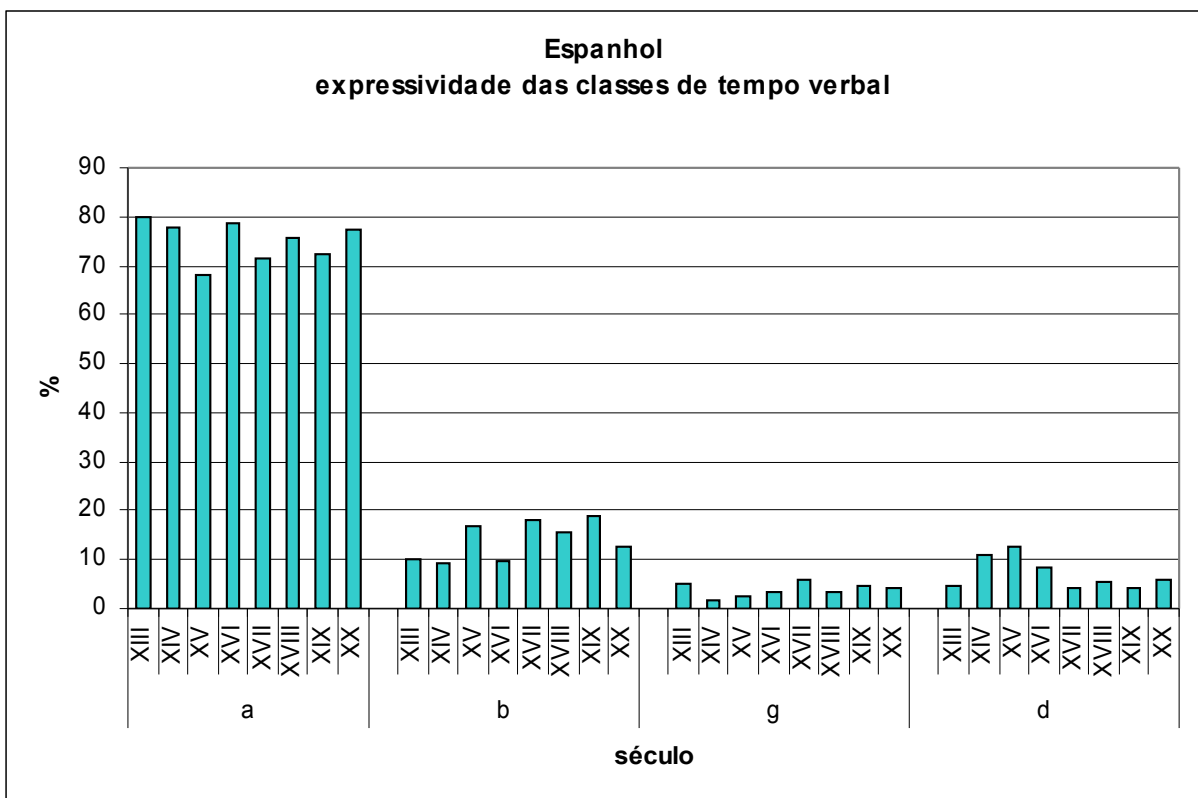


Gráfico E29 – perspectiva geral da distribuição das ocorrências do *corpus* espanhol segundo o critério *tempo do verbo*.

Uma observação mais detalhada dos comportamentos das classes de *tempo verbal* leva a resultados peculiares – nem tanto pelos perfis em si, mas devido à associação desse traço morfológico com tendências à mudança em algumas ordens.

Para as formas finitas do verbo, α (a categoria mais freqüente), identificam-se ordens preferenciais segundo as quais as ocorrências compõem o *corpus* espanhol. Assim é que as ordens **a** (SVC), **i** (VC) e **h** (VS) se sobressaem, apresentando tendências diferentes: dentre essas, as freqüências relativas de formas finitas do verbo tendem a aumentar nas ordens que têm o sujeito expresso (iRRL(SVC) = 0,86; iRRL(VS) = 1,62), e diminuem na ordem VC (iRRL = -1,70).

São ocorrências associadas a esse traço:

sabedes vos todos que se guarnesçen agora ellos quanto pueden contra nos [TRO]
que seis mill vírgines filavan e texían lo nescessario en los dichos ornamentos [MUG]
é desde sopo el infante don Juan é los otros que estavan en Sevilla quel [SAN]

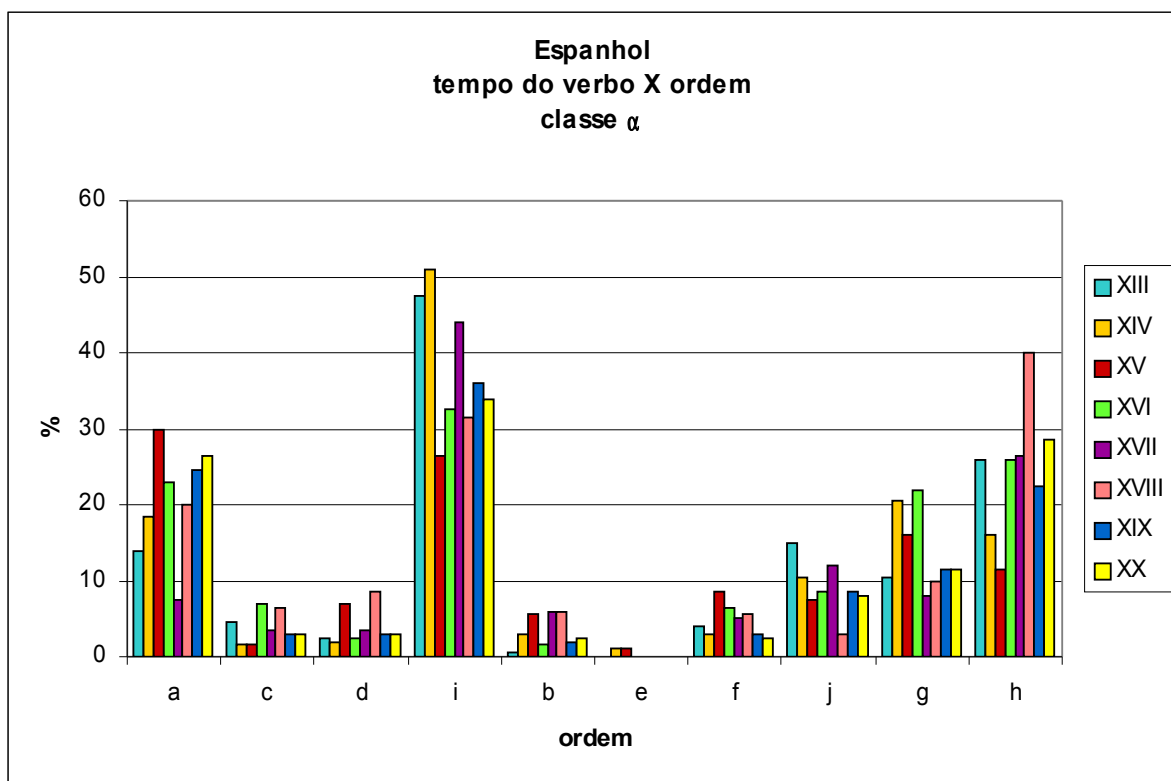


Gráfico E30a – distribuição das ocorrências de formas verbais finitas segundo a ordem dos constituintes.

As formas reduzidas do verbo, β , por sua vez, concentram-se preferencialmente na ordem **i** (VC), como se pode observar no gráfico abaixo. Esse subconjunto de ocorrências mostra uma inclinação francamente ascendente ($iRRL(VC) = 1,35$), e, dadas as proporções desprezíveis das outras ordens, o comportamento da classe β vai-se confundir com a distribuição dessas ocorrências na ordem **i** (VC).

Compreendem exemplos dessa categoria os listados abaixo:

e mal trecha, recibiendo muertes por muy crueles lides [CGE]

Siendo pues redondo todo el cuerpo del mundo [IND]

y sin saber lo que me sucedía me hallé en el gremio de los escolares [TOR]

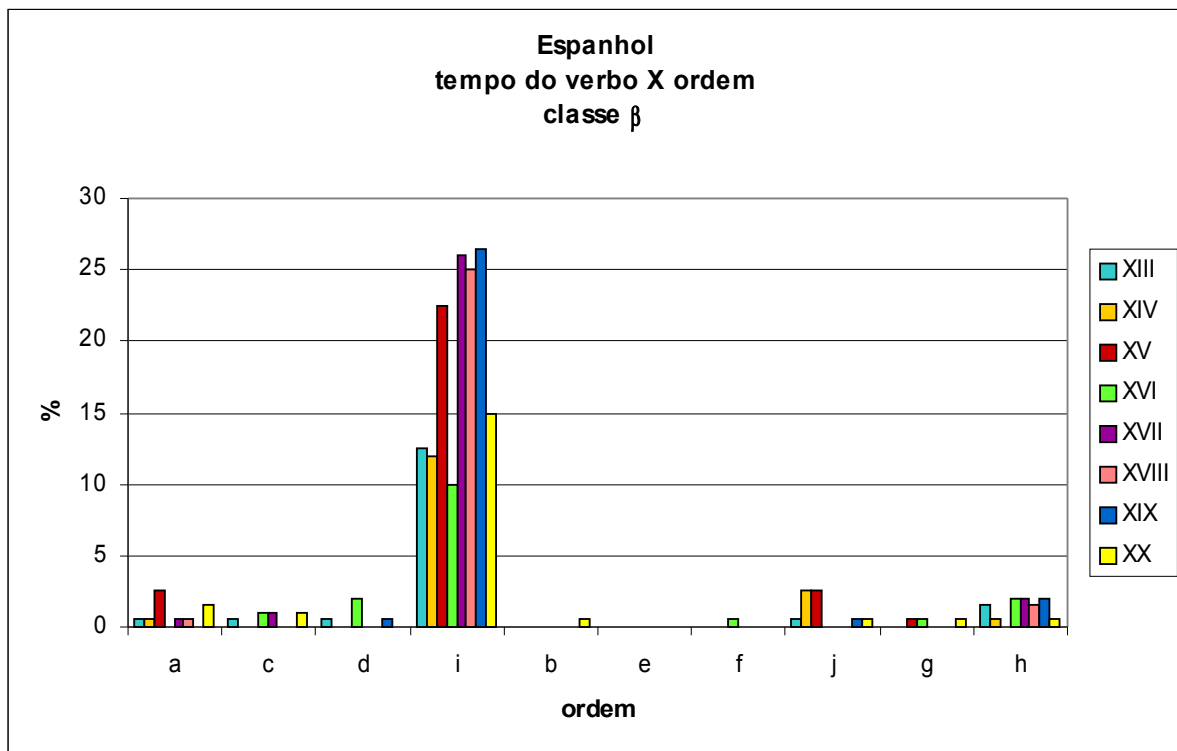


Gráfico E30b – distribuição das ocorrências de formas verbais reduzidas segundo a ordem dos constituintes.

As ocorrências associadas às locuções verbais, γ , são pouquíssimo freqüentes (a ordem **i** (VC) não chega aos 5%), razão pela qual é impróprio abordá-las com o mesmo instrumental metodológico utilizado para as demais categorias e apresentaremos apenas o gráfico que sumariza sua distribuição no *corpus*.

São exemplos relacionados a esse traço:

que por sí sola no podía alcanzar la gracia : perderla y cometer pecado sí [AGR]

Mis lectores, entendida la contienda, deberán juzgar tres cosas [PAZ]

A veces quiere la llanura ser la expresión del paisaje [SIL]

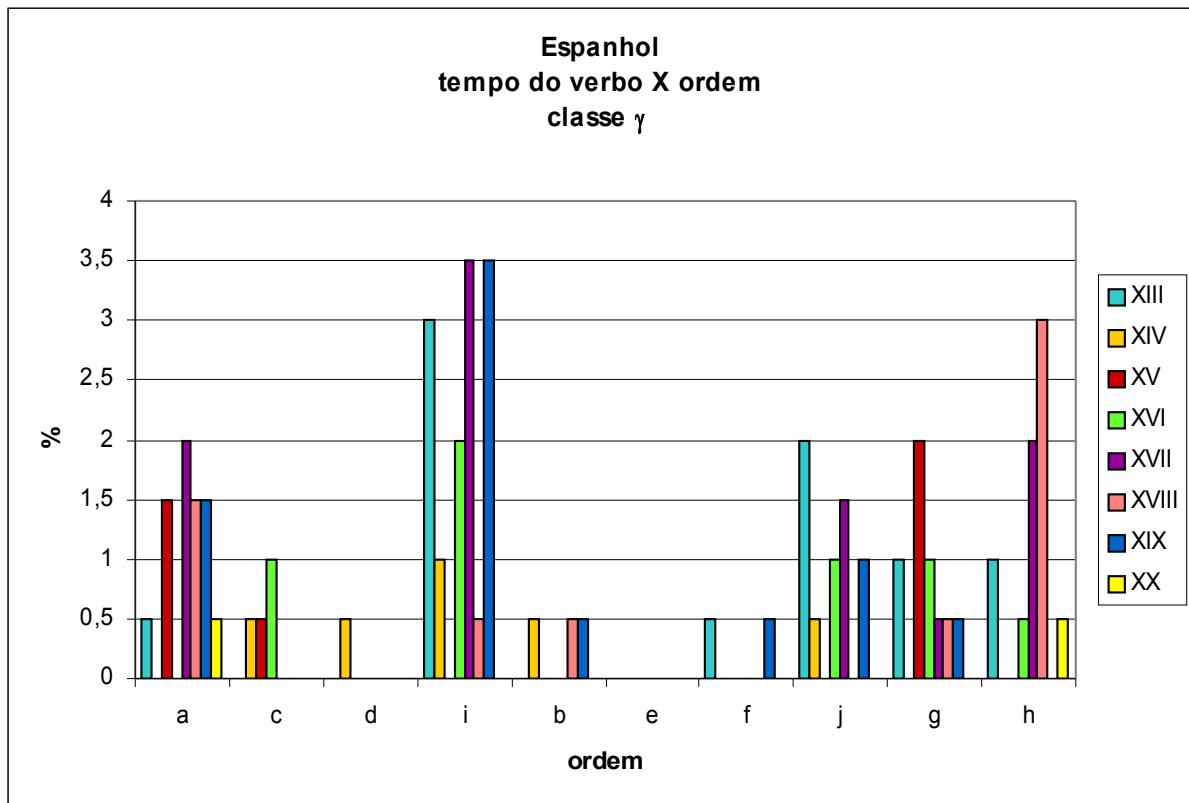


Gráfico E30c – distribuição das ocorrências de locuções verbais segundo a ordem dos constituintes.

A classe δ (tempos verbais compostos), que compreende muitas estruturas baseadas em participios passados, também tende à ordem **i** (VC). Ocorre em porporções baixas, mas não tanto quanto a classe γ (locuções verbais), o que permite inferir, ao menos com alguma segurança, um perfil descendente (iRRL (VC)= -0,61) para essa ordem mais comum.

Exemplos associados a esse traço são encontrados abaixo:

E este ... falló un hombre que le habia derribado una bestia en tierra [ENX]

Como había expirado con la misma destemplanza el pasado año [FEL]

Tú ... has salvado a muchos criminales de un encierro merecido [PROD]

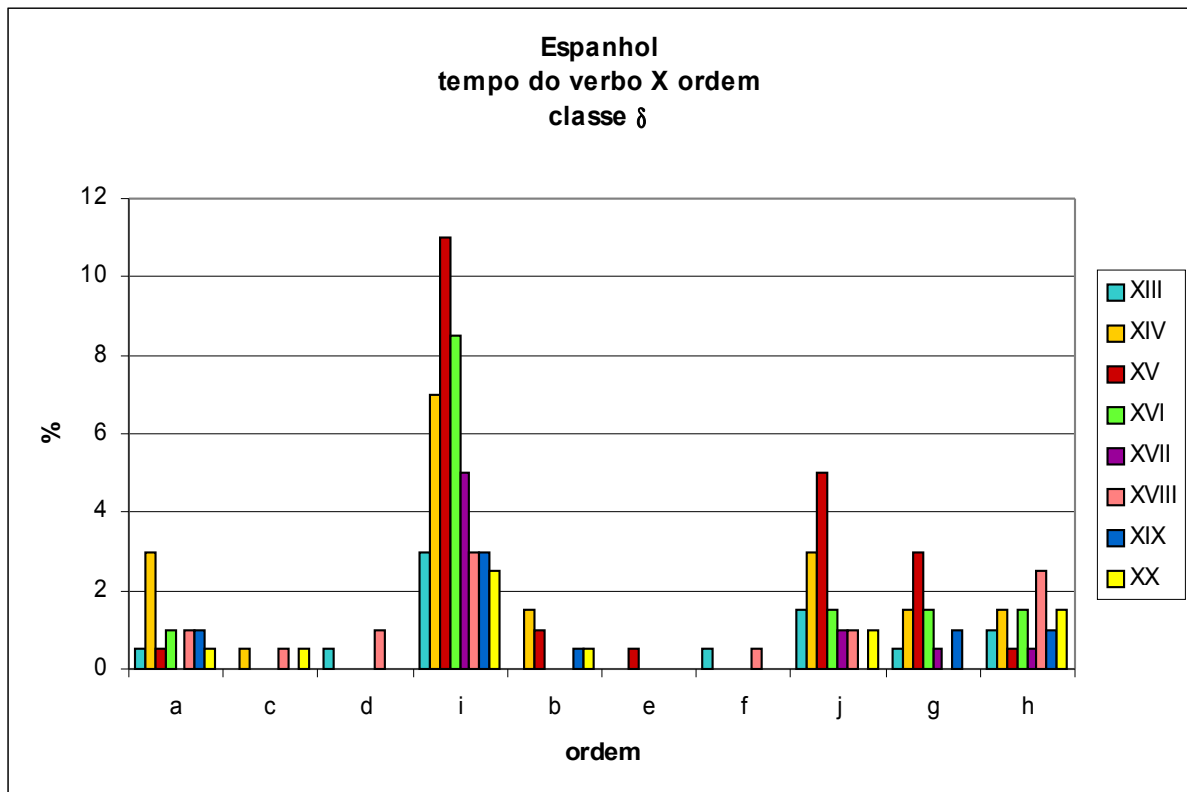


Gráfico E30d – distribuição das ocorrências de tempos verbais compostos segundo a ordem dos constituintes.

A fixação de um dos parâmetros – no caso, o *ambiente oracional* – mostra como as ocorrências se distribuem no *corpus* segundo as classes de *tempo do verbo*. Repetem-se,

aqui, os perfis encontrados no Gráfico E7 para todos os ambientes oracionais, o que confirma a proposta de que esse seja o fator que controla todos os demais critérios de análise da mudança arrolados neste estudo.

Para as orações principais, como se observa no gráfico abaixo, as proporções são praticamente constantes em torno dos 10% até o século XVI, a partir de quando ascendem ao patamar dos 25-28% até o século XIX e, por fim, aumentam outros 20% em frequência no século XX. Isso leva ao traçado da reta de regressão linear de inclinação positiva (iRRL (α) = 5,48) já encontrada em gráficos de mesma natureza para outros critérios. Tomem-se os exemplos:

- α – Dice sant Agustín en el libro de La Cibdad de Dios - ENX XIV
- β – Creer que hay infinitos mundos procedió de querer medir el mundo - IND XVI
- γ – quise tomar el apellido de mi padre andando sirviendo al Rey - CON XVII
- δ – No nos han llegado los franceses - PAZ XIX

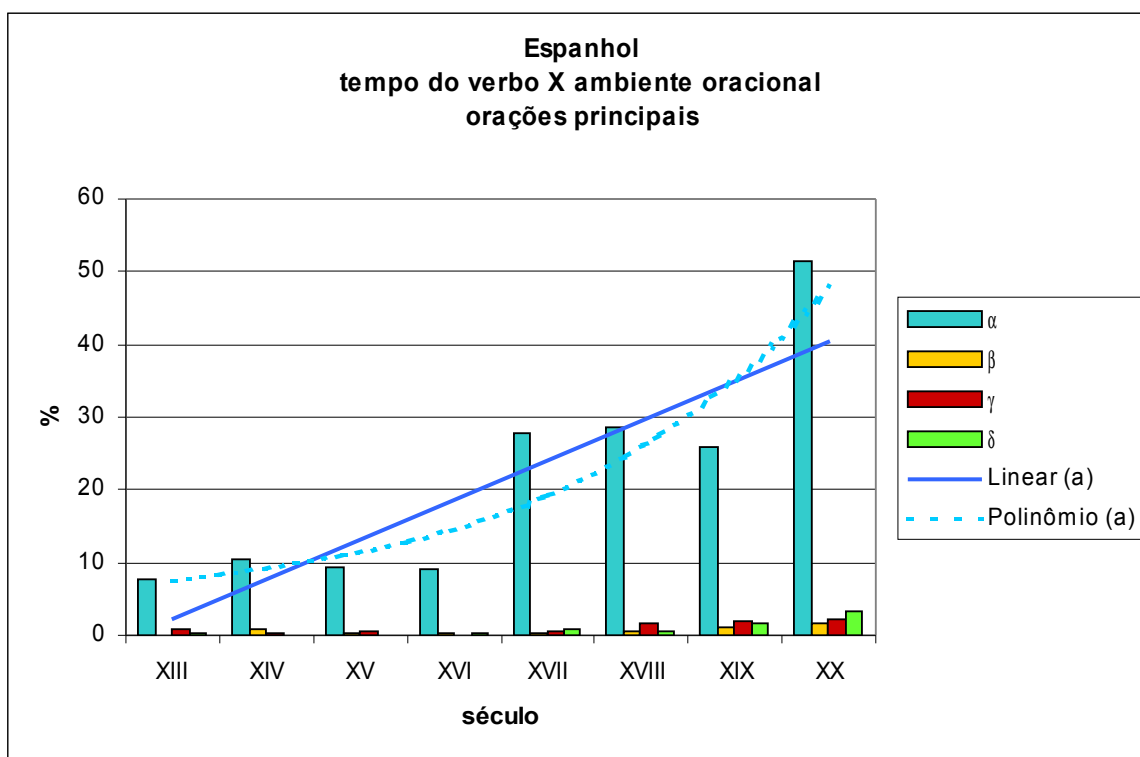


Gráfico E31a – distribuição das ocorrências de orações principais segundo o fator *tempo do verbo*.

No caso das orações coordenadas, o perfil da classe α (formas verbais finitas) é descendente, mas tem inclinação invertida ($iRRL = -2,52$) e de expressividade menor que a da RRL das orações principais apresentada no gráfico anterior. Isso significa que a taxa segundo a qual as formas verbais finitas aumentam em orações principais é mais alta do que a da perda da frequência dessas mesmas formas em orações coordenadas. Observe-se a predominância absoluta das ocorrências α em relação às demais, remetendo também ao Gráfico E29. Exemplos desse traço são reproduzidos abaixo:

α – et connoscieron las cosas que eran estonces - CGE XIII

β – et do non las fallare fechas, facerlas él buenas et derechas - CAB XIV

γ – y la repetición y el vencimiento les va trocando las inclinaciones - TOR XVIII

δ – le habían enseñado la arte de la persuasión - Prod XX

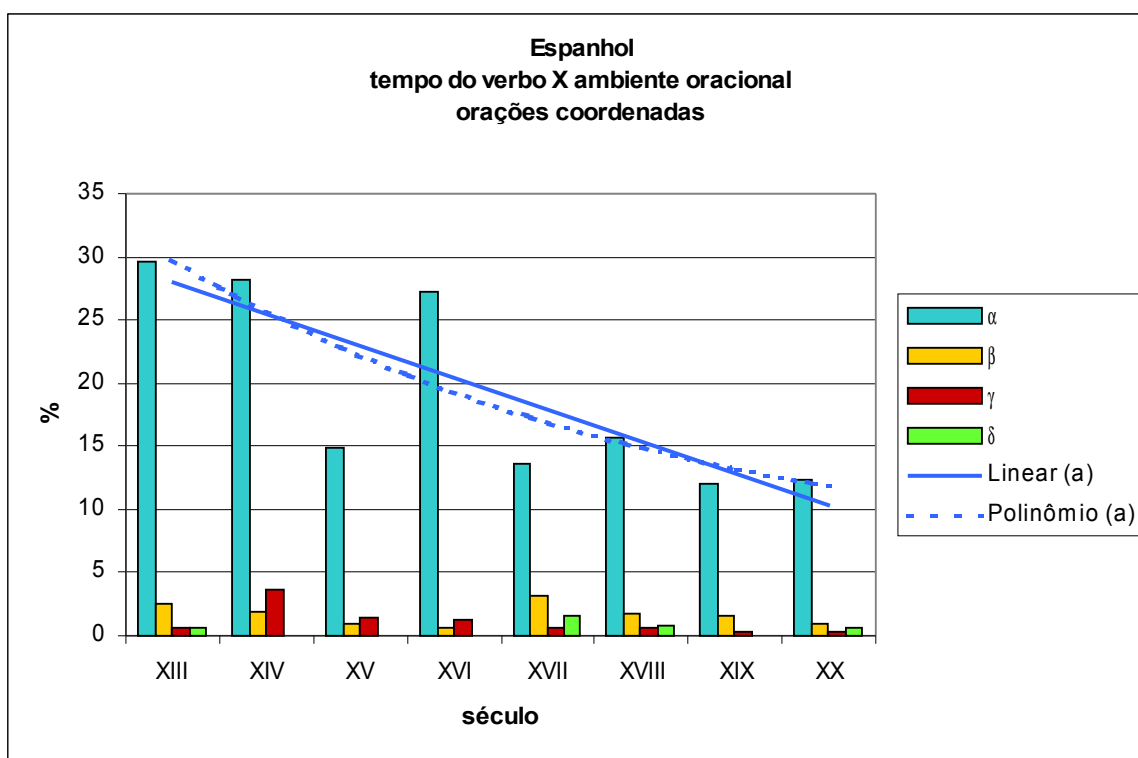


Gráfico E31b – distribuição das ocorrências de orações coordenadas segundo o fator *tempo do verbo*.

Interessante é o perfil de distribuição das classes de tempo verbal em orações subordinadas. Nota-se, outra vez, a influência do ambiente oracional no padrão delineado para a classe α , que têm $iRRL = -2,56$; porém, veja-se que aparecem, aqui, as formas verbais reduzidas do Gráfico E30b, relacionadas a β - ou seja, essas formas estão, em sua quase totalidade, relacionadas às orações subordinadas. Mais ainda: pelo mesmo Gráfico E30b, trata-se de ocorrências relacionadas à ordem i (VC). Definem-se, então, pelo menos três características de naturezas diferentes para esse grupo, cujos exemplos abaixo precedem o gráfico de distribuição dos dados no *corpus* espanhol.

α – por auer rrespuesta que fim aueria esta guerra - TRO XIII

β – costriniéndola por entrañable sentimiento ronper el silencio a la péñola - MUG XV

γ – commo el infante don Juan ... se queria alzar con Sevilla - SAN XVI

δ – la noche antes me había sabido bien la lumbre de la cocina - CON XVII

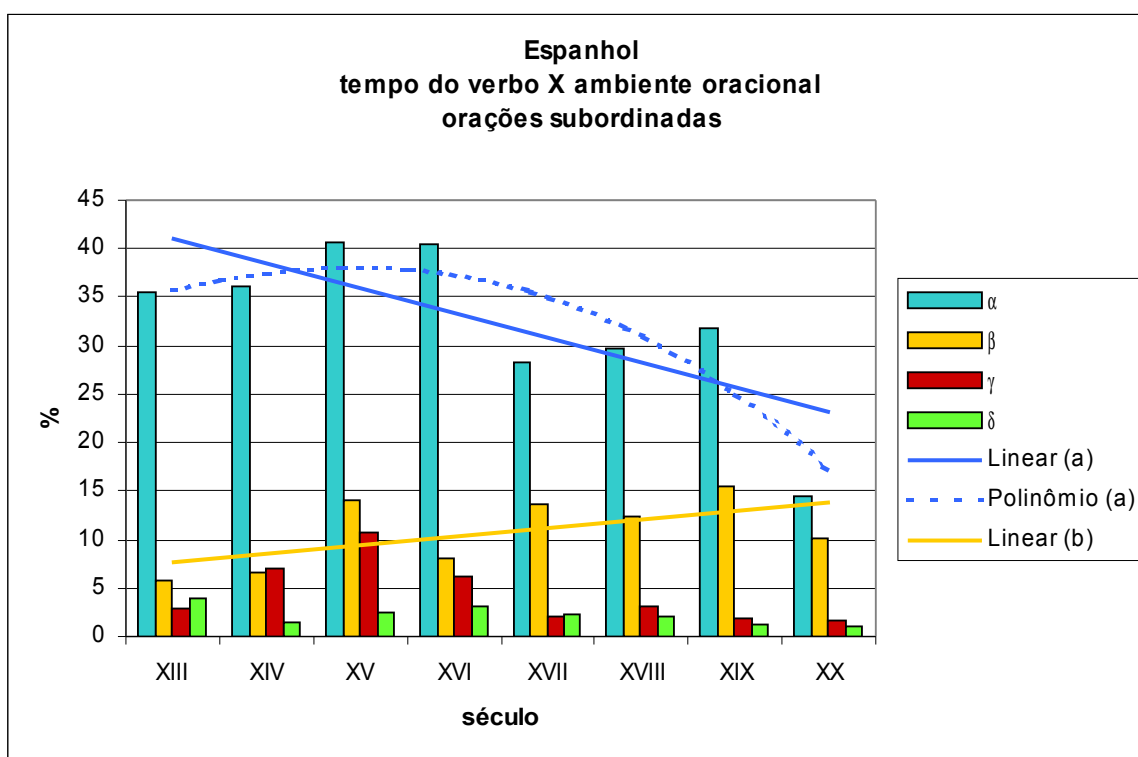


Gráfico E31c – distribuição das ocorrências de orações subordinadas segundo o fator *tempo do verbo*.

A transposição das matrizes utilizadas na plotagem dos dados dos gráficos anteriores leva à fixação do grau de liberdade *tempo do verbo*, e permite a observação do comportamento da amostra segundo o ambiente oracional para as classes de formas verbais.

Nota-se, inicialmente, a mesma espécie de complementaridade entre as freqüências de orações principais e subordinadas já observada em gráficos anteriores. Os polinômios referentes a esses subconjuntos de dados são marcadamente semelhantes, ainda que suas concavidades sejam opostas. Por sua vez, as ocorrências de formas finitas do verbo em orações coordenadas tendem a diminuir, caindo, no intervalo de amostragem, à metade dos níveis iniciais. A normalização não ofereceu perfis cujos resultados fossem comparáveis, devido ao fato de as curvas apresentarem padrões díspares e não-correlacionáveis.

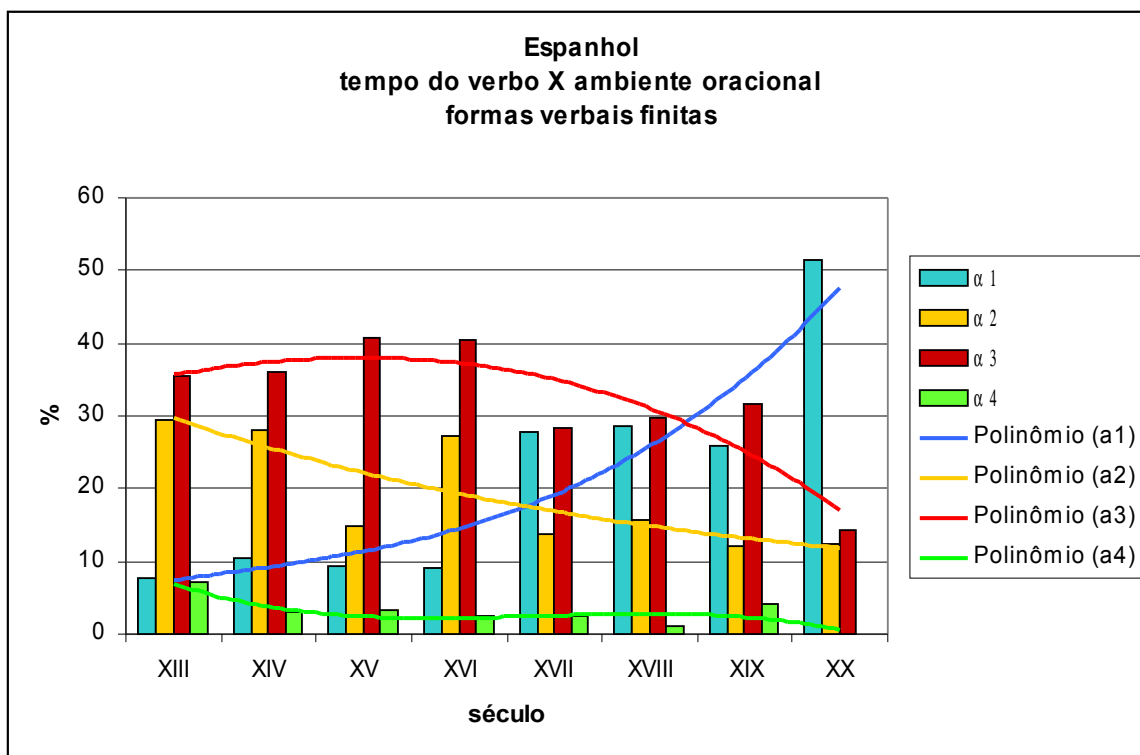


Gráfico E32a(T) – distribuição das ocorrências de formas verbais finitas segundo o fator *ambiente oracional*.

Para as formas verbais reduzidas, o perfil predominante é, como já salientado, o de orações subordinadas ($\beta 3$). A reta de regressão linear dessa subclasse esconde, entretanto, uma característica importante desses dados. Ela tem inclinação positiva ($iRRL = 0,90$), indicando o aumento das freqüências relativas das ocorrências de formas reduzidas em orações subordinadas; contudo, a curva polinomial de $\beta 3$ (tracejada) mostra que, na verdade, o perfil dos dados teria atingido um patamar de variação estável ou ultrapassado, de fato, um ponto máximo no século XVIII, após o qual estaria em declínio.

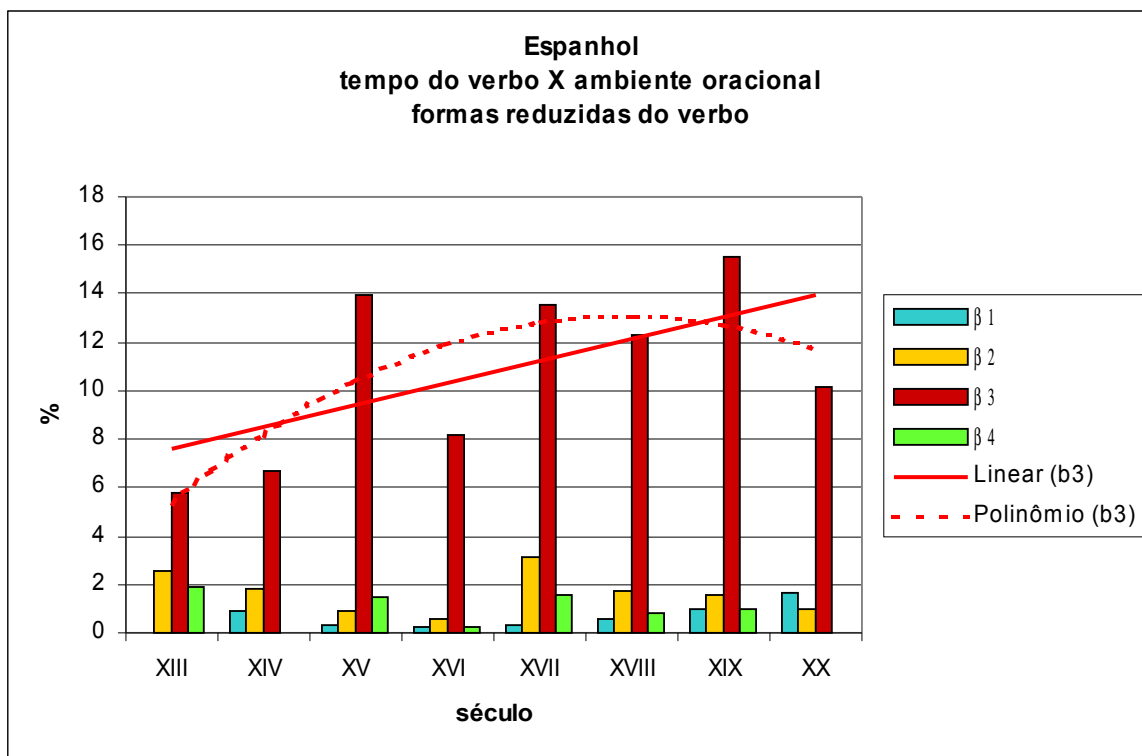


Gráfico E32b(T) – distribuição das ocorrências de formas verbais reduzidas segundo o fator *ambiente oracional*.

A normalização desses resultados mostra que há duas curvas de tendências praticamente paralelas – a de coordenadas, β_2 , e a de coordenadas em subordinação, β_4 –, cujos máximos estão compreendidos no século XVIII. A curva de dados referentes às orações subordinadas, β_3 , é decididamente ascendente, enquanto a de principais, β_1 , é muito modesta diante das demais.

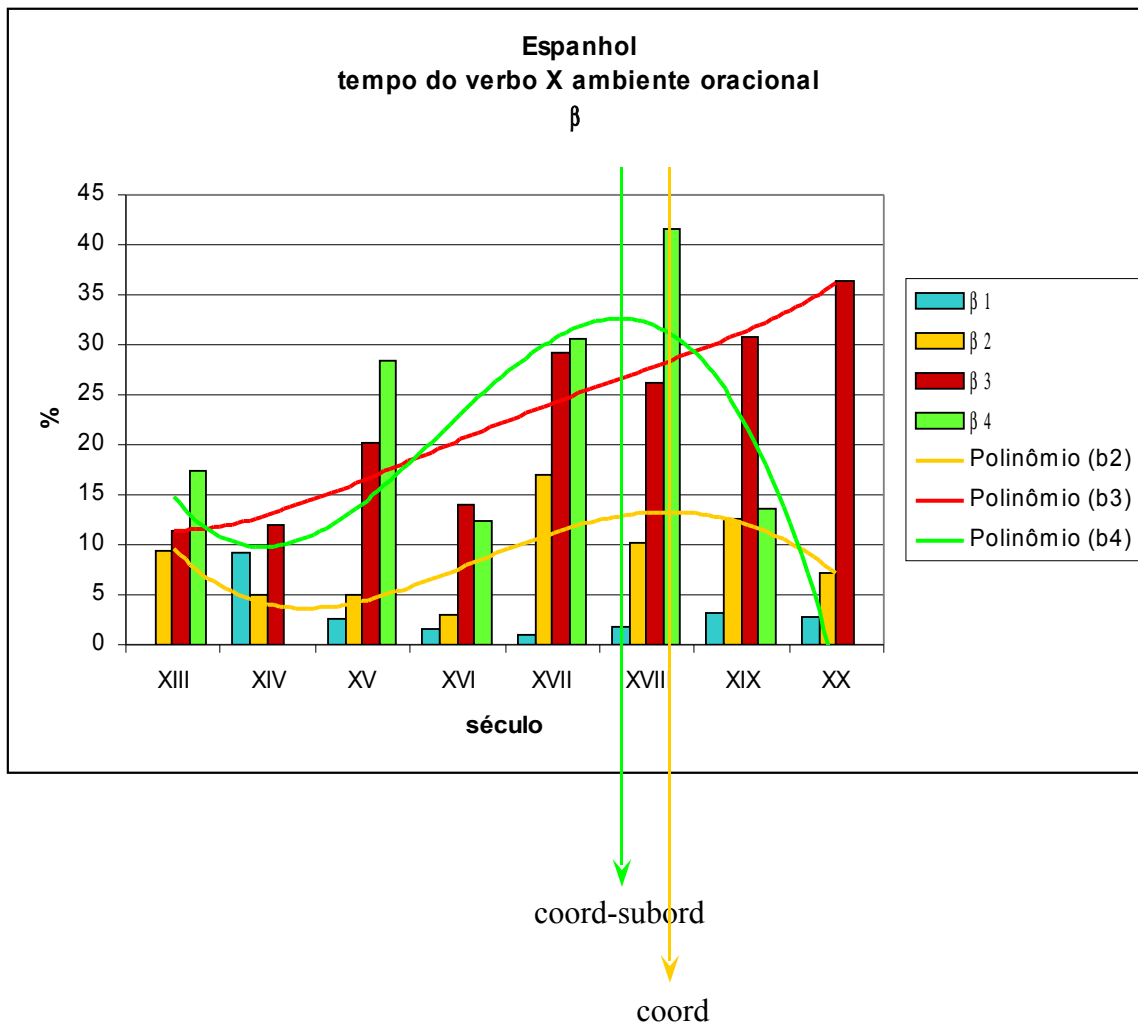


Gráfico E32b(T, normalizado) – distribuição das ocorrências de formas verbais reduzidas segundo o fator *ambiente oracional*.

As ocorrências de locuções verbais se associam, preferencialmente e em quase todo o intervalo de amostragem, às orações subordinadas, γ_3 ; apenas nos dois séculos finais a diminuição das freqüências dessa classe caem a ponto de serem interceptadas pelas proporções de γ_1 , que mostram um aumento muito sutil. Esboça-se também timidamente a curva associada às locuções verbais em orações coordenadas, γ_2 .

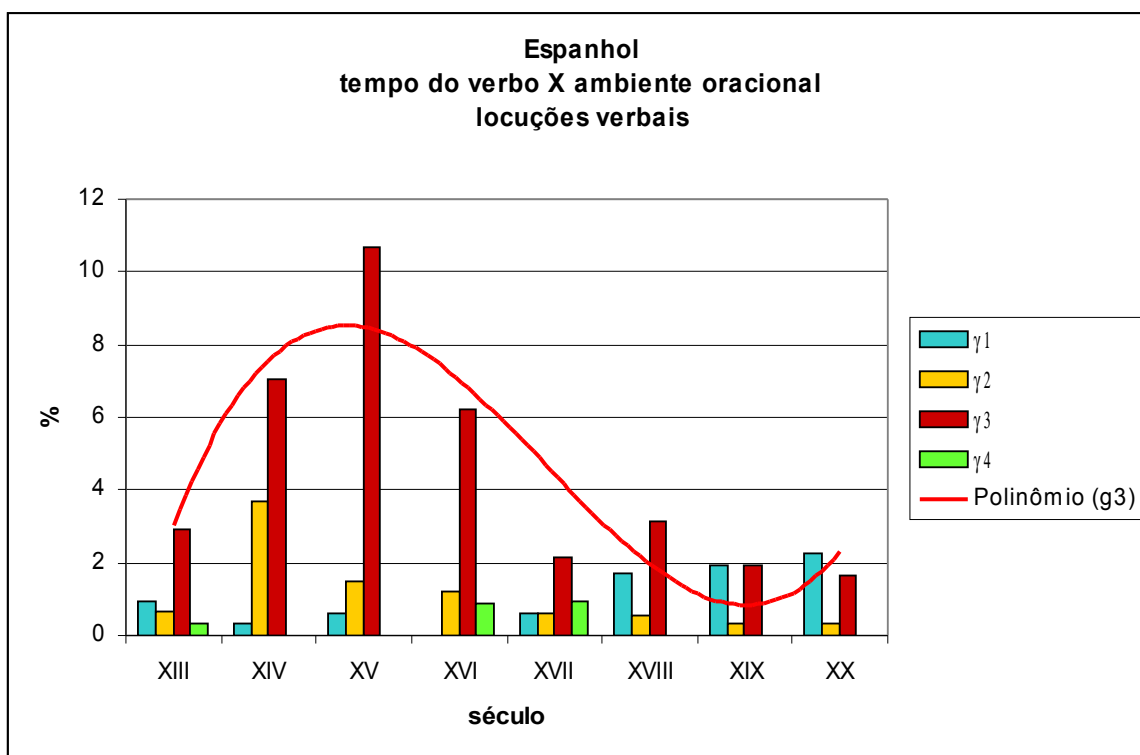


Gráfico E32c(T) – distribuição das ocorrências de locuções verbais segundo o fator *ambiente oracional*.

A partir da normalização dos resultados numéricos, observa-se que, de fato, as tendências para as curvas γ_2 e γ_3 são paralelas, a despeito das diferentes escalas em que essas mudanças de frequência se processam. Além disso, vê-se que a curva de γ_1 tem um comportamento contrário ao das outras duas, invertendo-se em relação a elas com pequena defasagem.

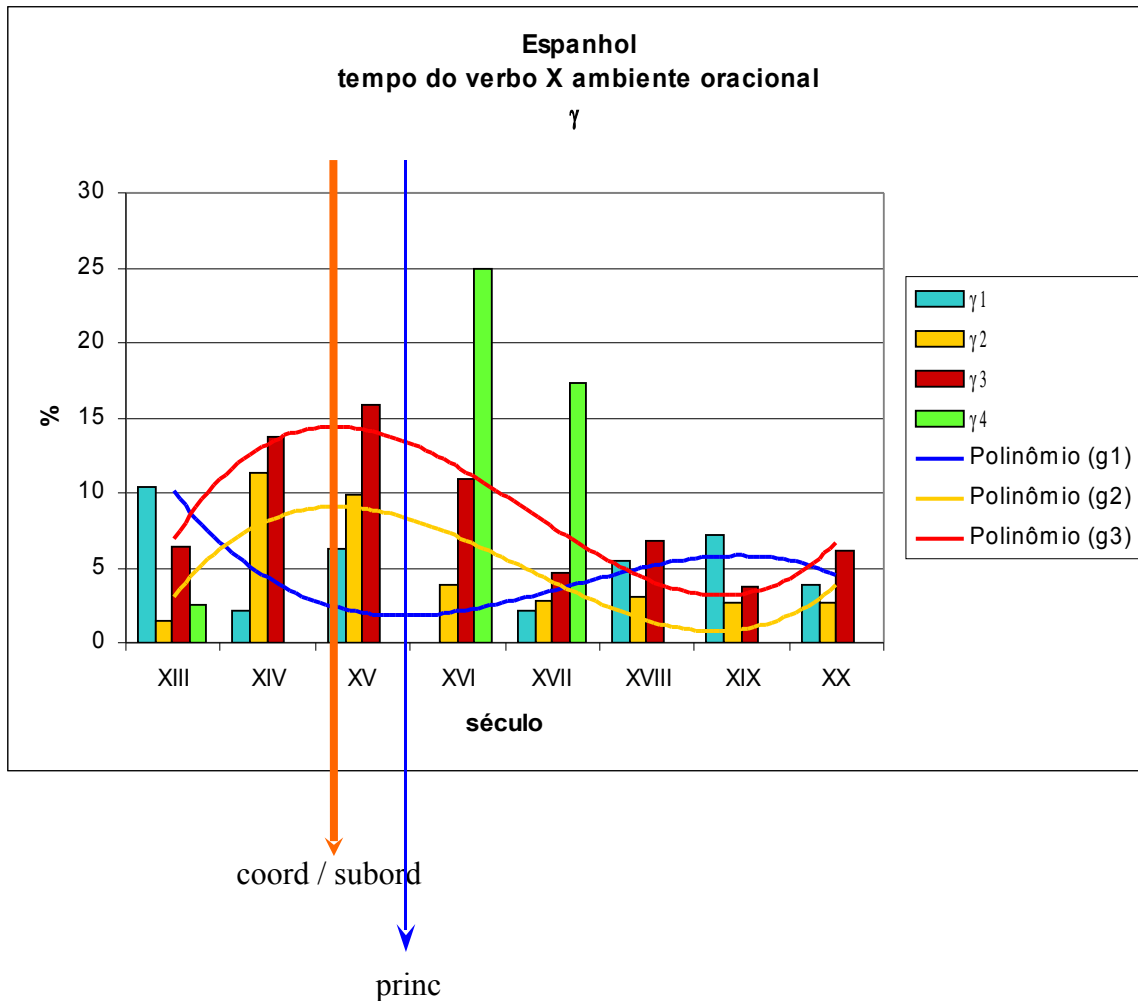


Gráfico E32d(T, normalizado) – distribuição das ocorrências de locuções verbais segundo o fator *ambiente oracional*.

As proporções segundo as quais ocorrem os tempos compostos no *corpus* espanhol são, como já foi dito, muito baixas. Não obstante, esboçam-se, também nessa escala reduzida, tendências contrárias para os comportamentos de orações principais (frequências crescentes) e subordinadas (frequências decrescentes), as duas classes mais comuns.

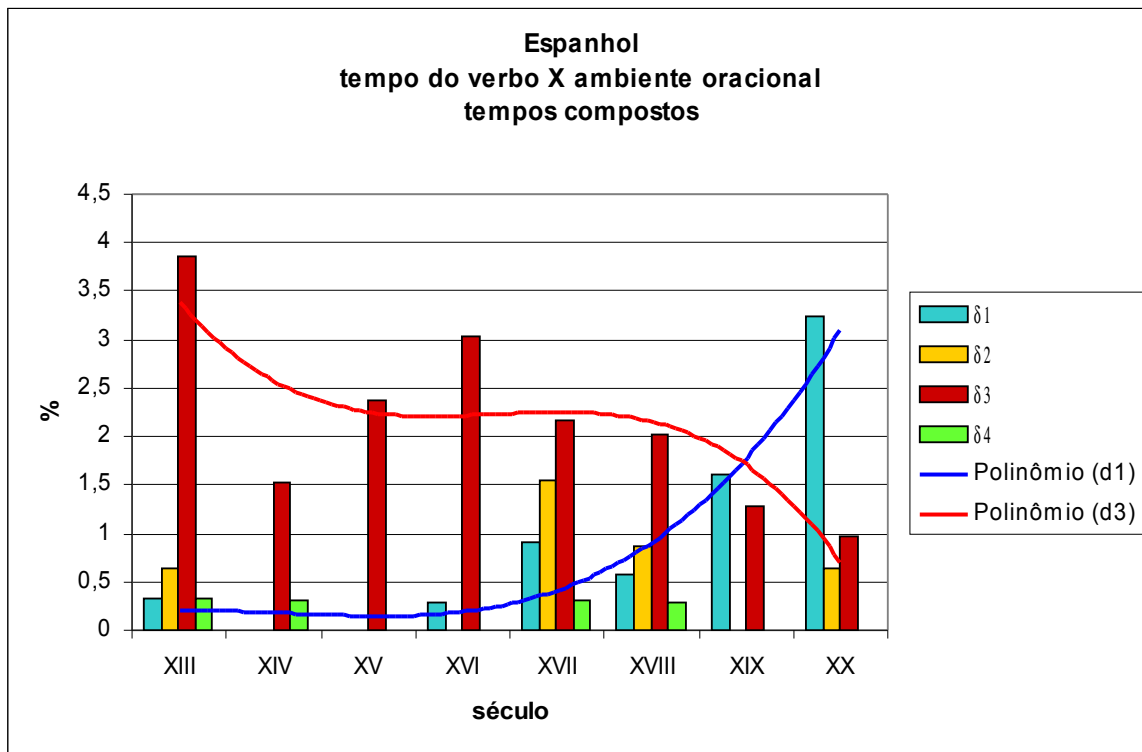


Gráfico E32e(T) – distribuição das ocorrências de tempos verbais compostos segundo o fator *ambiente oracional*.

As três conjugações mostram uma distribuição condicionada principalmente pela frequência das classes de ordenamento dos constituintes (cf. Gráfico E33a), além da distribuição das ocorrências segundo os ambientes oracionais (cf. Gráfico E7).

A distribuição das ocorrências segundo a classe flexional do verbo permite observar, inicialmente e de uma perspectiva mais ampla, que o sistema espanhol apresenta uma tendência ao favorecimento da primeira conjugação (I) diante da segunda (II) e da terceira (III), relação que se estabelece a partir do século XVI. A predominância das duas primeiras conjugações diante da terceira e a diminuição da diferença das frequências de I e de II também são características desse conjunto de ocorrências. Há, certamente, alguns itens lexicais que respondem majoritariamente por esse comportamento, como será exposto adiante; além disso, a menor expressividade da classe III é tão inequívoca quanto sua estabilidade em termos de frequência absoluta.

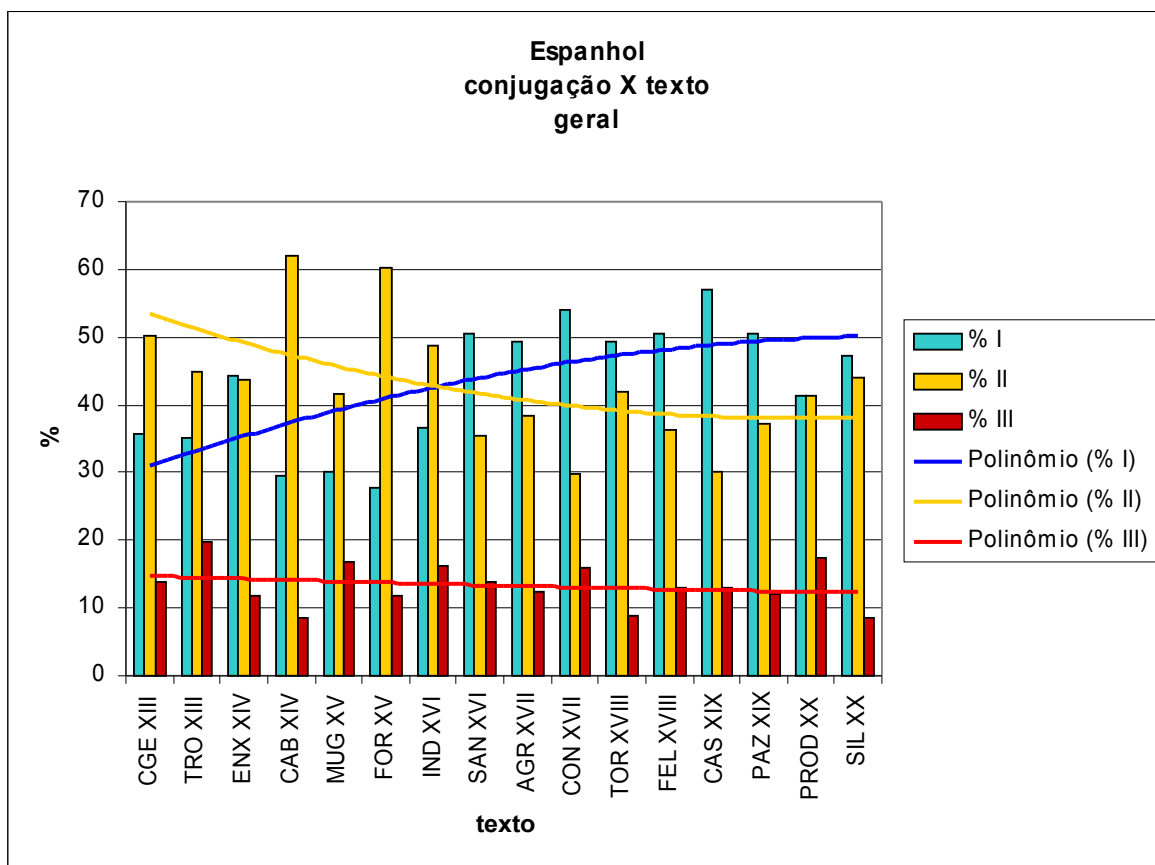


Gráfico E33a – distribuição das ocorrências segundo o traço *conjugação do verbo*.

Para a primeira conjugação, o tratamento numérico das ocorrências indica o aumento das frequências absolutas em estruturas **a** (iRRL(SVC) = .56) e **h** (iRRL(VS) = 2,49), mas um decréscimo discreto desses mesmos verbos em estruturas **g** (iRRL(SV) = -1,15) e naquelas cujo sujeito é elíptico, **i** (iRRL(VC) = -1,67) e **j** (iRRL(CV) = -1,13). Em termos de ambiente oracional e de ordem de constituintes, portanto, o aumento da proporção relativa da primeira conjugação no *corpus* espanhol (q.v. gráfico E7 acima) é determinado pelas orações principais e pelas ordens **a** (SVC) e **h** (VS).

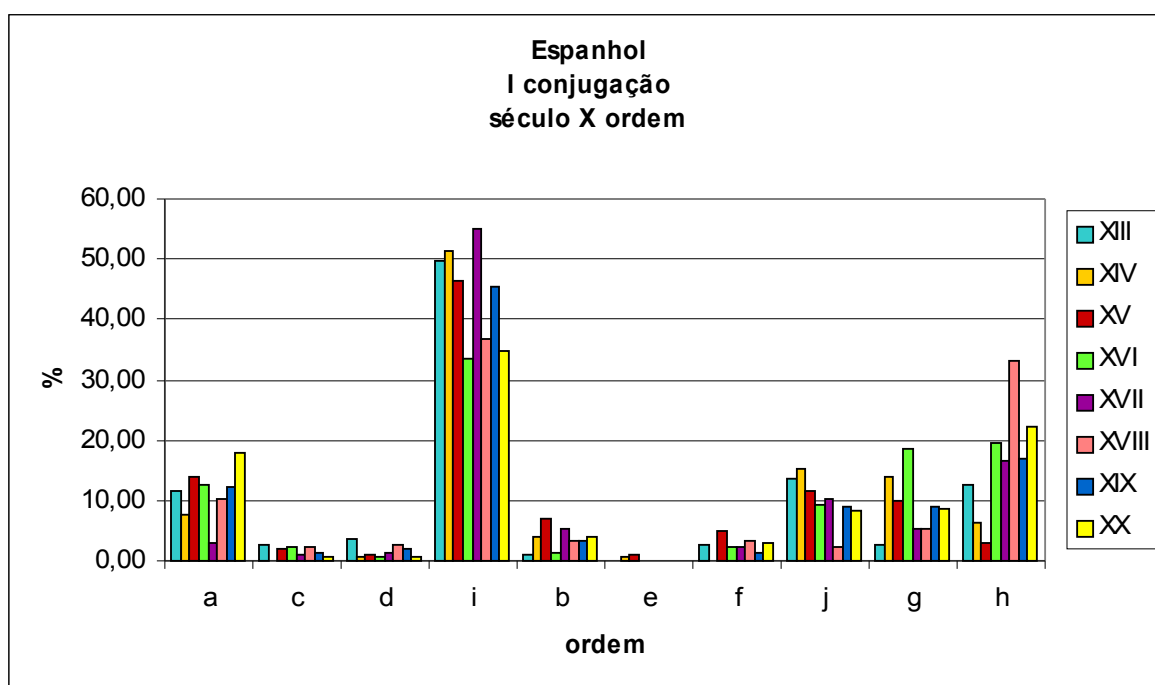


Gráfico E33b – distribuição das ocorrências de verbos de primeira conjugação segundo a ordem dos constituintes.

Tomando-se ainda a primeira conjugação, é interessante notar uma distribuição diferenciada com respeito ao ambiente oracional, como se pode depreender do exame do gráfico E33c. Nota-se que há um aumento da preferência por essa conjugação em orações principais, ao passo que, nas coordenadas e nas coordenadas em subordinação, há tendência à diminuição da frequência relativa de verbos de primeira conjugação; a mesma observação pode ser feita para as subordinadas, a despeito do máximo absoluto no século XV –

possivelmente em decorrência do mesmo efeito já tratado acima, no que tange à frequência relativa das subordinadas no *corpus*.

Compreendem exemplos associados a esse traço:

prosperó su hacienda, llegándola al aumento condeciente a su calidad [AGR]

No pocos infortunios agitaban el magnánimo corazón de Luis XIV [FEL]

como siguen los ojos el de una mariposa a examinar la región que habitan [CAS]

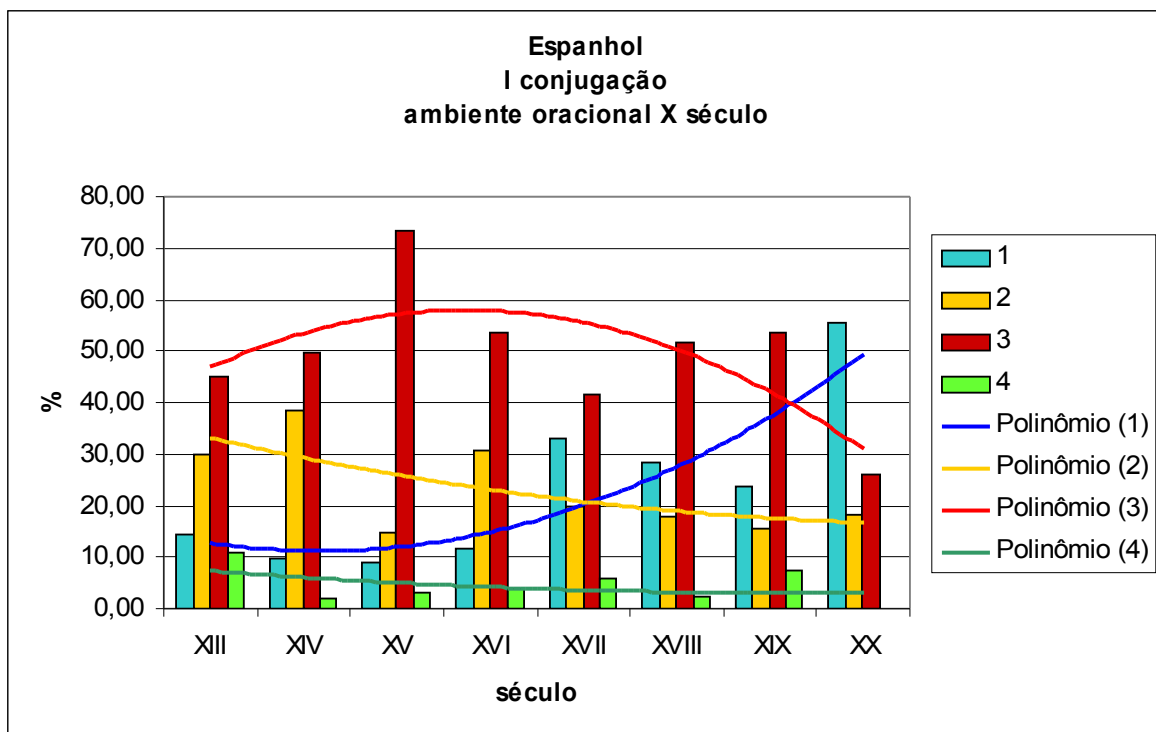


Gráfico E33c – distribuição das ocorrências de verbos de primeira conjugação segundo o ambiente oracional.

No que diz respeito à ordem dos elementos de orações baseadas em verbos da classe II, nota-se, ao exame do gráfico que representa o comportamento do conjunto, a mesma tendência ao aumento das proporções relativas das ordens SVC (**a**) e VS (**h**); entretanto, a amplitude da mudança para a ordem SVC é maior do que na primeira conjugação. As diminuições das frequências relativas das ordens VC (**i**), CV (**j**) e SV (**g**), também indicadas para a primeira conjugação, ocorrem na classe II, mas em menor intensidade. Veja-se ainda que, apenas considerando-se a segunda conjugação, há diminuição das proporções totais em relação às classes verbais I e III (q.v. gráfico E33a) – mostrando que o aumento da II conjugação em orações principais não foi suficiente para fazer frente à classe I.

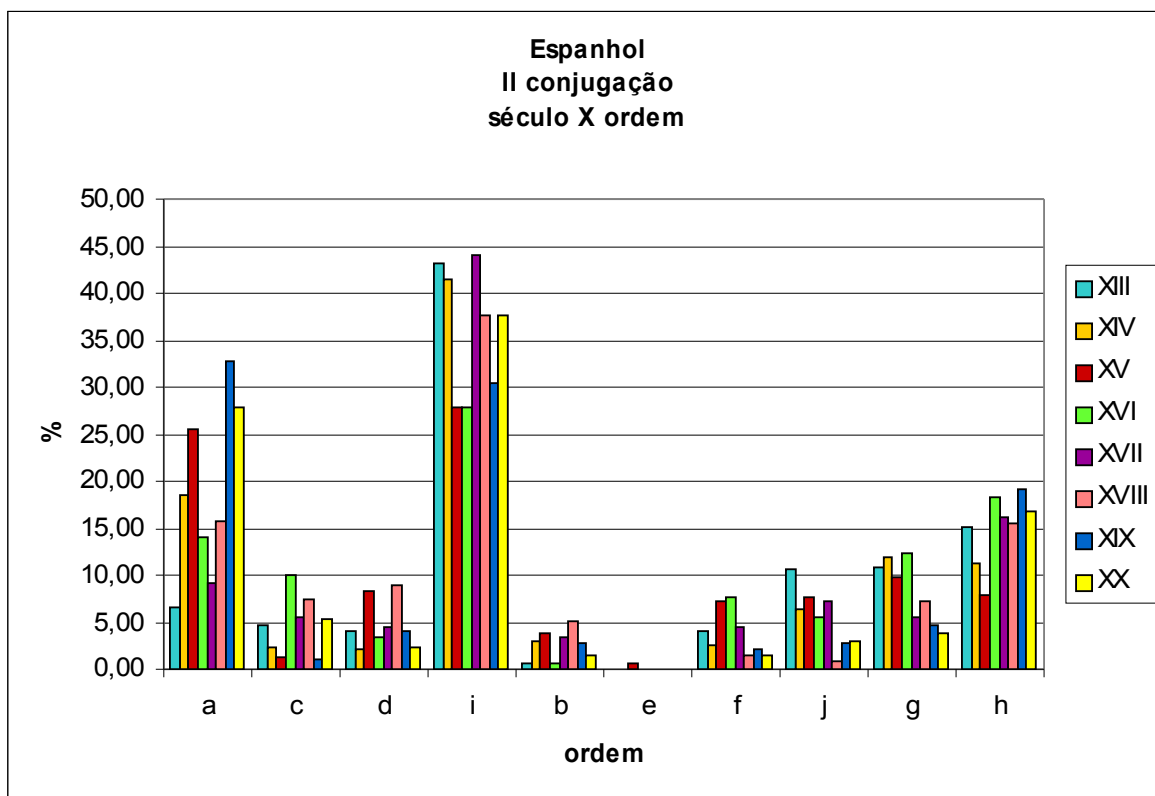


Gráfico E33d – distribuição das ocorrências de verbos de segunda conjugação considerando-se a ordem dos constituintes.

A segunda conjugação, quando analisada pelos mesmos critérios, mostra tendências diacrônicas bastante semelhantes às da primeira conjugação – marcadamente, o aumento considerável da frequência relativa dos verbos que a compõem em orações principais, o qual é acompanhado de decréscimo nos demais tipos oracionais. O aumento apontado, na segunda conjugação, para todos os tipos oracionais é de maior amplitude do que o observado na primeira, podendo-se fazer idênticas considerações acerca da diminuição obtida para os demais ambientes oracionais.

São exemplos de ocorrências associadas a essa conjugação:

No nacía este rebelión de aquel común alivio que sienten los muchachos con el ocio [TOR]

Tenía un despacho en la calle baja de San Pedro [PROD]

A la salida del pueblo aparece una gran pirámide truncada [SIL]

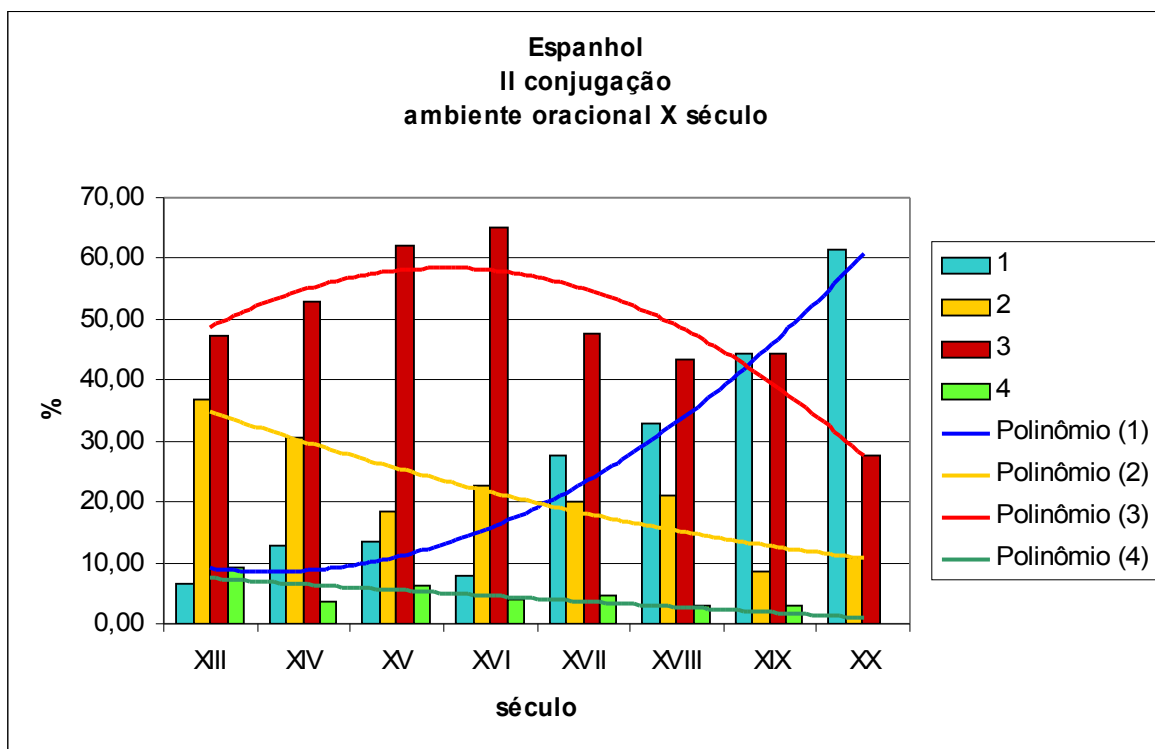


Gráfico E33e – distribuição das ocorrências de verbos de segunda conjugação considerando-se o ambiente oracional.

É interessante notar, com relação à classe III, que os comportamentos das ordens SVC (**a**), VC (**i**) e VC (**h**) são contrários aos observados para as classes I e II: as ordens **a** (SVC) e **h** (VS) verificam uma diminuição das proporções relativas de verbos da classe III, enquanto a ordem **i** (VC) mostra uma elevação dos valores de frequência relativa interna nos textos.

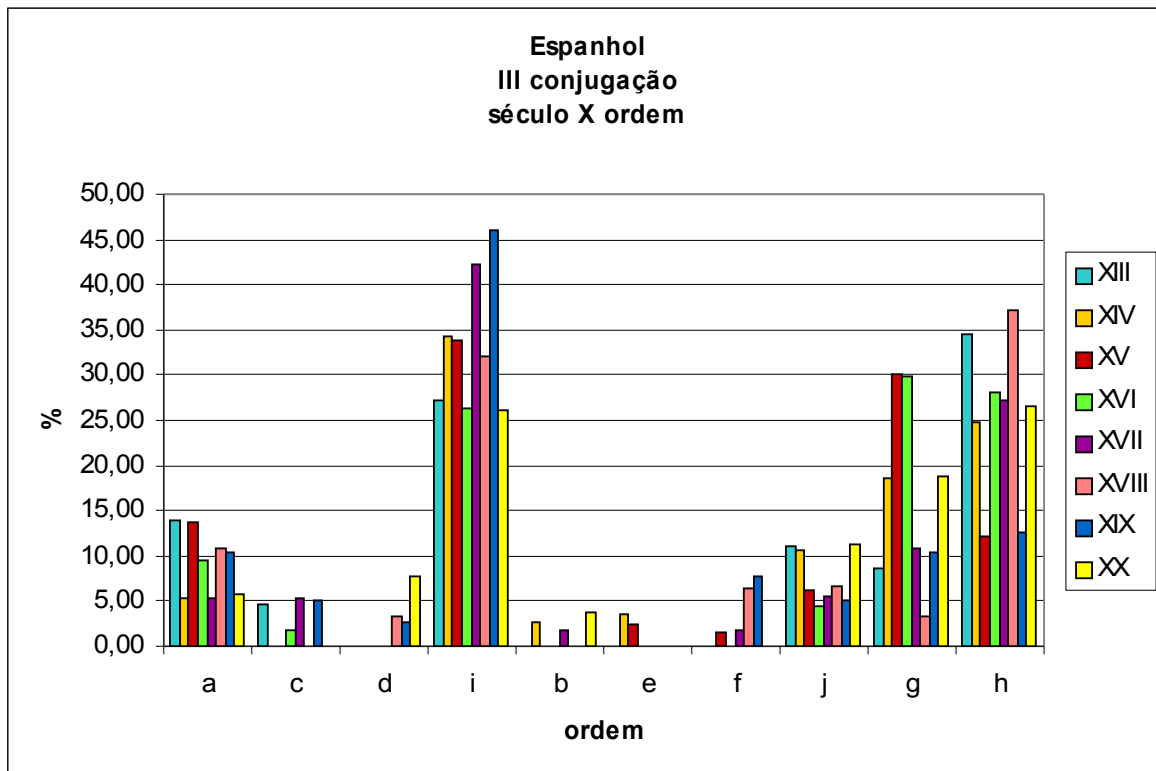


Gráfico E33f – distribuição das ocorrências de verbos de terceira conjugação segundo a ordem dos constituintes.

A frequência relativa dos verbos de terceira conjugação em função do ambiente oracional obedece aos mesmos padrões já apontados para as demais classes verbais: internamente, observa-se aumento das proporções relativas dessa classe em orações principais, concomitantemente à diminuição dessa mesma classe III nos outros ambientes oracionais. Essas observações devem ser feitas à luz do gráfico E33a, que mostra as relações de proporcionalidade existentes entre as classes I, II e III no *corpus* espanhol.

São exemplos atribuídos à terceira conjugação:

que por esta providencia se añadió después la contracción de Jesús [AGR]
y luego me recibí por paje de rodela el capitán Felipe de Menargas [CON]
la prosecución de este bien les prohíbe casi siempre publicar los sucesos [PAZ]

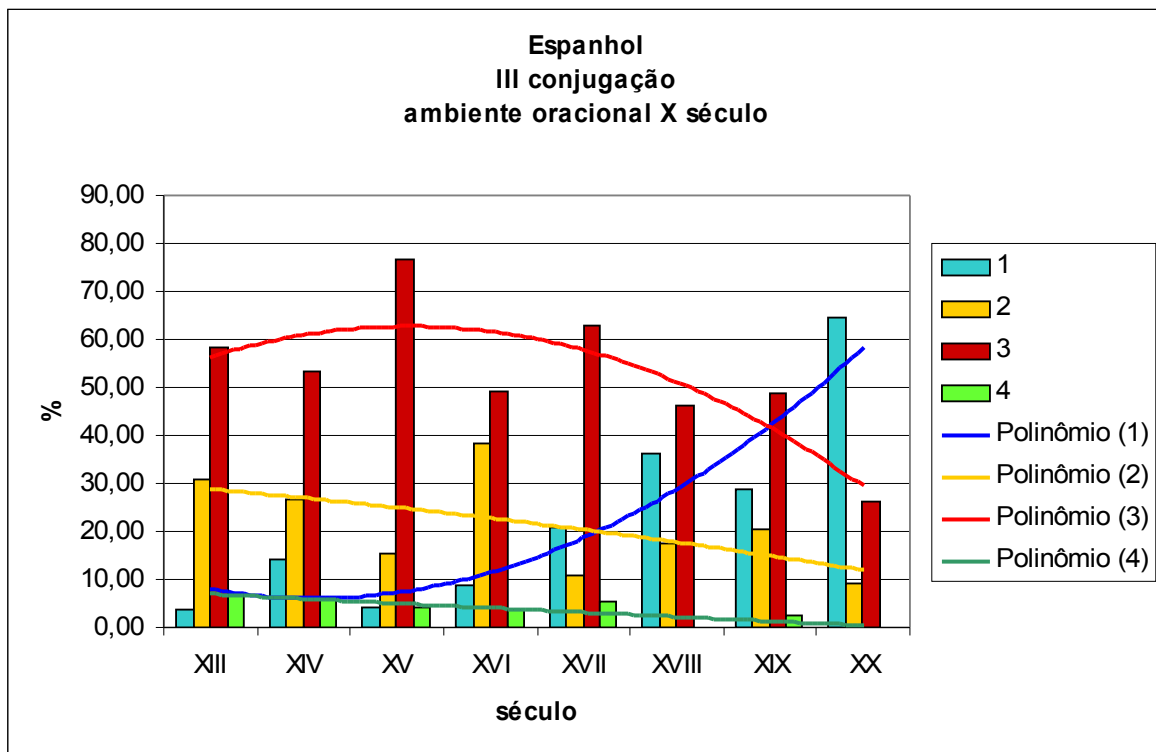


Gráfico E33g – distribuição das ocorrências de verbos de terceira conjugação considerando-se o ambiente oracional.

A fixação de um dos graus de liberdade da análise – *ambiente oracional* – oferece a possibilidade, como já feito anteriormente, de se estudar o peso desse fator na distribuição das ocorrências das três classes verbais. Vê-se, no Gráfico E33d, que o ambiente oracional é, mais uma vez e como proposto por Bybee 1991, um fator condicionador do comportamento da amostra: em primeiro lugar, há o mesmo padrão crescente, em degraus, para as classes mais expressivas do gráfico. Em decorrência disso, as curvas referentes às orações principais são, todas, ascendentes – apesar de haver uma diferença sensível na expressividade das conjugações. As curvas da primeira (C1) e da segunda (C2) conjugações se sobrepõem, mas a da terceira (C3) corre bem abaixo delas. Tomem-se os exemplos:

C1 – los troyanos otrosy allegaron muc[has] gentes de muchas partes - TRO XIII

C2 – si queréis fazer una arca, concurren estas quatro causas necesarias - FOR XV

C3 – de aquel común alivio que sienten los muchachos con el ocio - CON XVII

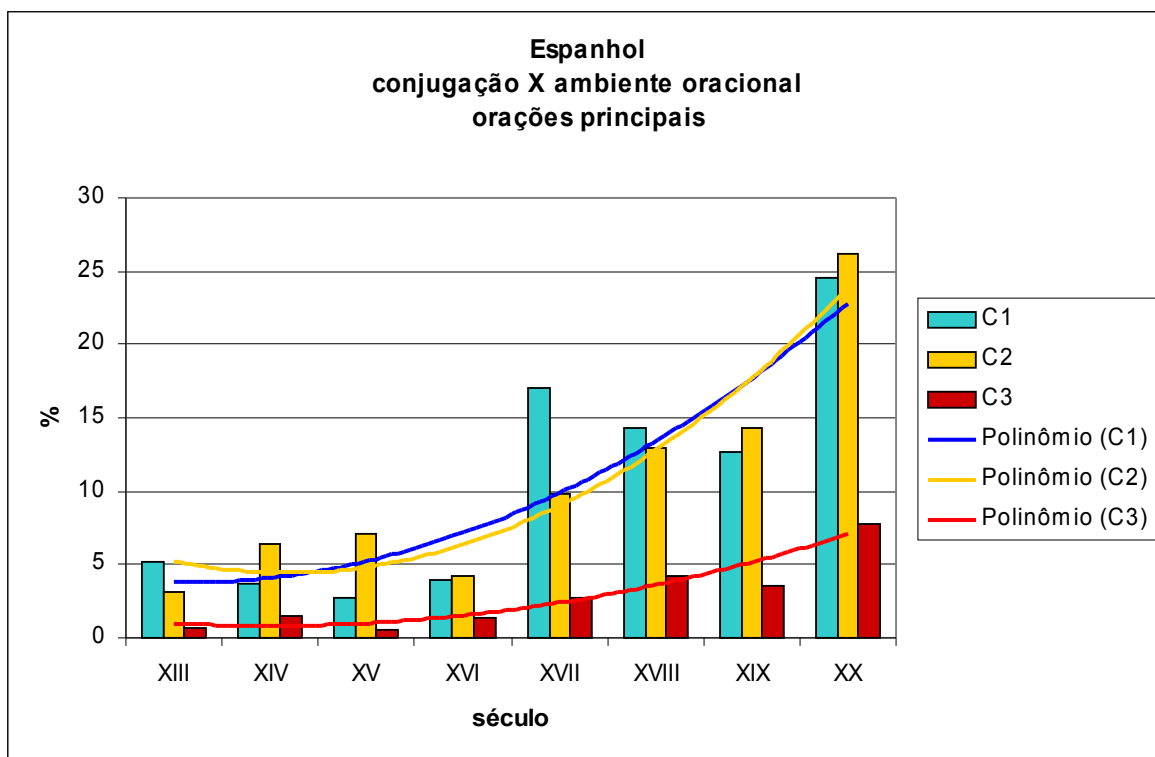


Gráfico E33h – distribuição das ocorrências de orações principais segundo a vogal temática do verbo.

No caso das orações coordenadas, o perfil descendente é mais acentuado para os verbos de segunda classe (C2), enquanto os demais (C1, C3) têm curvas mais suaves, subhorizontais. Isso indica que, para as coordenadas, a diminuição mais dramática das freqüências absolutas está associada a verbos de vogal temática -e-. Os exemplos seguintes foram colhidos no *corpus* espanhol:

C1 – et envíalo al infante don Johan, arzobispo de Toledo - CAB XIV

C2 – encubrirle los efectos de su asistencia y suspender sus regalos - AGR XVII

C3 – y se introdujo el hambre en los países más fríos - FEL XVIII

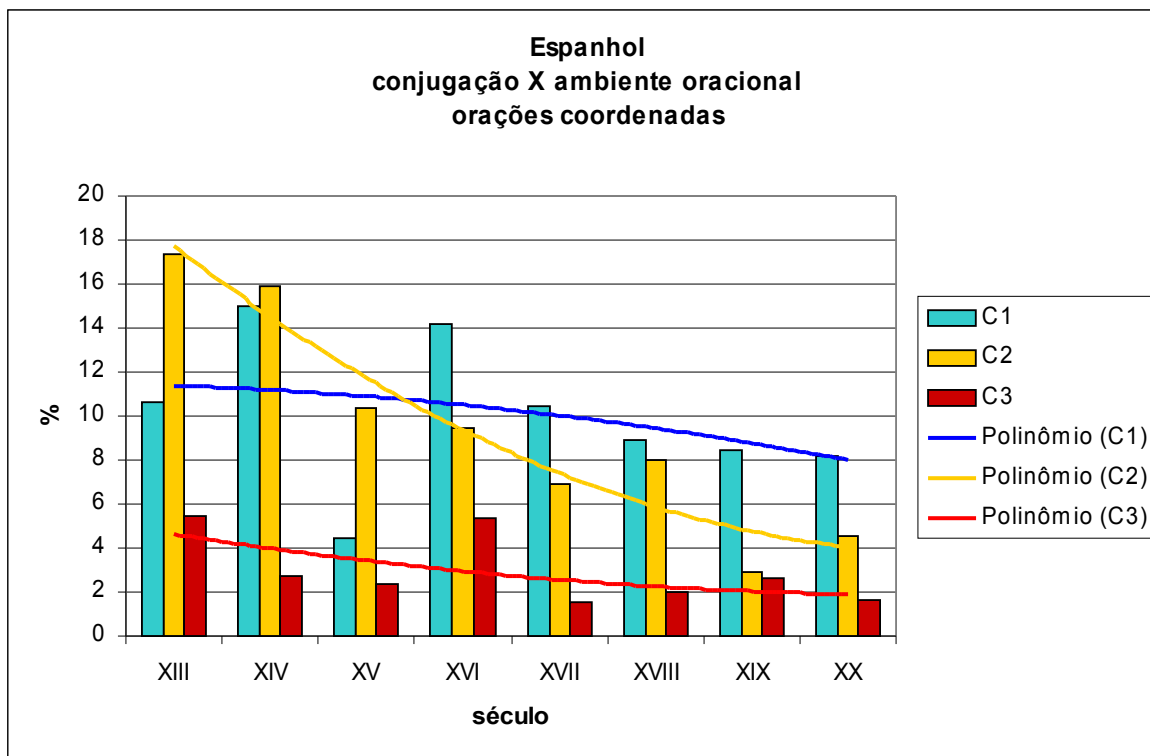


Gráfico E33i – distribuição das ocorrências de orações coordenadas segundo a vogal temática do verbo.

Para as subordinadas, tem-se um dado interessante: as curvas mais expressivas, relativas às classes temáticas I e II, mostram uma defasagem de cerca de $\frac{1}{4}$ de período – o que, no gráfico, equivale a aproximadamente 300 anos entre os máximos das curvas C2 e C1. Esse fato pode contribuir para explicar o motivo de se terem dois picos nas curvas de subordinadas em vários dos gráficos produzidos pela análise numérica. Vejam-se os exemplos e o gráfico de distribuição dos dados:

C1 – en que todas las cosas se engendran y crian del tamo y átomos - IND XVI

C2 – donde se formaron ... varias compañías para traer trigo de Levante - FEL XVIII

C3 – porque sólo acudían tipos patibularios al despacho - Prod XX

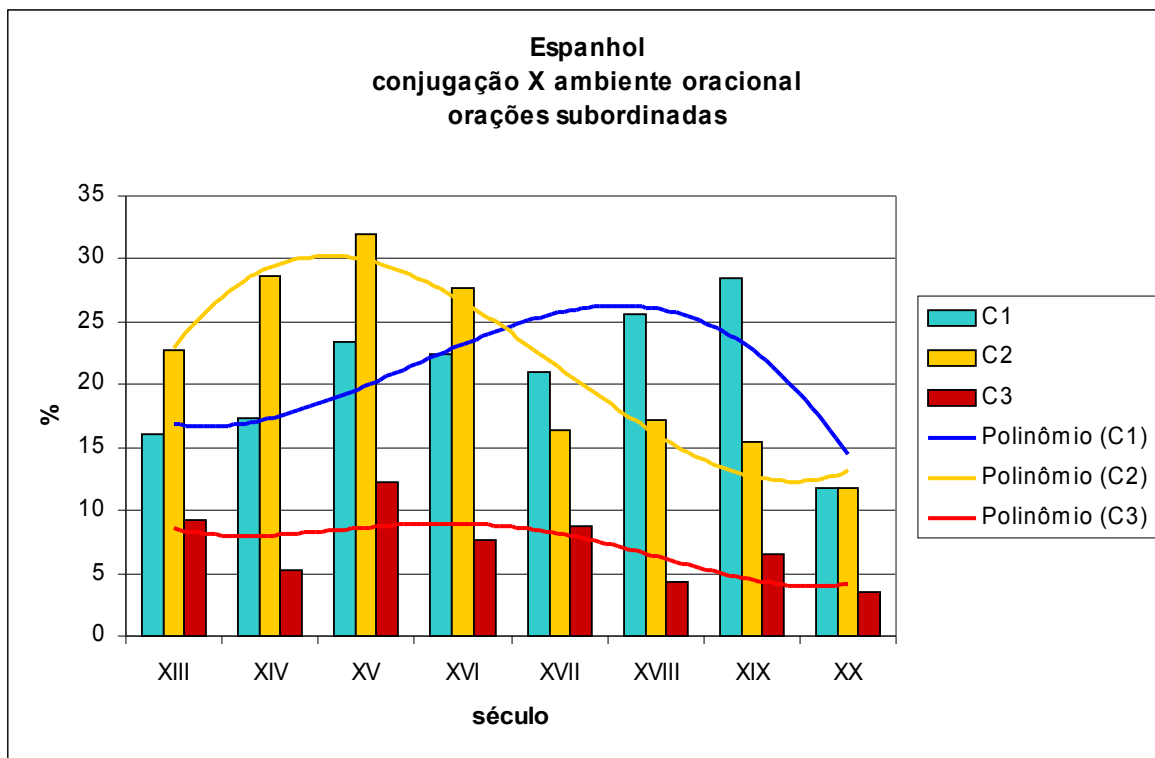


Gráfico E33j – distribuição das ocorrências de orações subordinadas segundo a vogal temática do verbo.

A Tabela E7 sintetiza os comentários feitos acima sobre o comportamento das classes de conjugação do verbo diante dos fatores considerados. Observem-se as tendências de III para as ordens **i**(VC) e **h** (VS), que são contrárias às das classes I e II.

<i>conj</i>	<i>1</i> (<i>princ.</i>)	<i>2</i> (<i>coord.</i>)	<i>3</i> (<i>subord.</i>)	<i>4</i> (<i>coord.</i> - <i>subord.</i>	<i>a</i> (<i>SVC</i>)	<i>i</i> (<i>VC</i>)	<i>j</i> (<i>CV</i>)	<i>g</i> (<i>SV</i>)	<i>h</i> (<i>VS</i>)
<i>I</i>	↗	↘	↘	→	↗	↘	↘	↘	↗
<i>II</i>	↗	↘	↘	→	↗	↘	↘	↘	↗
<i>III</i>	↗	↘	↘	→	↘	↗	.	.	↘

Tabela E7 – resumo das tendências gerais de comportamento das classes I, II e III de conjugação verbal segundo os critérios *ambiente oracional* e *ordem de constituintes*.

Com respeito aos itens lexicais, trataremos, aqui, apenas dos mais freqüentes – no *corpus* espanhol, **ser** y **hacer**. Observe-se que, da freqüência mais alta para a mais baixa, as cinco maiores são representadas pelo verbo **ser**; a sexta, por **hacer**. Dentre os itens mais freqüentes, portanto, a seqüência é **ser** > **hacer** > **saber** > **decir** > **tomar** > **haber**; os de vogal temática -e-, portanto, fazem larga frente aos de terceira e de primeira conjugações.

Pois bem, as duas curvas mais expressivas (as demais estão sempre abaixo dos 10% de freqüência absoluta) têm padrões gráficos bastante distintos um do outro: enquanto as proporções de **ser** variam mais de 100% e alcançam um mínimo entre os séculos XVII e XVIII, a curva de **hacer** tem um perfil sempre decrescente, mas mais suave. em outras palavras: **ser** varia em freqüência, mas segundo um padrão definido; **hacer** sempre diminui em freqüência absoluta.

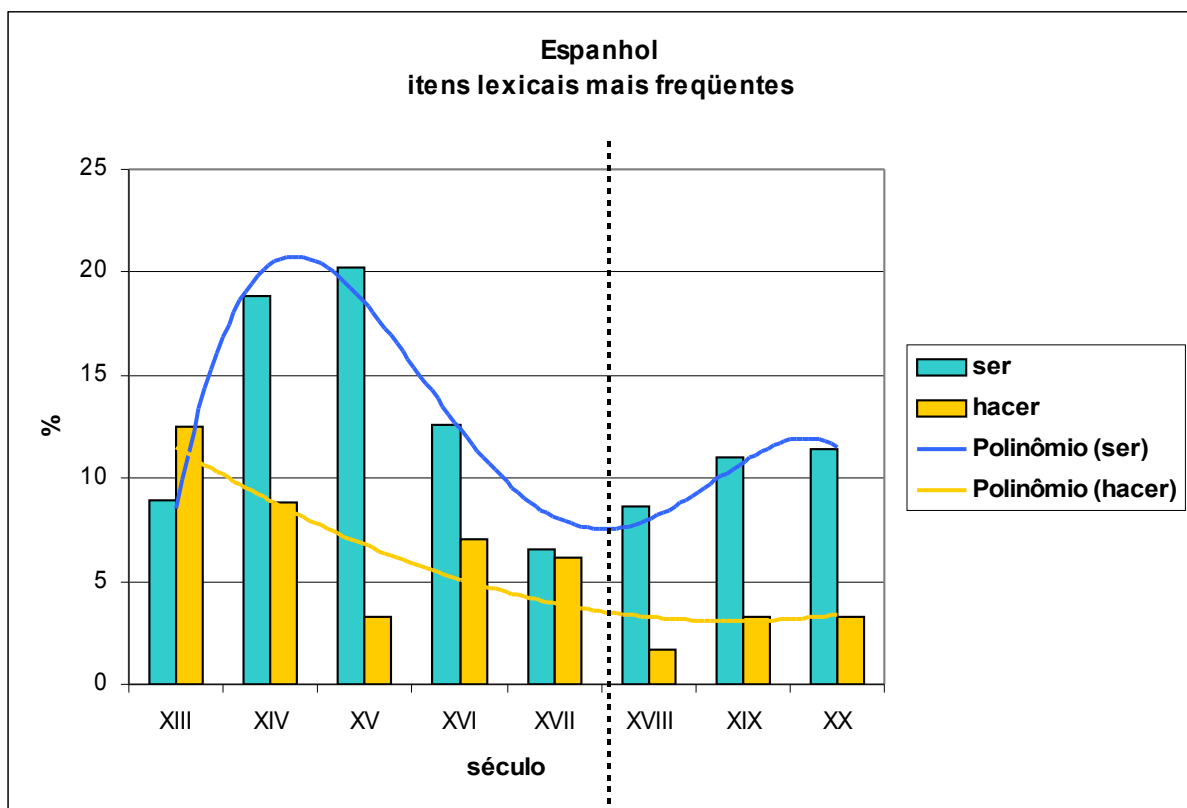


Gráfico E34 – distribuição dos itens lexicais mais freqüentes no *corpus* espanhol.

Em E35, apresento os resultados da distribuição dos dados segundo a frequência dos itens lexicais; o perfil hiperbólico é idêntico ao obtido para os dados de português (q.v. Gráf. P14l), e se aproxima do referente a uma função do tipo $f(x) = 1/x$.

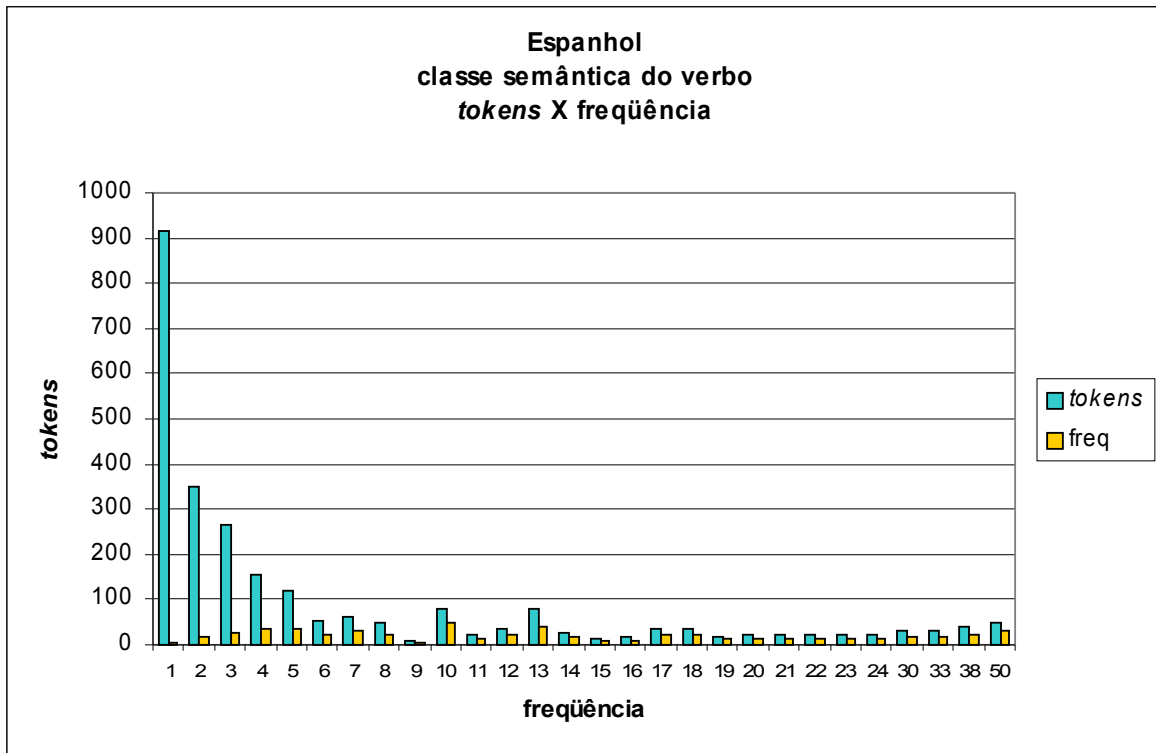


Gráfico E35 – distribuição dos verbos segundo sua frequência no *corpus* espanhol.

Com respeito à distribuição do item mais freqüente, 'ser', tem-se uma distribuição bastante semelhante ao perfil traçado no Gráfico E7 para o fator *ambiente oracional*, como

já assinalado, para o *corpus* português, no Gráfico P14m: aumentam, portanto, as frequências relativas desse item lexical em orações principais (1), enquanto se observa a diminuição das proporções em coordenadas (2) e em subordinadas (3). A similaridade dos Gráficos P14m e E36 sugere, mais uma vez, a necessidade de se considerar o *ambiente oracional* até mesmo como fator relacionado a processos que envolvam a frequência de itens lexicais. Como exemplos de ocorrências do verbo *ser* nos ambientes sintáticos selecionados, tomem-se os seguintes:

- 1 – Cada uno destes son muy buenos, en que - CAB XIV
- 2 – e causa destes es natura, onde estos se llaman dones - FOR XV
- 3 – dijeron que el rey don Sancho era en Sevilla, non le plogo - SAN XVI

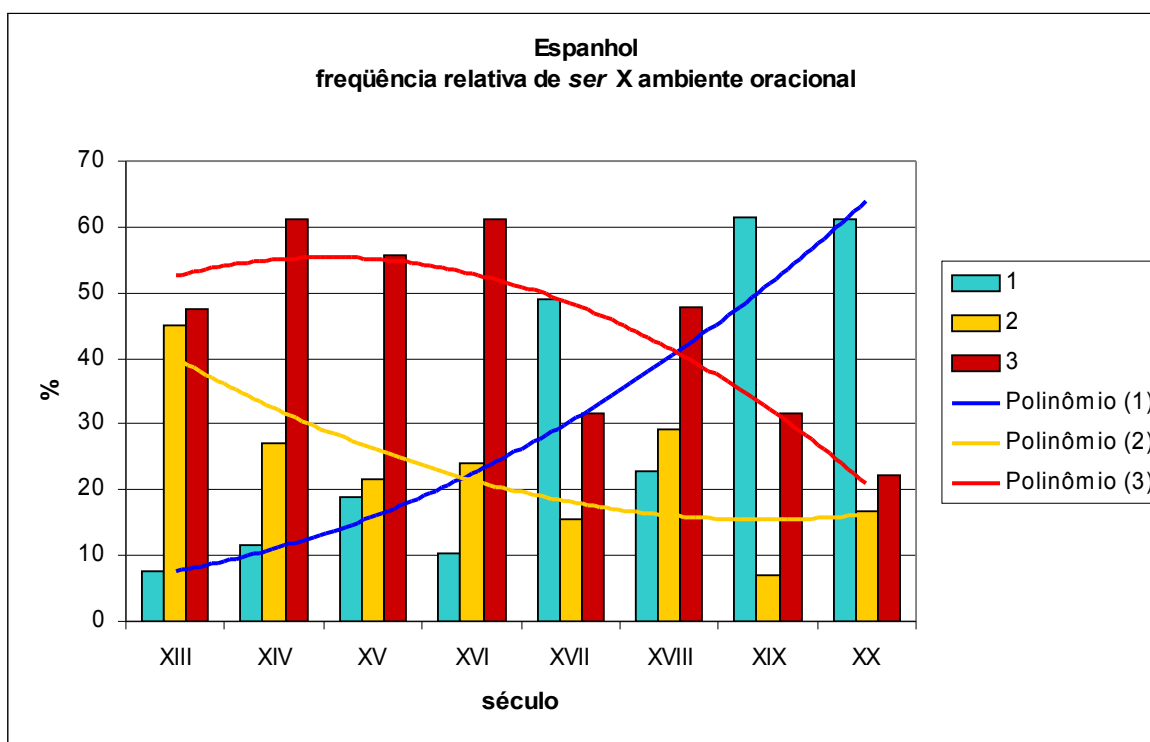


Gráfico E36 – distribuição de frequências relativas do verbo *ser* segundo o *ambiente oracional*.

O fator *classe semântica do verbo*, por fim, mostra uma especialização das ocorrências na categoria DISP – especialmente para as ordens [VC] mais comuns, i.e., **a** (SVC) e **i** (VC). As frequências absolutas desse subconjunto de dados mostram discreta tendência a aumentar ($iRRL(DISP,VC) = 0,41$), ou seja, essa classe semântica é, pouco a pouco, mais comum na amostra. Dado o elevado número de categorias, há uma pulverização dos dados, o que torna difícil discernir com segurança tendências de comportamento das ocorrências para uma base de dados de dimensões reduzidas. Entretanto, esboçam-se padrões ascendentes em ENAC ($iRRL(VC) = .44$) e descendentes em COGN ($iRRL(VC) = -.31$) e em DISP ($iRRL(CV) = -.30$). vejam-se ainda alguns exemplos dessas classes:

- ATTR – conuiene que uos digamos primero quamanna es Europa - CGE XIII
- COGN – Cuando él oyó esto, non pensando el engaño abscondido - ENX XIV
- DISP – con que ninguno dellos pueda hallar las obras que él mismo obró - IND XVI
- ENAC – siendo menester seguir aquellas quatro leguas mortales - CON XVII
- SEM – entre rectas y rasgos dibuja disimuladamente una cifra o una fecha - CAS XIX

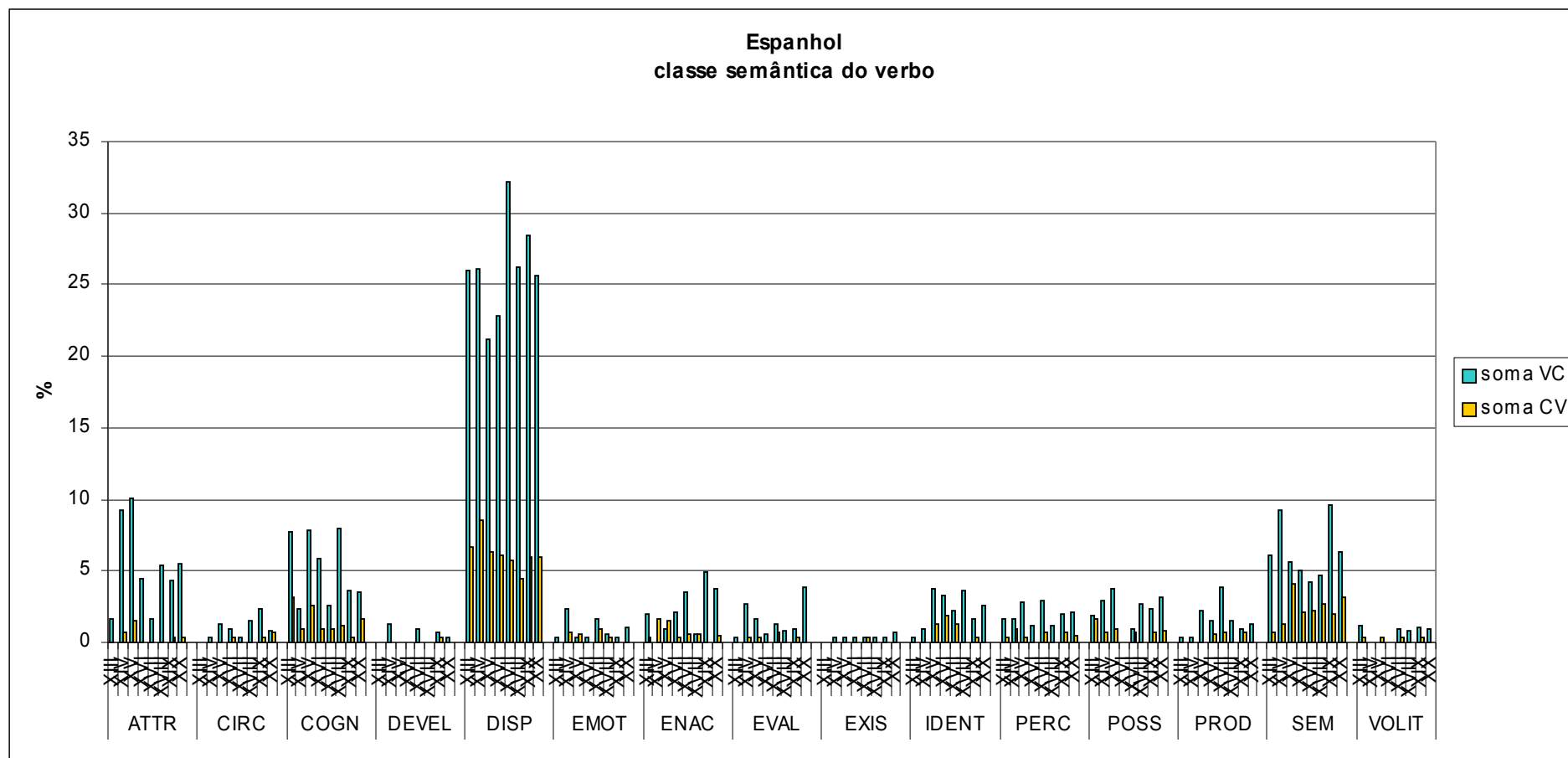


Gráfico E35 – distribuição das ocorrências segundo a classe semântica do verbo

Em linhas gerais, as tendências observadas no *corpus* espanhol são semelhantes às apresentadas pelo conjunto português: diversos fatores são claramente controlados pela ordem de constituintes – e.g., *agentividade* (cf. gráfico E1), *peso de S e de C* (cf. gráficos E21 e tabela E4), *regência do verbo* (cf. gráficos E28b-g e tabela E6), *conjugação do verbo* (cf. gráficos E33b, d, f e tabela E7), *tempo do verbo* (cf. gráficos E29) e *classe semântica do verbo* (cf. gráfico E35), distribuindo-se de acordo com as categorias sintáticas de **a** (SVC) a **j** (CV). Outros, além de mostrarem comportamento que é função da ordem de constituintes, também são condicionados pelo *ambiente oracional* – e.g., a própria *ordem de constituintes* (cf. gráficos E8-E11 e tabela E2), *peso de S e de C* (cf. a série de gráficos E22-E27 e a tabela E5), *tempo verbal* (cf. gráficos E31-E32), *conjugação do verbo* (cf. gráficos E33c, e, g-j e tabela E7) e *item lexical* (para o verbo *ser*, o mais freqüente na amostragem, como observado para o português; cf. gráfico E36). A exemplo do que foi mencionado na análise dos dados do português, o *ambiente oracional* é função apenas do tempo, como mostra o gráfico E7. Os perfis de evolução das freqüências relativas dos demais fatores mostram que este é determinante na caracterização das mudanças também no sistema espanhol.

Tratamento numérico das ocorrências: o *corpus* italiano

No que respeita à posição relativa dos constituintes A e N internos ao SN, o *corpus* italiano exhibe o mesmo comportamento que seus pares português (Cohen 1986/1990) e espanhol (Totaro 1996). A base de dados compreende cerca de 2000 ocorrências¹³ colhidas nos seis séculos da amostragem, perfazendo pouco mais de 150 SNs por texto. São exemplos de ocorrências encontradas neste estudo:

século	texto	ocorrência
XV	SERM	Incominciano i deuotissimi sermoni del diuo Bernardo Abbate di sicontiene la sōma di tutte leuirtute necessarye 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 :
	PIGN	col ritrouarsi in tutte quelle importanti espeditioni ,che si faranno a i suoi di, ogni Et sopra il tutto douerà far ogni opera per interuenire ne' consigli secreti .
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 essercito : e giunto verilmente gli era uscito contra con l' essercito romano , si conuennero
XVIII	PRIN	tre principali punti ; che si dia Provedenza divina ; che l'anime umane sien' e che si debbano moderare l' umane passioni con la Giustizia , e da quella si
	LETT	raramente sono giudici competenti , illuminati , e caldamente spassionati Mille ostacoli impediscono il retto giudizio degli altri ; ma , freddato interamente
XIX	ITIN	di Roma . Essa è d' ordine corintio , tutta incrostata di buoni marmi , e di Roma . Essa è d' ordine corintio , tutta incrostata di buoni marmi , e
	GEOG	non abbiamo ancora alcun pezzo osteologico , in discreto stato di conservazione non abbiamo ancora alcun pezzo osteologico, in discreto stato di conservazione
XX	GRAA	contraddica la possibilità che la non esplicita identificazione del testo de Chrétien riassume con chiarezza i momenti cruciali relativi al Graal e alla lancia
	NOME	due capi, che ispirarono loro delle false teorie , un sacerdote che era stato L'immagine di un uomo dalla vita avventurosa , capace anche di uccidere

Tabela II – exemplos de ocorrências de AN e de NA nos textos do *corpus* italiano.

As proporções individuais de AN e de NA foram somadas século a século, e as médias foram lançadas no gráfico II abaixo.

¹³ Seria interessante estudar também os condicionadores estruturais da variação entre adjetivos e formas encabeçadas por preposição, como 'regioni del mare' X 'regioni maritime'.

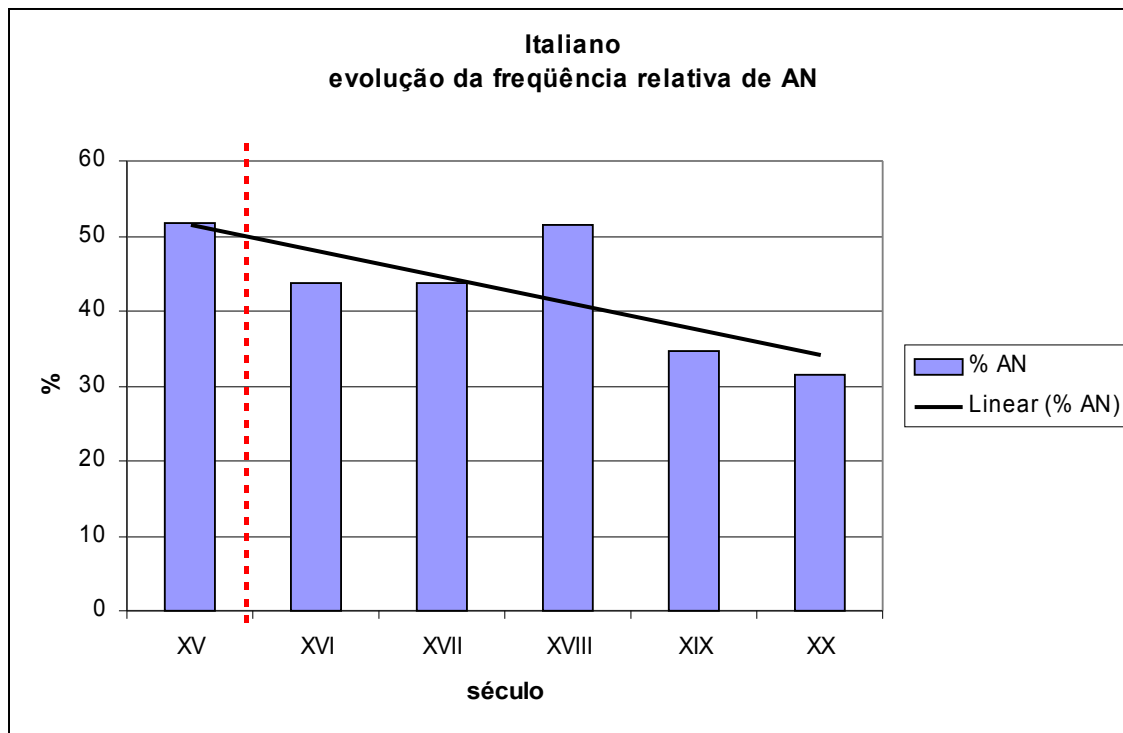


Gráfico II – evolução das médias gerais de AN por século nos textos utilizados para amostragem.

Nota-se, primeiramente, que a proporção de AN diminui no sentido positivo das abscissas, o que é ressaltado pela reta que exprime a regressão linear da distribuição das ocorrências (no gráfico, em preto). Além disso, vê-se que essa mesma reta intercepta a linha de 50% próximo à passagem do século XV para o XVI (no gráfico, pontilhado em vermelho); esse resultado é interessante, uma vez que o francês cruzou esse mesmo patamar no século XVI (von Wartburg 1966); o espanhol, no século XVII (Totaro 1998), e o português, no século XVIII (Cohen *op. cit.*).

É necessário salientar, ainda, que o pico correspondente ao máximo absoluto da amostra coincide com o século XVIII. Trata-se do efeito¹⁴ da proporção elevada de AN em um único texto – de fato, da proporção mais alta dentre todos os textos aqui utilizados –, *Del Principe e delle lettere*, que se mostra, diante dos demais, realmente excêntrico com respeito a esse traço sintático.

¹⁴ A exclusão desse texto da amostra teria como efeitos mais relevantes a) a resolução dos dados segundo um perfil francamente linear e b) o recuo, para meados do século XV, do momento previsto para a mudança tipológica AN → NA.

O gráfico I1 foi elaborado a partir das proporções brutas de AN e de NA; deixa-se em aberto o questionamento sobre a distribuição fina dos constituintes A e N segundo fatores controladores de sua posição relativa, como os estudados por Cohen *op. cit.*, Totaro *op. cit.* e Lima 2003.

No nível oracional, verifica-se a predominância das ordens SVC (**a**), VC (**i**), SV (**g**) e VS (**h**). As duas primeiras, que trazem o complemento verbal explícito, são mais freqüentes e contribuem para que o grupo [VC] seja bastante expressivo; deve-se ressaltar que o valor positivo obtido para a inclinação da reta de regressão linear (iRRL) do grupo [VC] (0,98) é devido à combinação da inclinação positiva da ordem a (iRRL (SVC) = 2,04) e da negativa de **i** (iRRL (VC) = -0,76); em outras palavras, ainda que a freqüência da ordem **i** (VC) tenda a diminuir, a proporção absoluta de ocorrência da superclasse [VC] aumenta graças à elevação da freqüência das ocorrências de estrutura **a** (SVC).

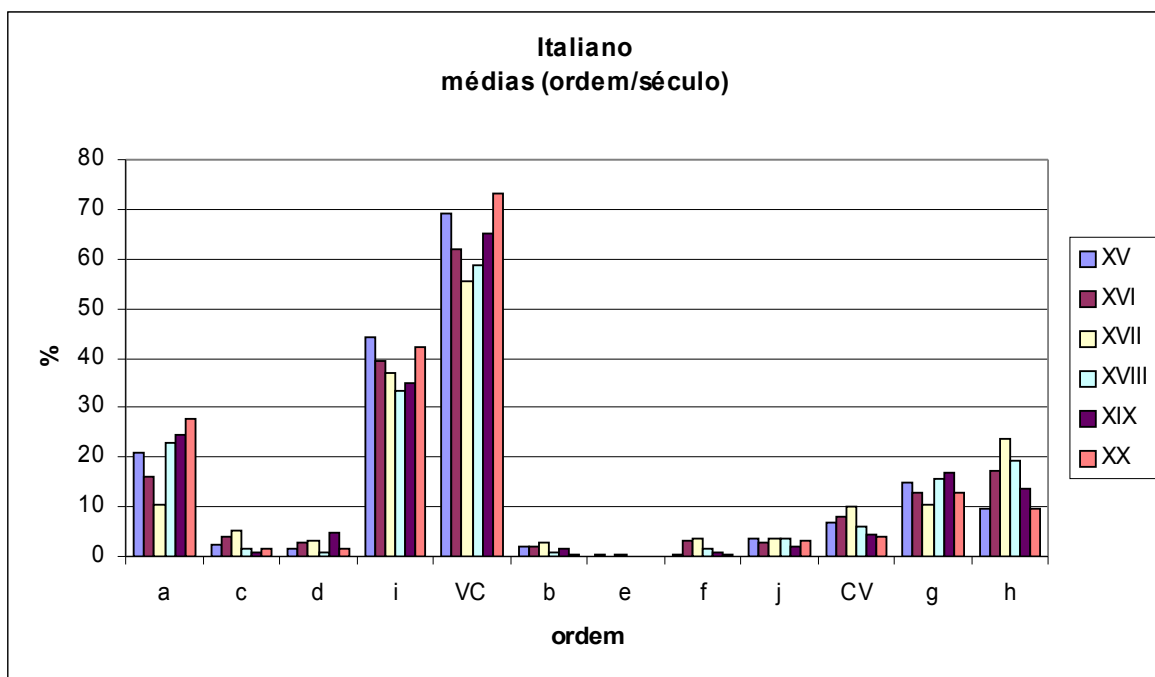


Gráfico I2 – médias gerais por século segundo a ordem dos constituintes S, V e C e dos grupos VC e CV.

O fator *agentividade*, em linhas generalíssimas, distribui-se preferencialmente segundo as classes **Ah** e **Pn**, e ambas apresentam freqüências gerais crescentes no *corpus*: $iRRL(VC) = 1,83$; $iRRL(CV) = 1,51$. Observe-se que esses resultados indicam uma tendência oposta à do *corpus* português, em que os valores de **Ah** e de **Pn** são decrescentes; o *corpus* espanhol também mostra decréscimo generalizado em **Ah**, mas aumento em **Pn** – como se, de fato, os conjuntos de dados português e italiano corresponderem a posições polares no universo de amostragem, e o espanhol estivesse a meio caminho. Nesse caso, contudo, não está claro se a tendência à mudança em **Ah** precedeu a que se observa em **Pn** ou se foi o contrário; a identificação de uma possível direção de mudança entre esses traços demanda, portanto, estudos posteriores.

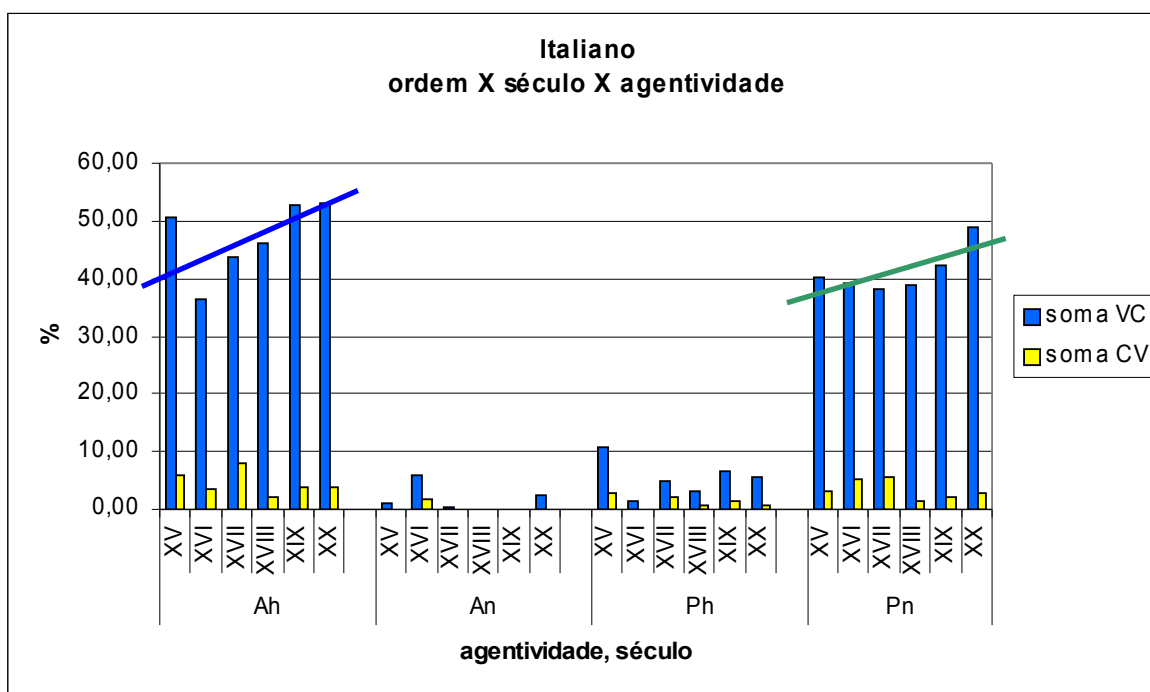


Gráfico I3 – médias das freqüências de dados segundo o traço *agentividade* e a pertinência a uma das classes gerais VC ou CV.

Por outro lado, a combinação dos traços *ordem* e *agentividade* permite discernir com mais propriedade os grupos de ocorrências mais importantes na constituição desses perfis. Tome-se, em primeiro lugar a ordem **a** (SVC), que mostra o mesmo perfil crescente, tanto para **Ah** quanto para **Pn**, exibido no Gráfico I3; isso significa que as ocorrências que contêm os traços **Ah** e **Pn** são mais freqüentes, em parte, graças a sua associação com a

ordem **a** (SVC). As curvas sigmóides que resolvem o comportamento dos dados dessas duas classes têm mínimos coincidentes, mas apresentam pequena defasagem nos máximos; em outras palavras, a combinação dos traços *agentividade* e *ordem SVC* aponta um avanço mais rápido de **Ah** que de **Pn**.

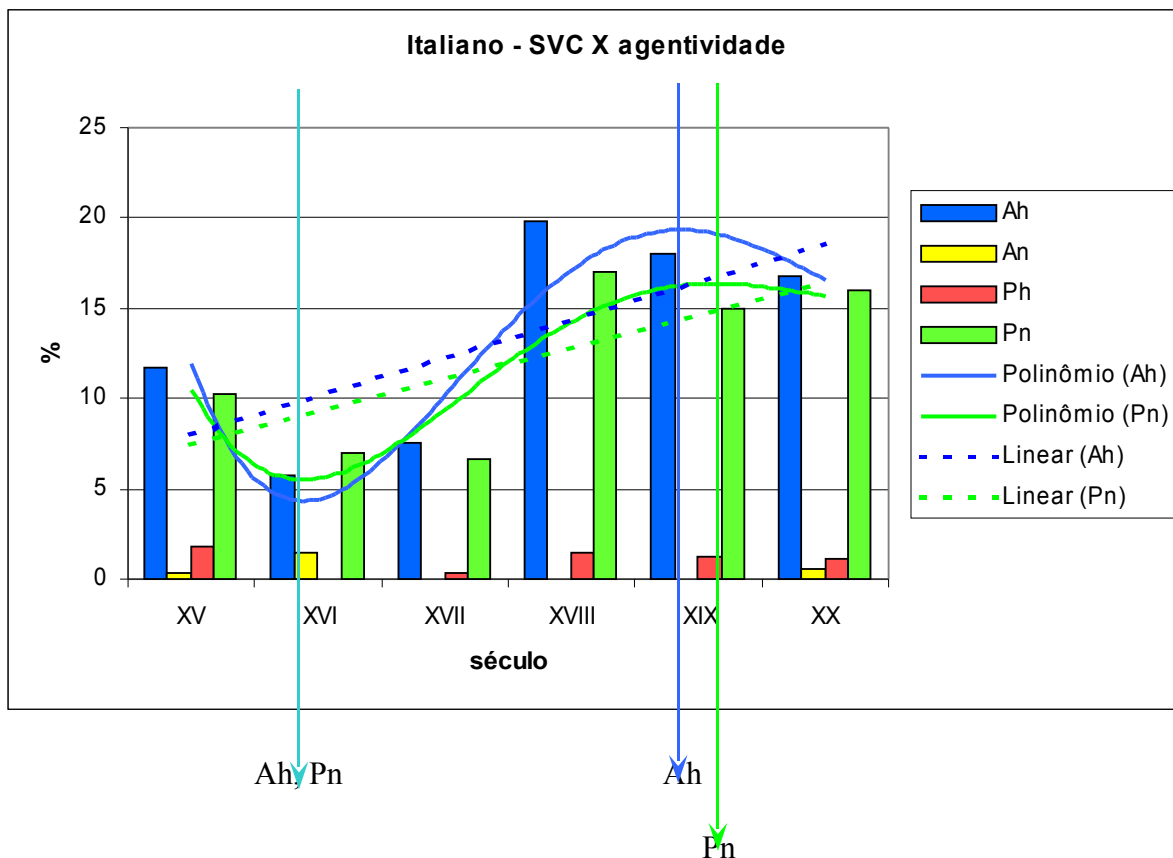


Gráfico I4 – médias de freqüências relativas para o cruzamento dos traços *ordem – SVC* e *agentividade*.

São exemplos de ocorrências associadas a esses traços:

<i>texto (século)</i>	<i>ocorrência</i>
Expositione (XV)	et desidero che lui augmenti iltuo stato in maggior gloria
LingToscana (XVI)	perche il nome contiene in cinque casi tutte le persone, e la seconda
Viaggi (XVII)	perche il Gran Signore non voleua altro General, che lui
Lettere (XVIII)	nelle quali uno scrittore latino... gratuitamente accorda la palma della eloquenza ai Greci, o a chi la vorrà;
Graal (XX)	Longino appoggia le dita insanguinate sugli occhi e riacquista

Tabela I2 – exemplos de ocorrências SVC colhidas no *corpus* italiano.

Para a ordem **i** (VC), observa-se uma flutuação de até cerca de 15% das médias de frequência de **Ah** e de **Pn**, de modo que os comportamentos globais se resolvem, nos dois casos, por retas subhorizontais. Em casos como os baixos índices relativos de Pn-XVIII, é preciso investigar se se trata do efeito provocado por um texto isolado, cujo perfil é muito distinto dos demais; alternativamente, a amostragem de um número maior de ocorrências colhidas em textos diferentes dos utilizados neste estudo diluiria esse efeito individual. Veja-se, todavia, que, à exceção do século XVI (praticamente coincidindo, portanto, com a mudança de AN para NA nesse sistema), as proporções de **Ah** são sempre mais altas que as de **Pn**. As curvas têm mínimos mais espaçados e perfis diferentes, indicando maneiras distintas de implementação de mudanças nessas duas categorias mais comuns.

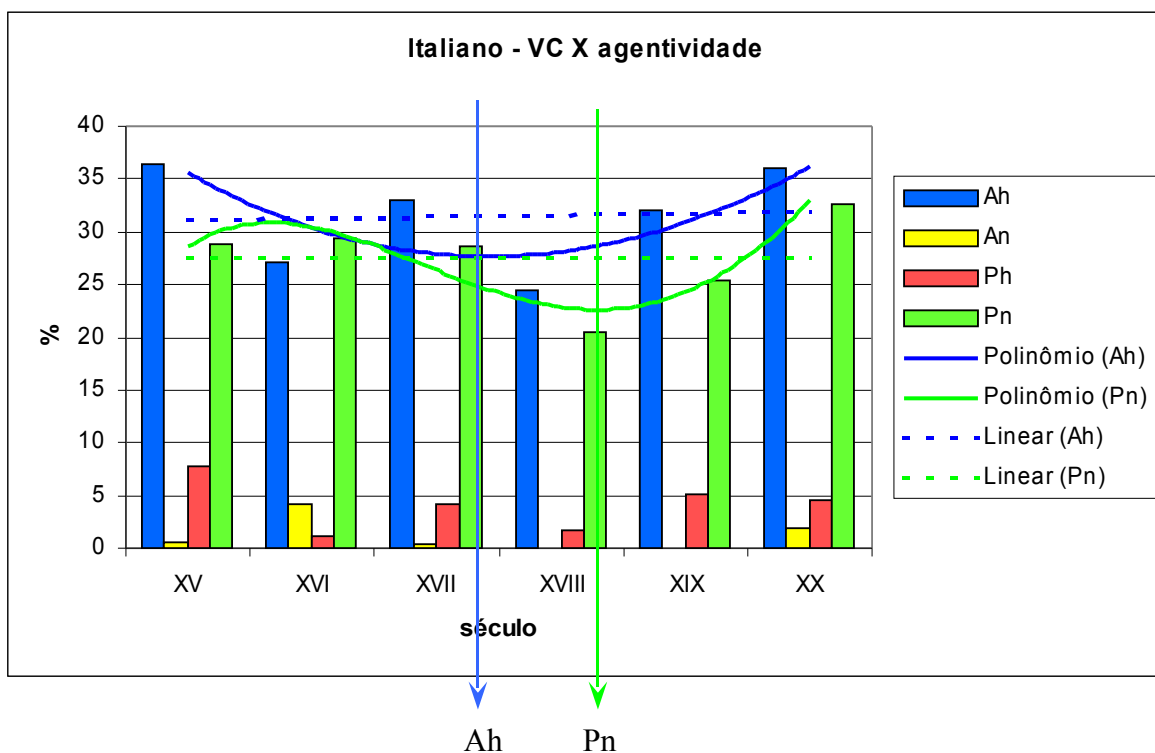


Gráfico I5 – frequências relativas para o cruzamento dos traços *ordem – VC* e *agentividade*.

São exemplos de ocorrências associadas a esses traços:

Expositione (XV)	che amaua la uergine Maria come contento dogni sua salute
Compendio (XVI)	scaturiscono ne'vasi, nō hauendo veloce esito; restano
Lettere (XVIII)	e non avreste così dovuto arrossire in riceverlo."
Geografia (XIX)	per abbreviare l'agonia degl'infermi accelerandone la morte!

Tabela I3 – exemplos de ocorrências VC no *corpus* italiano.

Para a ordem *g* (SV), em que a maior parte das ocorrências também se concentra nas classes Ah e Pn, os perfis indicam leve aumento (1% < variação < 5%) para o traço humano do agente e uma tendência subhorizontal (variação de até 2%) para o traço não-humano do paciente. As iRRLs dessas classes são, respectivamente, .54 e .32; entretanto, as curvas mostram que os dois subconjuntos de dados têm perfis distintos.

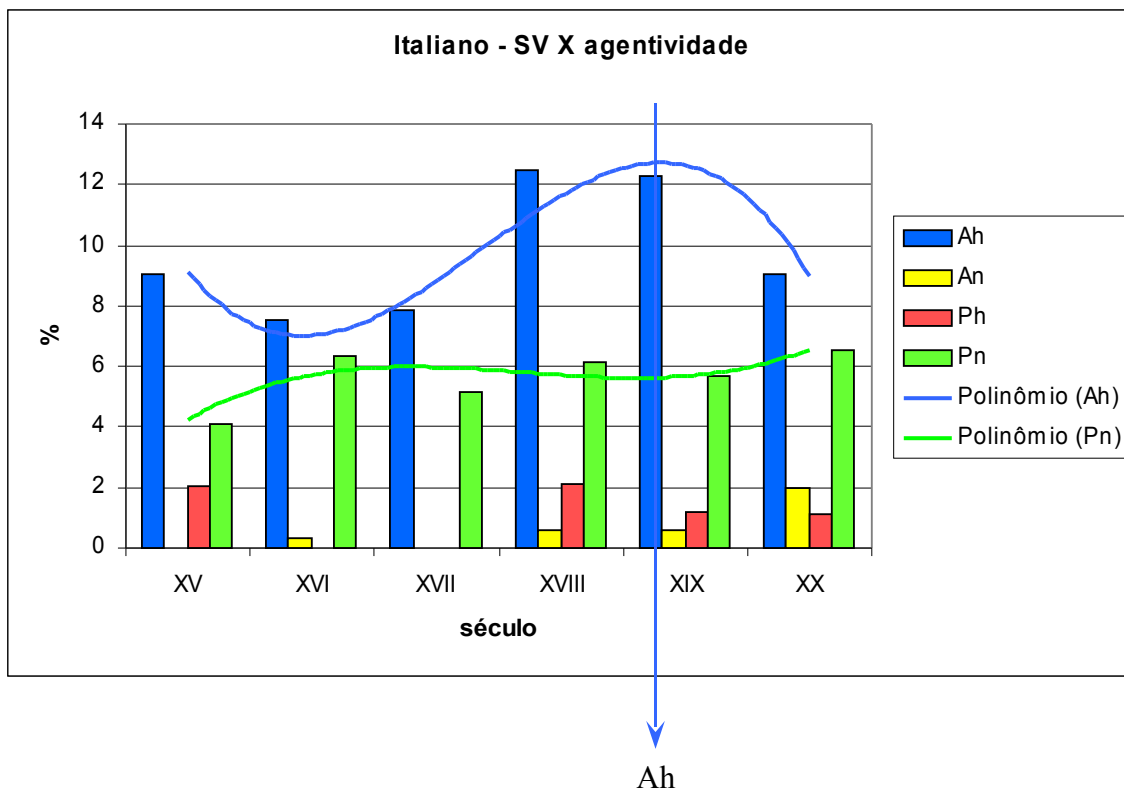


Gráfico I6 – médias de freqüências relativas para o cruzamento dos traços *ordem* – SV e *agentividade*.

São exemplos de ocorrências associadas a esses traços:

Dyalogo (XV)	almeno per alquante herbe odorifere potrebbe bastare per horto
Compendio (XVI)	Imperò questa cognitione, & consiglio debbe essere nel saggio , & sperimentato giudicio
Christoforo (XVII)	E Colombo nella morte sua fece come buon patriotta, perche lasciò
Principj (XVIII)	di cotal sua posizione . Esso certamente conviene in ciò , ch'è vero
Geografia (XIX)	ed un altro Titi appare più tardi sotto Ramses III
Graal (XX)	La metamorfosi finale di quest'immagine appare nell'arte romanica

Tabela I4 – exemplos de ocorrências SV encontradas no *corpus* italiano.

As ocorrências de **Ah** e de **Pn** são, também para a ordem **h** (VS), mais frequentes que **An** e **Ph**; a despeito da flutuação nas proporções gerais dessas classes, a RRL de **Ah**/VS mostra tendência subhorizontal levemente decrescente, enquanto a de **Pn**/VS é horizontal. Deve-se considerar o fato de um dos textos do século XVII (*Viaggi del pellegrino*) não se alinhar com os que lhe são coetâneos em relação ao traço Ah/VS – efeito que pode ser minimizado através da ampliação da base de dados.

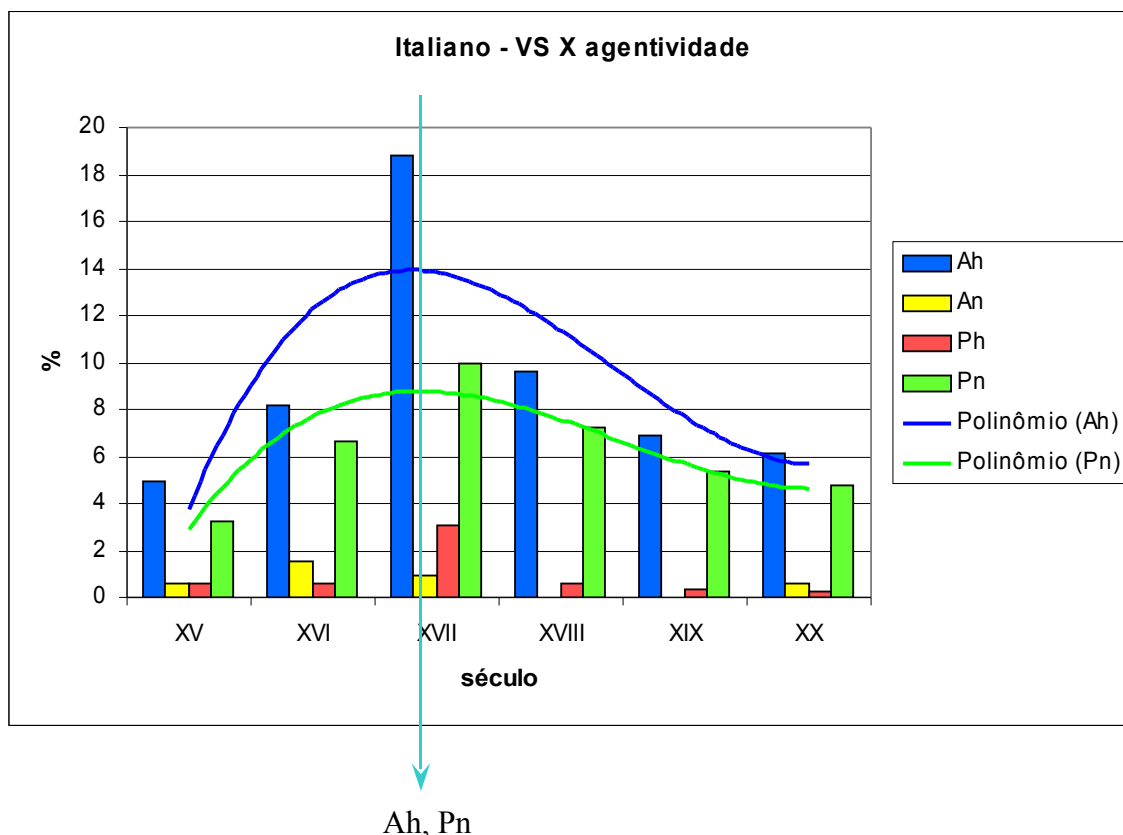


Gráfico I7 – médias de freqüências relativas para o cruzamento dos traços *ordem – VS* e *agentividade*.

São exemplos de ocorrências associadas a esses traços:

Expositione (XV)	Come fece il nostro Salvatore a sancto Pietro: il quale hauendo
Christoforo (XVII)	non prima (darei io) per non auuilire più di quello
Lettere (XVIII)	Alle prove. Discende Enea nell' inferno e gli vien fatta
Geografia (XIX)	ma, come argutamente osservò il De Rougé, nei testi egiziani
Eco (XX)	Penso che l'errore sia di credere che prima venga l'eresia, poi

Tabela I5 – exemplos de ocorrências VS encontradas no *corpus* italiano.

A tabela I6 sumariza os comportamentos das classes **Ah** e **Pn** segundo as ordens mais freqüentes. Observe-se que as ordens mais favoráveis para o aumento da proporção relativa de **Ah** e de **Pn** são as que têm sujeito pré-verbal, SVC (**a**) e SV (**g**), ao passo que, nas demais, há manutenção do número relativo de ocorrências, com indicativos sutis de aumento (SV/**Pn**, p.e.) ou de diminuição (CV/**Ah**, **Pn**; VS/**Ah**) dessas proporções. Esse resultados devem, naturalmente, ser analisados diante dos Gráficos I2 e I3, que apresentam as proporções absolutas das classes sintáticas no *corpus*.

ordem	Ah	Pn
SVC	↗	↗
VC	→	→
CV	→	→
SV	↗	→
VS	→	→

Tabela I6 – comportamento das ordens mais freqüentes para os traços Ah e Pn.

Em relação aos *corpora* ibéricos, as classes dispostas na Tabela I6 mostram clara tendência à estabilidade, relacionada às RRLs de inclinação subhorizontal no intervalo de amostragem. Isso indica, para o cruzamento dos fatores *ordem* e *agentividade*, um caráter diferente do italiano diante do português e do espanhol, já que se verifica nestes últimos uma propensão à diminuição das proporções absolutas de várias categorias (cf. Tabelas P1 e E1 acima).

O fator *ambiente oracional* mostra uma divisão desigual das ocorrências no *corpus* italiano, como se pode notar ao exame do gráfico abaixo, a qual corrobora os resultados de Bybee 1991 no que tange ao comportamento diferenciado de classes sintáticas distintas. É interessante observar que as orações subordinadas (3) predominam em todo o intervalo de amostragem, mas são resolvidas por RRL e curva polinomial que indicam um decréscimo dessas freqüências absolutas ($iRRL(1) = 2,66$; $iRRL(3) = -2,62$); essa diminuição é concomitante ao aumento nas freqüências de orações principais (1), indicando uma tendência à redução de estruturas frasais complexas. As proporções relativas de orações coordenadas (2) são praticamente constantes no *corpus* analisado, e apresentam variação da ordem de 5%.

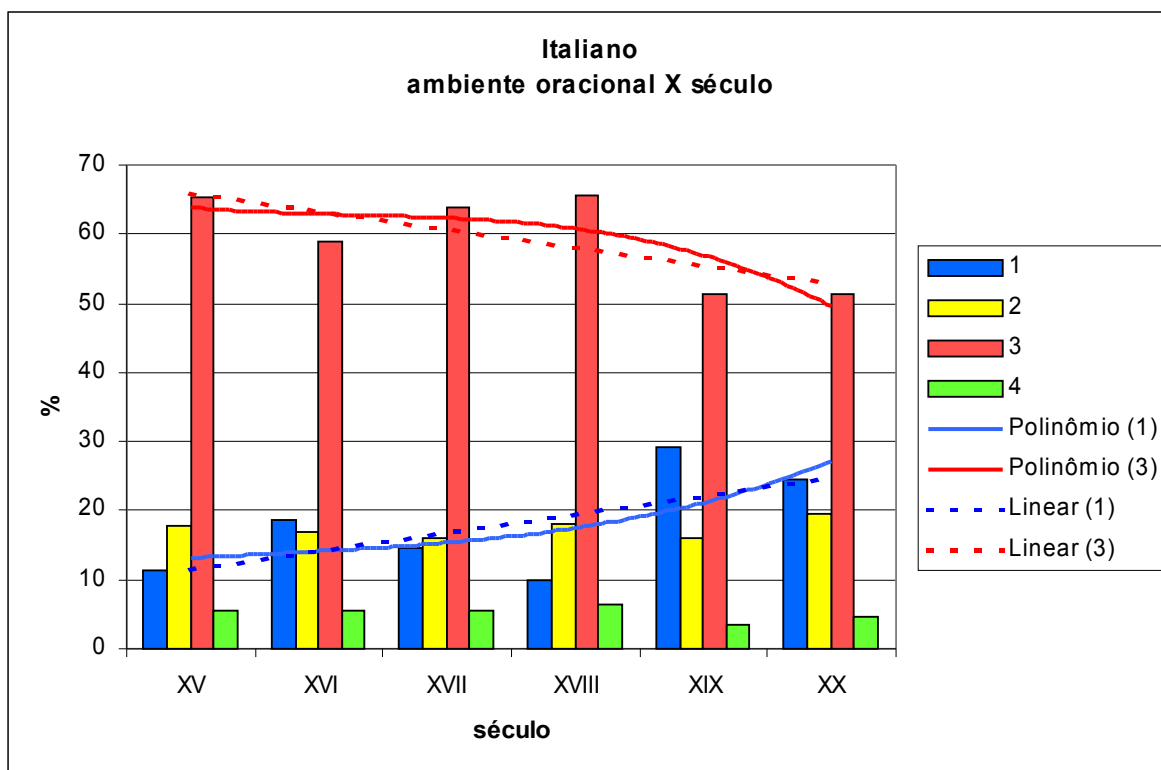


Gráfico 18 – freqüências absolutas das classes de *ambiente oracional* no *corpus* italiano.

O cruzamento dos fatores ordem e ambiente oracional para o *corpus* italiano fornece resultados semelhantes aos encontrados para os dados ibéricos. Para as orações principais, em primeiro lugar, observa-se que os conjuntos **a** (SVC) e [VC] apresentam proporções relativas crescentes em frequência (inclinações de RRL de 1,99 e 2,23, respectivamente), da mesma forma que se passa com as de **g** (SV) e **h** (VS). A exemplo do observado para as orações principais do português (Gráf. P6a acima), a classe **a** (SVC) é a mais freqüente no grupo VC, seguida da classe **i**. Por fim, nota-se que as médias gerais de frequência absoluta de orações principais para os três *corpora* são de mesma ordem, girando em torno dos 10%.

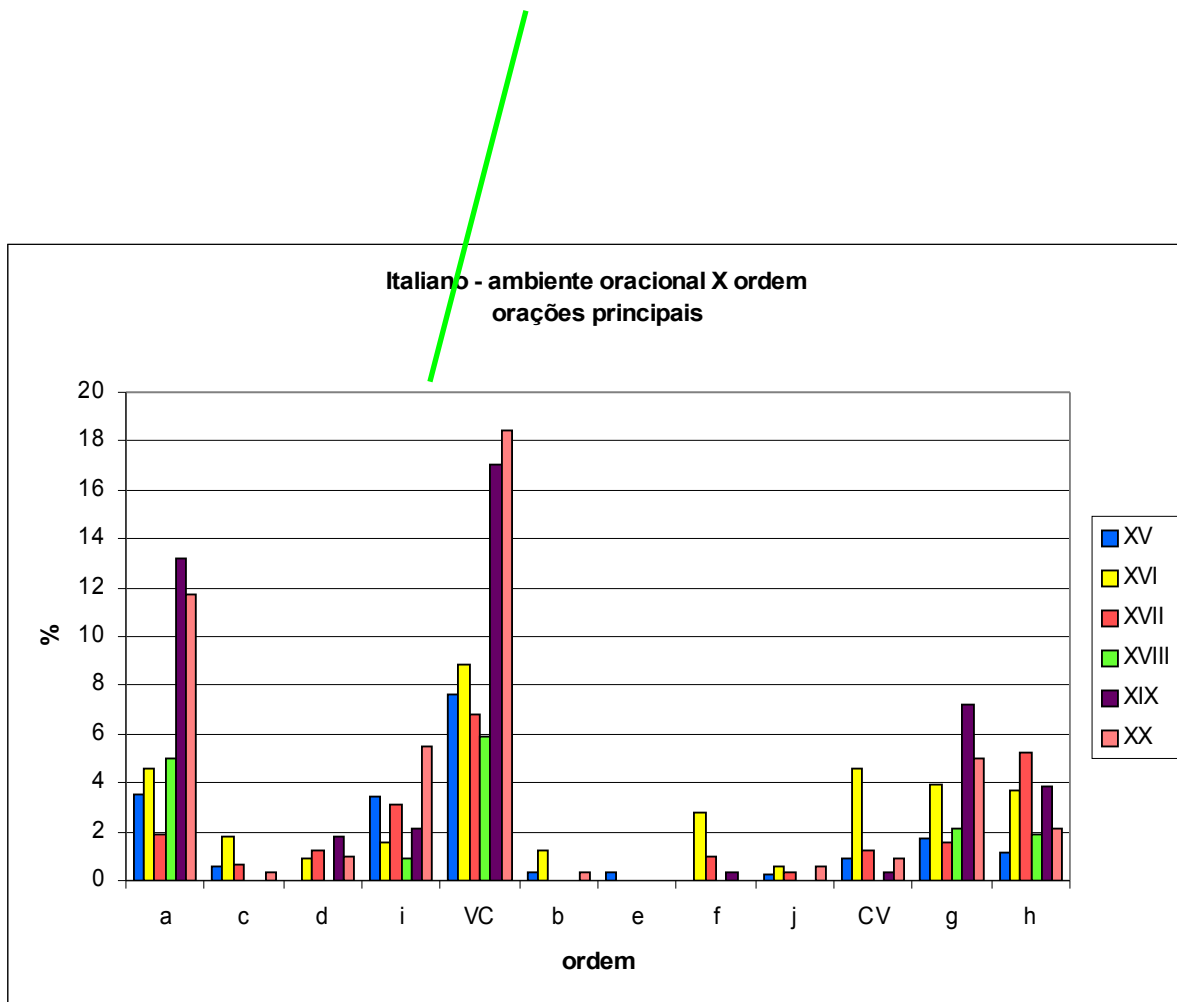


Gráfico I9 – proporções absolutas de orações principais no *corpus* italiano.

São exemplos de ocorrências associadas a esse traço:

Compendio (XVI)	Scaturisce detta acqua minerale, particolarmente in 7.luoghi
Principj (XVIII)	Tanto bene tutti i Giureconsulti... hanno considerato la Legge Publilia , con la quale
Graal (XX)	Una serie di tavolette d'avorio adotta questa iconografia che si potrebbe

Tabela I7 – exemplos de ocorrências com orações principais.

Para as orações coordenadas, verifica-se um comportamento semelhante ao identificado para os dados de português (Gráf. P6b acima), em que há predominância da classe **i** (VC) no grupo [VC] e tendência deste ao aumento. A inclinação das retas de regressão linear desses conjuntos [VC], entretanto, é maior no português (0,89) do que no italiano (0,57). Para o espanhol (Gráf. E9 acima), as coordenadas diminuem em frequência.

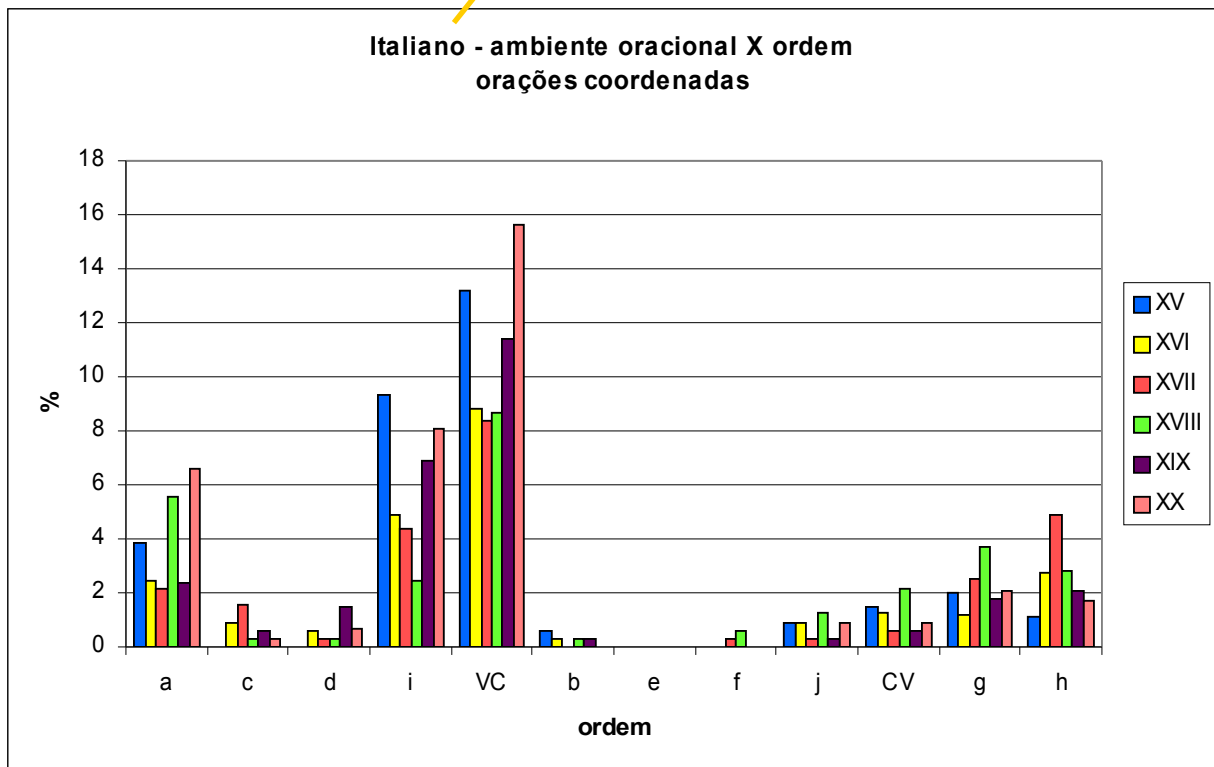


Gráfico I10 – proporções absolutas de orações coordenadas no *corpus* italiano.

São exemplos de ocorrências associadas a esse traço:

Dyologo (XV)	Ma grande merauglia adiuenne:che vicendo lanima del venerabile Anastagio
Viaggi (XVII)	ma portauano ben gli archi loro, e le frecchie, & erano bizarrissimi a vedere
Graal (XX)	ma... troviamo a sinistra Ecclesia con in mano la coppa, e a destra Sinagoge

Tabela I8 – exemplos de ocorrências com orações coordenadas.

As orações subordinadas no *corpus* italiano também apresentam comportamento similar ao observado para esse traço sintático no português (cf. Gráf. P6c acima): há tendência à diminuição nas proporções relativas de construções hipotáticas ao longo da amostra. Algumas das classes dominantes, como **i** (VC), **g** (SV) e **h** (VS), exibem perfis decrescentes, a exemplo do que se dá com o grupo [VC]; outras, como **a** (SVC), mostram flutuação estável; as demais classes exibem freqüências relativas sensivelmente baixas, mas, de maneira geral, apresentam discreta diminuição de freqüência.

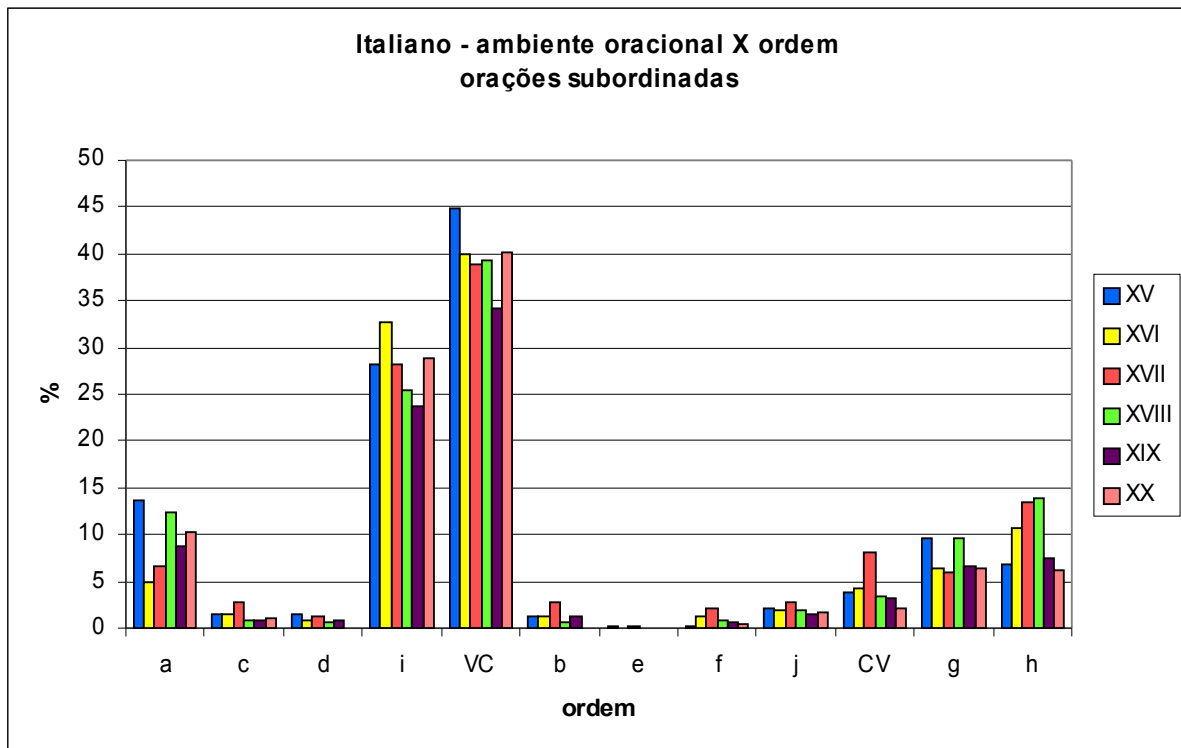


Gráfico I11 – proporções absolutas de orações subordinadas no *corpus* italiano.

São exemplos de ocorrências associadas a esse traço:

Compendio (XVI)	nel tēpo che l'Illustriss.Marchese di Mantoua vi andò, & se ne parti sano
Campanella (XIX)	scritti più tardi da' sacerdoti che avevano assistito alcuni di quegl' infelici, inserti poi
Eco (XX)	avevano costituito una chiesa in cui si distinguevano i perfetti dai semplici credenti

Tabela I9 – exemplos de ocorrências com orações subordinadas.

Mesmo que o perfil para a superclasse [VC] seja decrescente, nota-se que a reta de regressão linear para os dados de italiano tem uma inclinação maior (-1,16) do que as referentes aos dados de português (-0,47) e de espanhol (-0,37); em outras palavras, a taxa de decréscimo da frequência das orações subordinadas é mais alta em italiano do que no grupo ibérico.

As orações subordinadas em coordenação apresentam, em italiano, um perfil cuja RRL tem inclinação muito baixa. A ordem mais frequente para essas orações é *i* (VC); as demais ordens têm frequências demasiado baixas para permitir inferências seguras sobre o comportamento das ocorrências a que se referem.

São exemplos de ocorrências associadas a esse traço:

<i>texto (século)</i>	<i>ocorrência</i>
Dyalogo (XV)	che quello pocho di spacio che la voce tacete e poi lo chiamoe:significa che egli
LingToscana (XVI)	al meno secondariamente ad esser segno, ed accennar quella cosa , che sarebbe accennata
Viaggi (XVII)	più tosto da seruigio per maneggiar terra, tagliar legna, e cose simili
Principj (XVIII)	era di starsi tutti soli nelle loro grotte , e curare le loro famiglie
Geografia (XIX)	e che questi Libii biondi sieno Celti, venuti dalla Spagna
Eco (XX)	più o meno perfette, e pensano che questa sia quella e quella questa."

Tabela I10 – exemplos de ocorrências com orações coordenadas em subordinação.

De maneira geral, o comportamento das ocorrências do *corpus* italiano é paralelo, no que tange ao *ambiente oracional*, ao observado para o português e para o espanhol: as frequências absolutas de orações principais e de coordenadas aumentam na mesma medida em que se verifica a diminuição das proporções de orações subordinadas. Isso indica uma alteração no padrão sintático básico do período composto, que tenderia a uma estrutura

menos complexa e orientada, primariamente, pelas orações independentes, e secundariamente pelas coordenadas.

<i>ambiente oracional</i>	<i>VC</i>
1 (principais)	↗
2 (coordenadas)	↗
3 (subordinadas)	↘
4 (coordenadas-subordinadas)	→

Tabela I11 – comportamento da ordem VC
segundo o ambiente oracional.

Fixando-se o fator *ambiente oracional*, observa-se o comportamento das classes de agentividade propostas para análise. Para as orações principais, nota-se que as curvas mais expressivas, **Ah** e **Pn**, têm perfis ascendentes, indicando uma tendência de aumento das frequências absolutas dessas classes. Os mínimos para essas curvas são ligeiramente defasados (diferença de pouco mais de 1/3 de século), e, como ambas têm concavidade para cima em todo o intervalo considerado, os pontos de máximo ainda não foram atingidos. Observe-se também o ponto excêntrico para as duas classes no século XVIII, que podem indicar a necessidade de se considerarem mais textos a fim de se diluírem os efeitos de uma única obra no cálculo das médias de frequência. Tomem-se os exemplos:

Ah – o archangelo Gabriel nõ disse maria nominandola per - ExposRev XV

An – Parte de' Pronomi adunque anno la persona distinta; - LingTosc XVI

Ph – aveva erroneamente identificato i Mas'uas'a coi Moschos - GeogrEtn XIX

Pn – Il prologo di Perlesvaus riassume con chiarezza i momenti cruciali - Graal XX

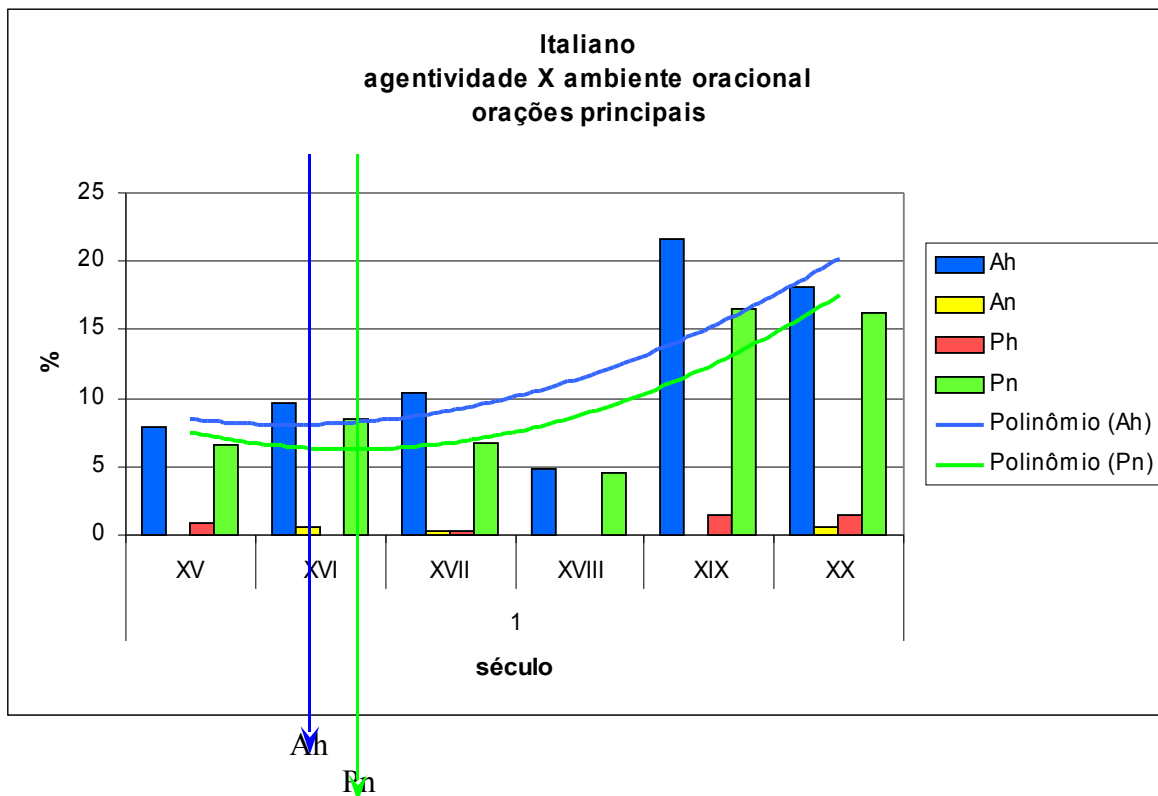


Gráfico I12 – médias de frequências absolutas de ocorrência das classes de *agentividade* para orações principais.

Já o comportamento das classes de *agentividade* em orações coordenadas é diferente do observado no gráfico anterior. A escala dentro da qual a variação das proporções absolutas de coordenadas varia ao longo do tempo é reduzida em relação à das orações principais, e as tendências indicam uma estabilidade muito grande das frequências de coordenadas no intervalo de amostragem. É interessante, contudo, notar que as curvas para as classes mais comuns, **Ah** e **Pn**, apresentam tendências complementares: um aumento leve das proporções de **Pn** é acompanhado de uma diminuição de mesma ordem de **Ah**, e vice-versa. Tomem-se as ocorrências reproduzidas abaixo como exemplos:

Ah – ma gli suoi costumi studiaua sempre di comportare con grande - Dyalogo XV

An – variato dal singulare, altri non l'anno da quel distinto – LingTosc XVI

Ph – Et. doue lascio io il Mag. M.Aluigi Vitturi nobile Venetiano - Regole XVI

Pn – E perciò egli ha fatto il suo libro assai minore di quello che - PrincLett XVIII

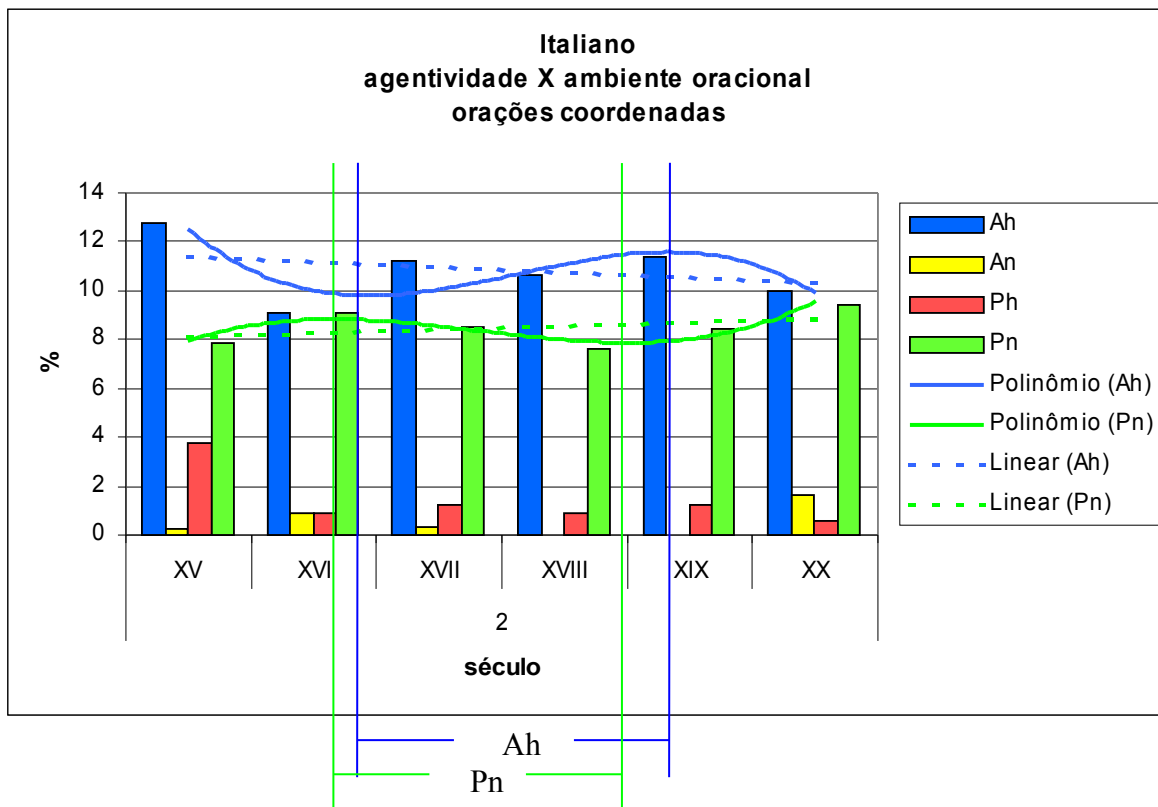


Gráfico I13 – médias de freqüências absolutas de ocorrência das classes de *agentividade* para orações coordenadas.

O traço *agentividade*, para as orações subordinadas, apresenta praticamente o mesmo perfil das coordenadas, mas a escala de variação das freqüências absolutas no Gráfico I14 é maior. Nota-se, entretanto, que as RRLs são subhorizontais, indicando uma flutuação estável no intervalo considerado. Compete ressaltar ainda que, ao contrário do que se observa para as coordenadas, as subordinadas mostram uma liderança dos pontos críticos da curva **Ah**, e não de **Pn**. Exemplos de ocorrências nesse ambiente:

Ah – Vnaltra volta lauado lo p/dicto Nonnosio le sue lampade di vetro - Dyalogo XV

An – saluo che il Pronome ha distinte le persone per varietà di caratteri - LingTosc XVI

Ph – di mandar questo esercito contro'l Rè di Persia - ViagPell XVII

Pn – e numerose, che mettevano paura a' Padri - Vico XVIII

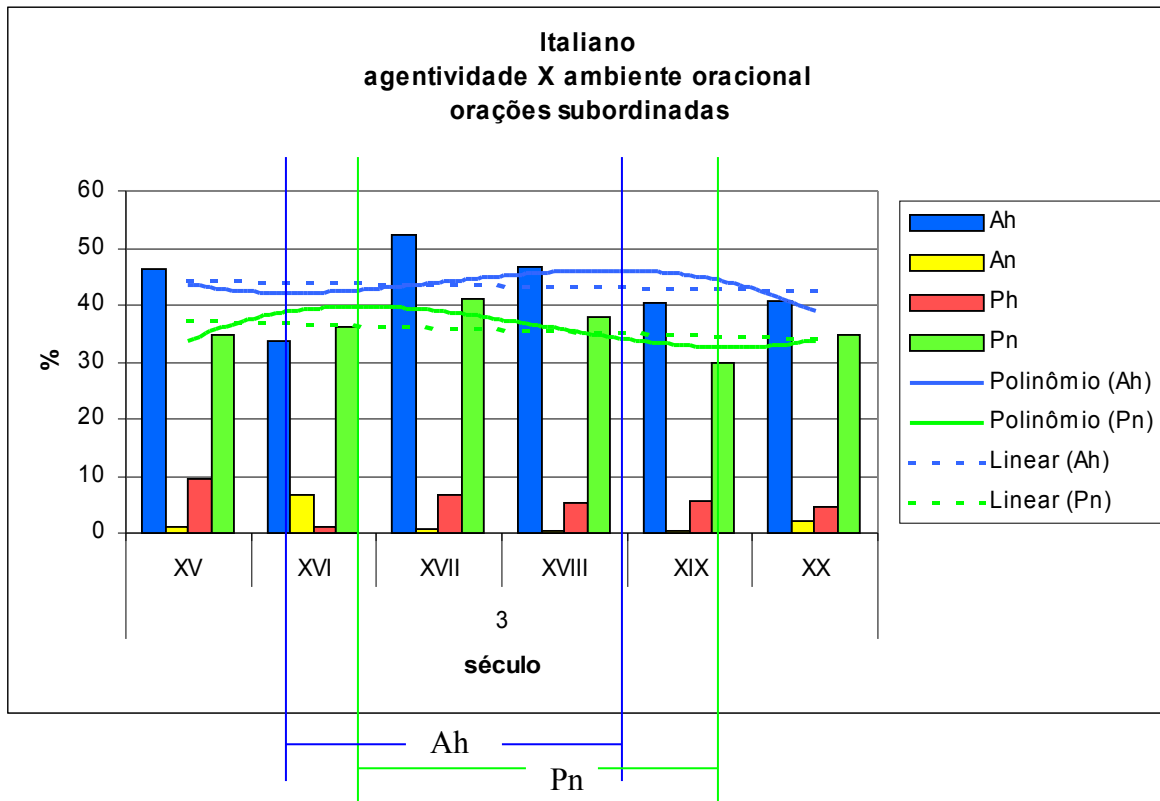


Gráfico I14 – médias de freqüências absolutas de ocorrência das classes de *agentividade* para orações subordinadas.

Isolando-se, por fim, o traço *agentividade*, podem-se comparar os comportamentos das classes mais freqüentes, **Ah** e **Pn**, em ambientes oracionais distintos. Observe-se, inicialmente, que o perfil da curva Ah1 é idêntico ao obtido para a distribuição de orações principais (q.v. Gráfico I12), indicando que, nesse caso, o ambiente oracional é determinante para a evolução de **Ah**. O traçado da curva **Ah3** é distinto de **Ah1** por apresentar um mínimo local no século XV; assemelha-se, portanto, à curva **Ah** no Gráfico I14.

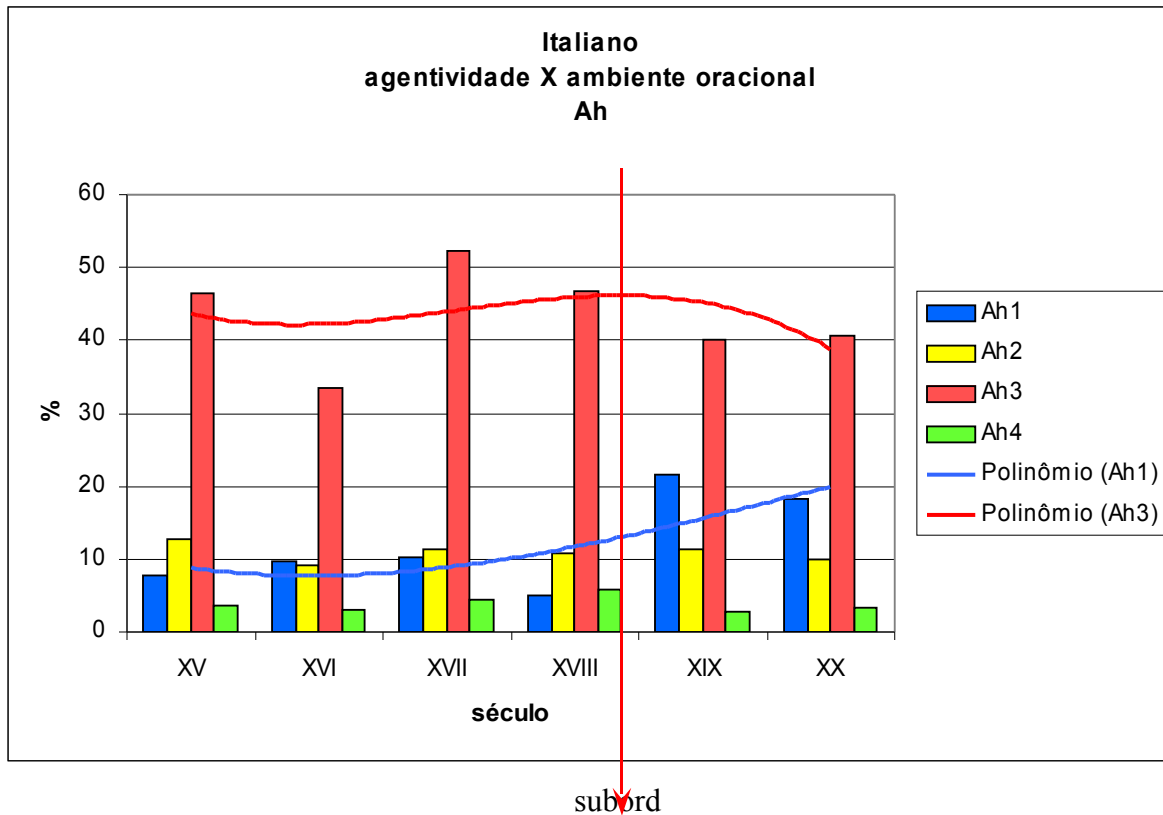


Gráfico I15 – médias de frequências absolutas de ocorrência das classes de *ambiente oracional* para Ah.

As curvas para **Pn** também remetem aos Gráficos I12 e I14, numa confirmação da relevância do ambiente oracional para a definição dos padrões de comportamento da amostra. Assim é que a curva **Pn1** tem um perfil ascendente característico de orações principais, enquanto **Pn3** tem o mesmo traçado praticamente senoidal já identificado acima. Como as configurações desses perfis são diferentes, é difícil argumentar em favor da influência de um ambiente oracional sobre outro; isso não impede que haja uma relação entre os mesmos no sentido de o aumento em orações principais ser condicionado, por exemplo, a uma certa proporção de orações subordinadas, isoladamente ou em combinação

com outros fatores: veja-se, a esse propósito, que o mínimo da curva **Pn1** coincide com o máximo de **Pn3**. A despeito da oscilação das freqüências de orações subordinadas, a ascensão das freqüências de principais, uma vez acionada, não se deteve até o final do intervalo considerado neste estudo.

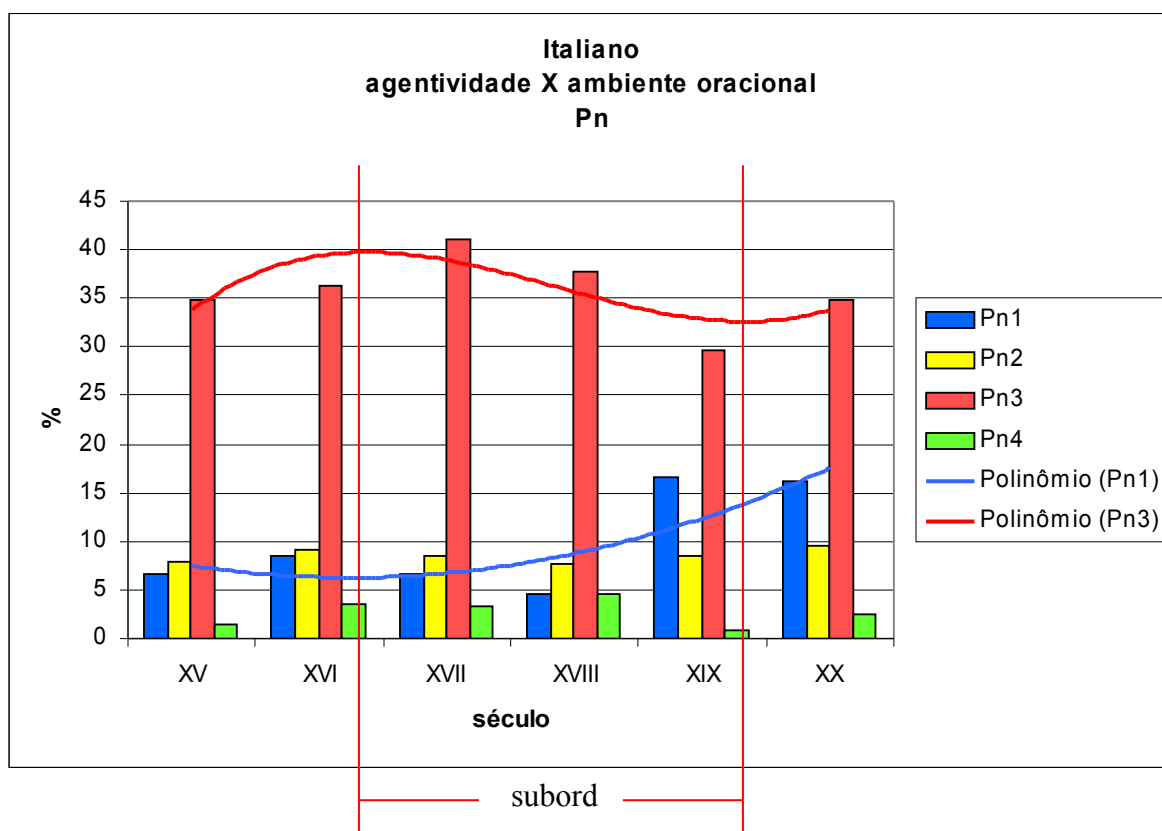


Gráfico I16 – médias de freqüências absolutas de ocorrência das classes de *ambiente oracional* para Ah.

A normalização das freqüências absolutas das classes ilustradas no Gráfico I15 permite verificar mais de perto as relações topográficas das curvas que representam o comportamento da amostra. É possível observar, portanto, que a diminuição das freqüências de subordinadas (**Ah3**) é seguida de igual efeito em coordenadas (**Ah2**) e, no século seguinte, em orações principais (**Ah1**); por outro lado, a partir destas, a recuperação das freqüências é invertida, pois tem início pelas coordenadas e só então passa às subordinadas. Em outras palavras, a flutuação na freqüência do traço **Ah** é liderada, no período estudado, pelas orações subordinadas, quando há decréscimo, e pelas coordenadas, quando há aumento das freqüências.

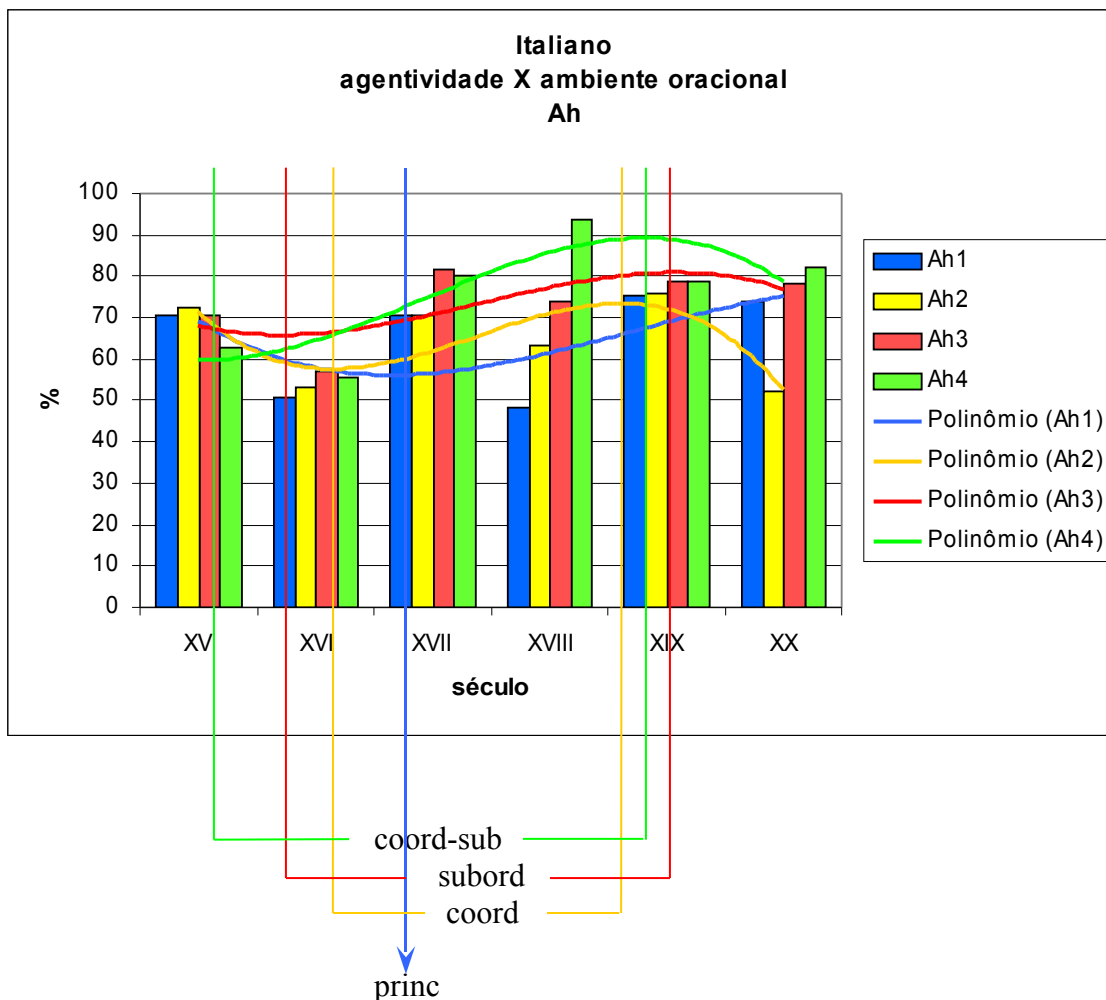


Gráfico I17 – médias de freqüências absolutas de ocorrência das classes de *ambiente oracional* para Ah.

Normalizando-se também as curvas de I16, verificam-se relações semelhantes para Pn: o ponto crítico da curva **Pn1** está situado entre os das curvas **Pn2** e **Pn3**. Observe-se, contudo, que os máximos locais desses dois polinômios são praticamente coincidentes no início do século XVI, ao passo que há um espaçamento maior entre os mínimos locais no século XIX. Interessante também é o fato de que, para **Ph** e diferentemente do que se viu para **Ah**, as subordinadas precedem as coordenadas nesses pontos críticos.

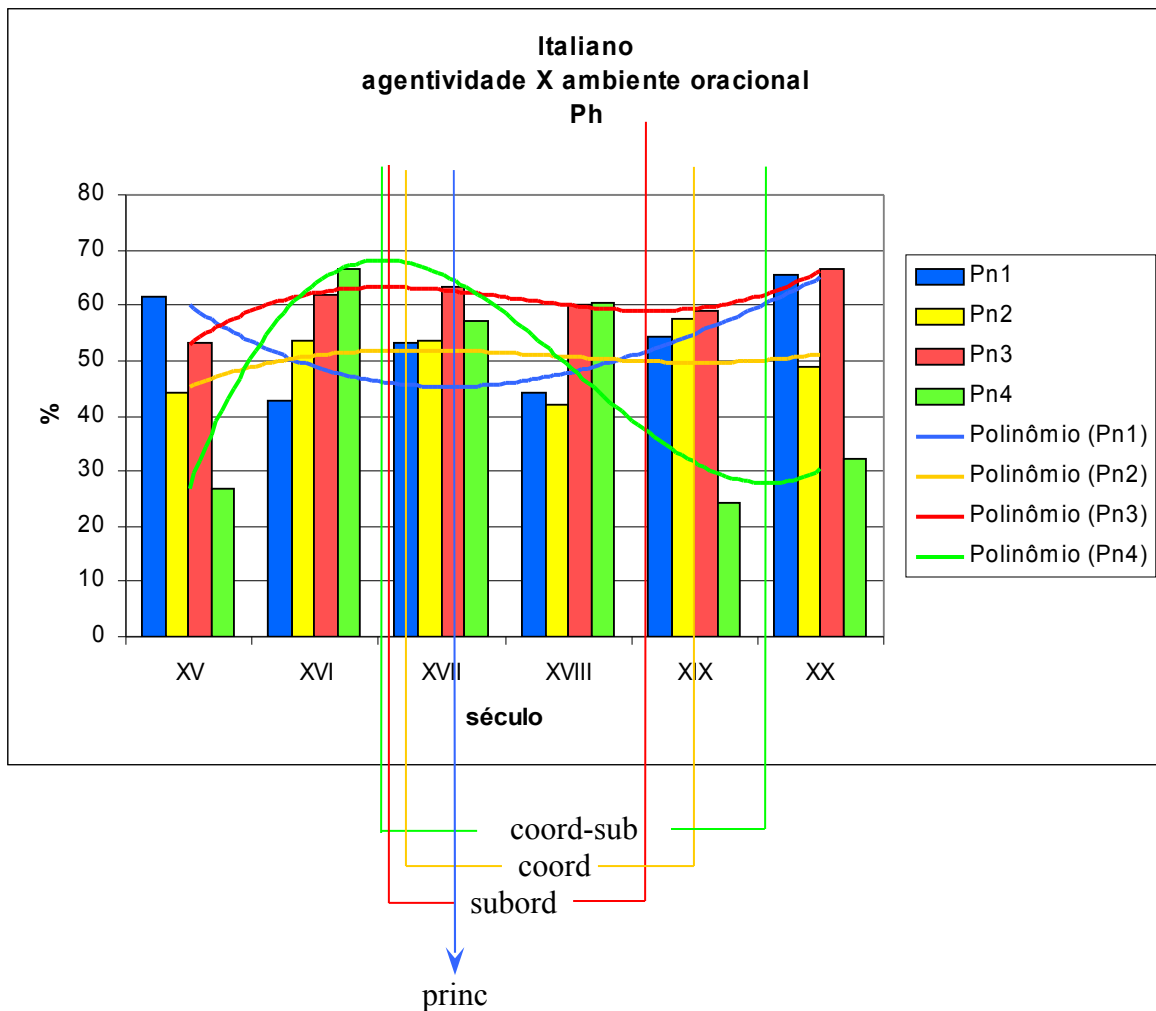


Gráfico I18 – médias de freqüências absolutas de ocorrência das classes de *ambiente oracional* para Ah.

A classe **k** (sujeito pleno) do fator peso de S mostra uma nítida especialização em termos de ordem de constituintes: expressa-se preferencialmente nas categorias **a** (SVC), **g** (SV) e **h** (VS), aumentando nas que têm sujeito pré-verbal e diminuindo na que o apresenta em posição pós-verbal. As inclinações das RRLs (iRRL) para essas ordens são, respectivamente, 4,94, 1,44 e -16. Comparem-se esses resultados com os de português, que mostra predominância das mesmas categorias sintáticas e aumento das proporções relativas de todas as três (apesar de as freqüências relativas das ordens de complemento não-expresso em português mudarem mais rapidamente do que em italiano).

São exemplos de ocorrências associadas a esse traço:

Expositione (XV)	Et similmente usa lachiesa sancta dimutare ilnome ad coloro che sono facti papa
Compendio (XVI)	I pesci paludosi,& che hãno sapore di fango... in breue spacio diuentano saporiti,& buoni
Campanella (XIX)	Erano allora in gran voga, e giustamente, i Padri Ministri degli infermi: lo stesso venerabile Camillo
Graal (XX)	Quando però, l'indomani, i Giudei aprirono i sigilli con l'intenzione di uccidere

Tabela I12 – exemplos de ocorrências de sujeito pleno.

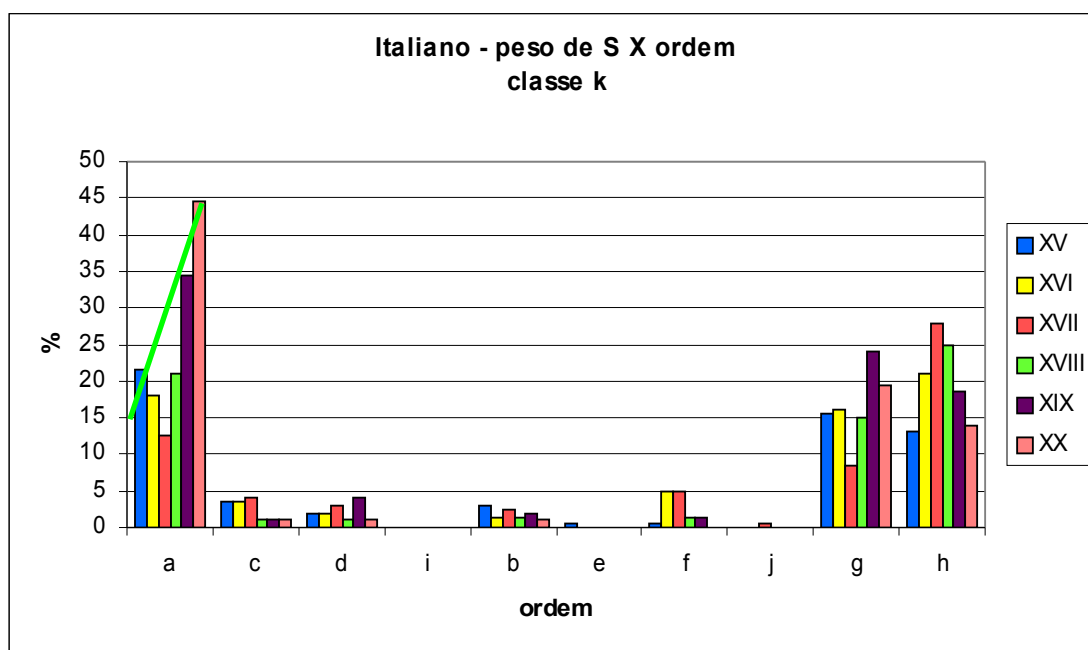


Gráfico I19a – proporções relativas de sujeito pleno no *corpus* italiano.

A classe **I** (sujeito pronominal) tem expressividade menor do que **k** (sujeito pleno), mas apresenta o mesmo perfil básico desta: as ordens mais proeminentes são **a** (SVC), **g** (SV) e **h** (VS) – todas com tendências decrescentes¹⁵ –, e as demais ocorrem em frequências mais modestas ou praticamente nulas.

São exemplos de ocorrências associadas a esse traço:

Dyologo (XV)	Questo che io tho dicto certame[n]te se puo prouare:che certamente
--------------	--

¹⁵ Apesar da flutuação nos valores da ordem **h** (VS), a regressão linear para esse subconjunto de ocorrências retorna um valor negativo (-.67) para a reta que resolve a distribuição dos dados.

LingToscana (XVI)	Negano molti (come dicemmo a suo luogo) la persona nel nome, ma
Christoforo (XVII)	non prima (direi io) per non auuilire più di quello
Principj (XVIII)	Egli è fondato sopra quella Dignità ; che cangiandosi gli huomini , ritengono
Campanella (XIX)	ed egli intanto avrebbe letto questi esami per valersene a tempo opportuno. - Come ben s' intende,
Eco (XX)	"Tu hai chiara la visione della costituzione del popolo di Dio. Un grande gregge, pecore buone

Tabela I13 – exemplos de ocorrências de sujeito pronominal.

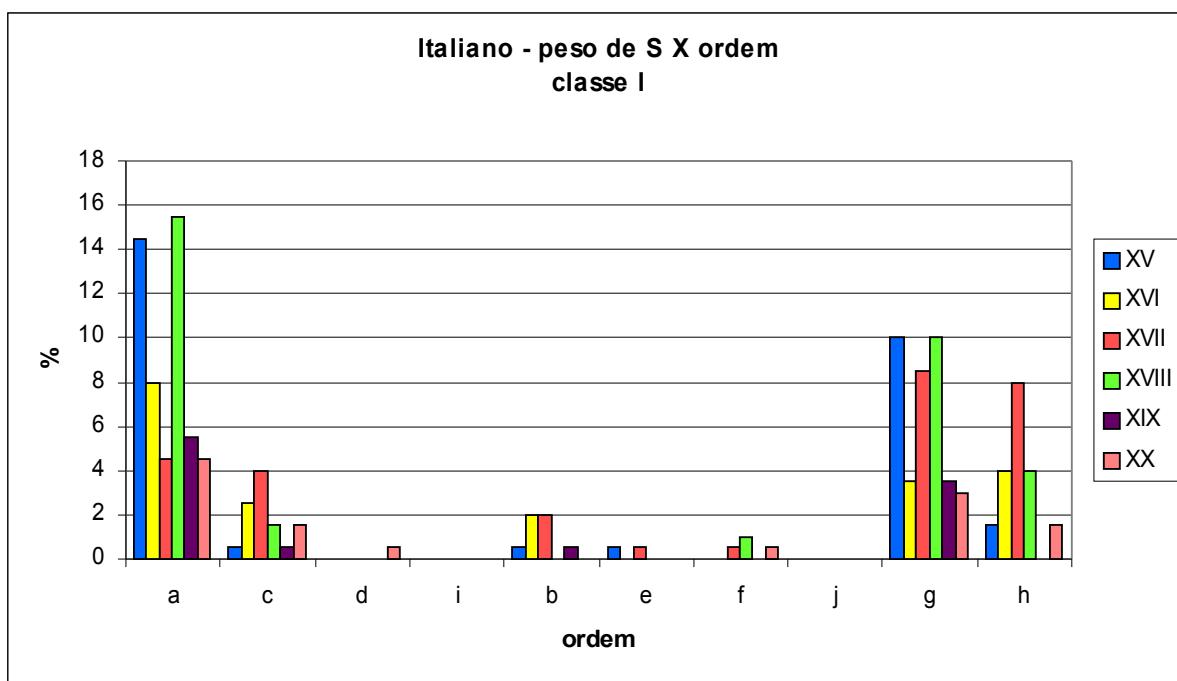


Gráfico I19b – proporções relativas de sujeito pronominal no *corpus* italiano.

O sujeito oracional ocorre em proporções baixas, mas parece tender a uma especialização nas ordens VCS (**d**) e VS (**h**) - i.e., com sujeito preferencialmente posposto ao verbo, independente da presença do complemento verbal. Entretanto, essa distribuição mostra tendências diferentes, na medida em que a ordem **d** (VCS) aumenta levemente em frequência (iRRL = .23), ao passo que **h** (VS) apresenta perfil subhorizontal (iRRL = -.02).

São exemplos de ocorrências associadas a esse traço:

Expositione (XV)	Dipoi piacque alla sancta chiesa agiugnere laltre parte
------------------	---

Compendio (XVI)	ch'è vno stupore, il vedere vn castello tãto ristretto si sito; dare tãta somma di vetronaglie,nõ
Viaggi (XVII)	è ben vero, che per la strada anderà sempre ingrossando
Lettere (XVIII)	Parmi innegabile , che Virgilio in questo luogo, e in mille altri simili , abbia voluto piacere ad Augusto
Geografia (XIX)	è certo che la razza antica occupante l'Africa settentrionale serviva di tipo agli
Graal (XX)	È noto che il Nuovo Testamento... si è formato e trasformato con il passare

Tabela I14 – exemplos de ocorrências de sujeito oracional.

A classe **n** (complemento pleno) tem peculiaridades interessantes em relação às demais. Observe-se, inicialmente, a concentração dessas ocorrências nas ordens **a** (SVC) e **i** (VC), ambas trazendo o complemento em posição pós-verbal. Em segundo lugar, os perfis dessas duas categorias sintáticas coincidem com os observados para o *corpus* espanhol, mas são distintos do português: a ordem **a** (SVC) aumenta em frequência relativa (iRRL = 3,16), enquanto **i** (VC) diminui (iRRL = -.54).

São exemplos de ocorrências associadas a esse traço:

Dyologo (XV)	che per sua oratione mosse vno monte da vno luo/cho ad vno altro
LingToscana (XVI)	considerando che elle non accennan'vn Nome, ma vn Pronome
Christoforo (XVII)	e qual Colomba altresì... annūciar buona nuoua dell'Euangelica pace
Principj (XVIII)	e le loro famiglie per le loro patrie , che lor conservavano le famiglie ; e per esso erano
Campanella (XIX)	Il Nunzio non tardò a trasmettere a Roma le pretensioni del Vicerè, tanto sul modo de formare il tribunale, quanto
Graal (XX)	e se i settatori di un movimento hanno commesso un delitto, questo delitto sarà attribuito a ciascun

Tabela I15 – exemplos de ocorrências de complemento verbal pleno.

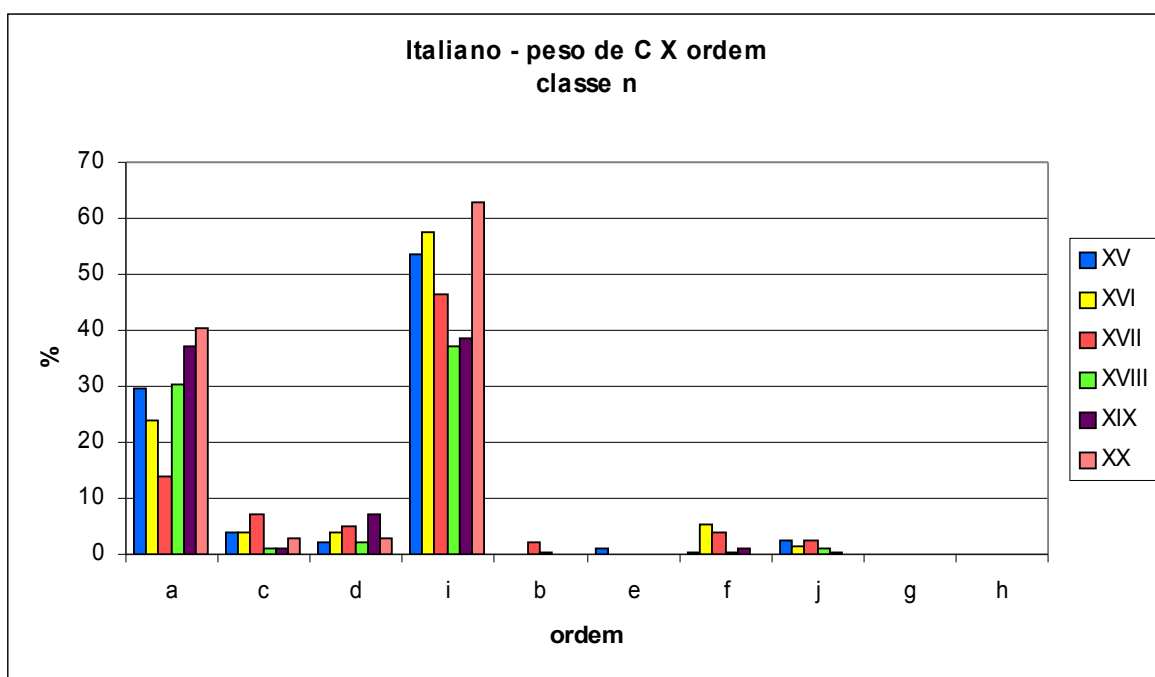


Gráfico I19c – proporções relativas de complemento verbal pleno no *corpus* italiano.

A distribuição da classe **o** (complemento pronominal), por sua vez, distingue o italiano dos conjuntos ibéricos: especializando-se em torno das ordens **i** (VC), **j** (CV) e **b** (SCV), o italiano mostra proporções mais expressivas (ainda que diminuam ao longo do intervalo de amostragem) da ordem **i** (VC) para objetos pronominais, enquanto o português e o espanhol concentram-se mais fortemente na ordem **j** (CV), que, nesses *corpora*, é decrescente.

São exemplos de ocorrências associadas a esse traço:

<i>texto (século)</i>	<i>ocorrência</i>
-----------------------	-------------------

Expositione (XV)	ma se uogliamo parlare di quello Re non lo nomineremo signore in Ferrara
LingToscana (XVI)	e altri simili, i quali dimostrano la cosa accennata : insegnandola quasi col dito .
Viaggi (XVII)	e per far, che si mandasse altri in luogo suo
Principj (XVIII)	Ma ci piace <i>ad exuberantiam</i> confutarlo dagl' impossibili , e dagli assurdi di cotal sua posizione .
Geografia (XIX)	prima di farlo sentiamo in dovere d'espore tutt'i dati di fatto
Graal (XX)	La regina di questo castello l'ha fatta incastonare in oro e pietre preziose

Tabela I16 – exemplos de ocorrências de complemento verbal pronominal.

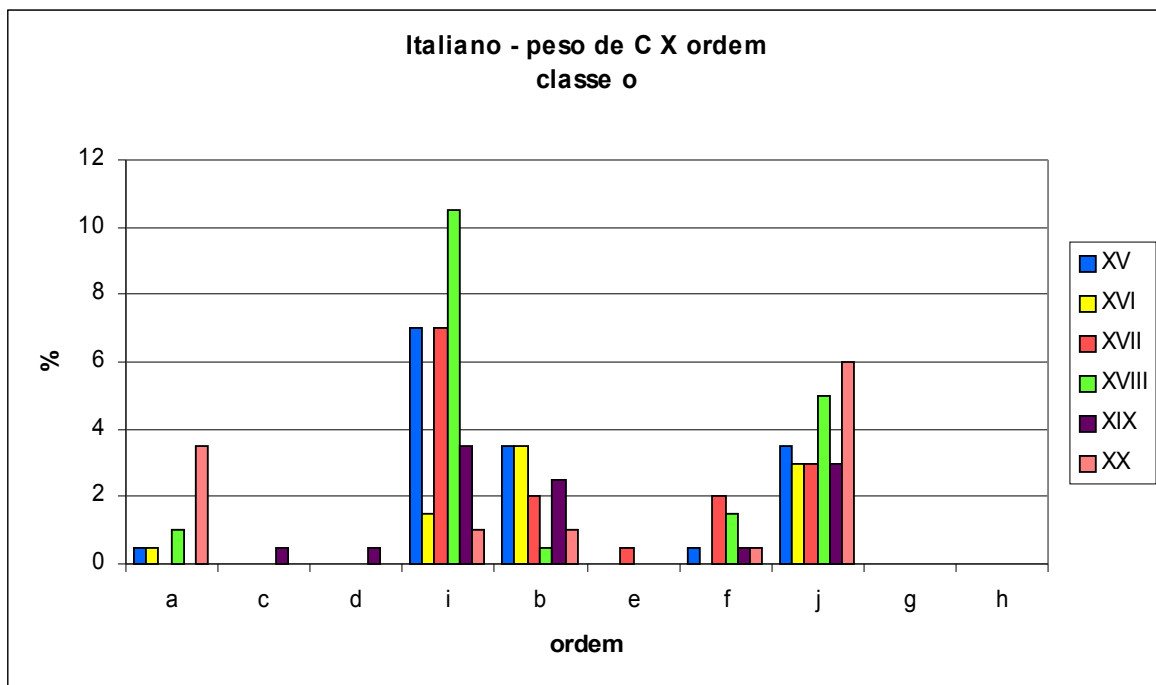


Gráfico I19d – proporções relativas de complemento verbal pronominal no *corpus* italiano.

Também a classe **p** (complemento verbal oracional) reflete uma peculiaridade do *corpus* italiano diante dos ibéricos: ainda que as ordens mais comuns sejam **a** (SVC) e **i** (VC), como acontece para os dados de português e de espanhol, verifica-se uma queda na frequência relativa de ambas, ao passo que a ordem **a** (SVC) dos conjuntos ibéricos tende a ser mais frequente.

São exemplos de ocorrências associadas a esse traço:

texto (século)	ocorrência
----------------	------------

Dyalogo (XV)	si mōstra:che la predestinatione p priegi:e per oratione si adempie:quando
Compendio (XVI)	Sapendo per certo , che si come sono dette acque & fango gioueuoli a' corpi ben disposti
Christoforo (XVII)	e la negligenza insieme del Giustiniano, che dice non saperne la cagione
Lettere (XVIII)	di volersi far tale , io credo che prima d'ogni altra cosa egli debba piacere a se stesso ; e per ciò , innanzi
Campanella (XIX)	di S.ta Severina, annunziando che dirigevasi a Roma per dar conto del suo operato, ed egli intanto
Eco (XX)	E dopo un poco vedi che molti vengono a te, anche da terre lontane, e ti considerano un profeta, o un

Tabela I17 – exemplos de ocorrências de complemento verbal oracional.

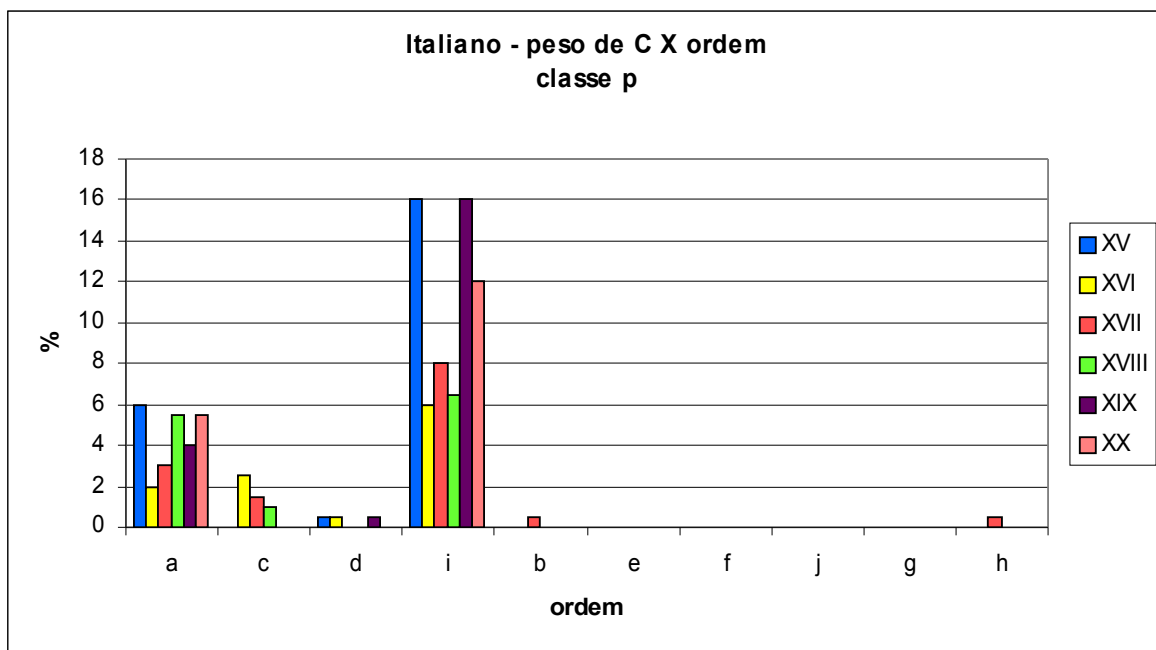


Gráfico I19e – proporções relativas de complemento verbal oracional no *corpus* italiano.

A Tabela I18 mostra que alguns traços parecem controlar o comportamento de vários grupos de ocorrências, como o sujeito pleno (**k**) – sempre crescente em frequência –, o sujeito pronominal (**l**) e o complemento oracional (**p**) – ambos decrescentes nas ordens em que são mais salientes; a única ordem determinante parece ser, até onde o *corpus* permite discernir, **i** (VC), cujas frequências diminuem para todo tipo de complemento. Os

padrões distintos de comportamento das classes de argumentos verbais em função de seu peso corroboram as observações feitas sobre o francês por Abeillé & Godard 1997.

ordem	k (S pleno)	l (S pron.)	m (S orac.)	n (C pleno)	o (C pron.)	p (C orac.)
a (SVC)	↗	↘		↗		↘
d (VCS)			↗			
i (VC)				↘	↘	↘
b (SCV)					→	
j (CV)					↗	
g (SV)	↗	↘	→			
h (VS)	↗	↘				

Tabela I18 – comportamento das ordens mais freqüentes
segundo o peso de S e de C.

O cruzamento dos fatores *peso de S e de C e ambiente oracional* mostra, uma vez mais, que este último tem maior relevância na definição dos comportamentos dos subconjuntos de dados. Chamam a atenção no Gráfico I20a, por exemplo, as semelhanças entre os perfis das curvas de sujeito pleno, **k**, em orações principais (1) e o encontrado no Gráfico I8, que trata das freqüências absolutas dessas mesmas orações. O traçado das subordinadas (3) e a reduzida variação de freqüência da classe **k** (sujeito pleno) em orações coordenadas (2) também são semelhantes aos encontrados dos Gráficos I13 e I14. Vejamos os exemplos e o gráfico de distribuição dos dados:

- 1 – Questo nome Maria e / etiam sancto in lei maximamente cioe puro - ExposRev XV
- 2 – & di qui tutto quel bagno prese nome di acque della Masera - CompRegole XVI
- 3 – perchè queste quattro nazioni con un senso uniforme sentirono - Vico XVIII
- 4 – mentre era richiuso in prigione e di come lo avesse liberato - Graal XX

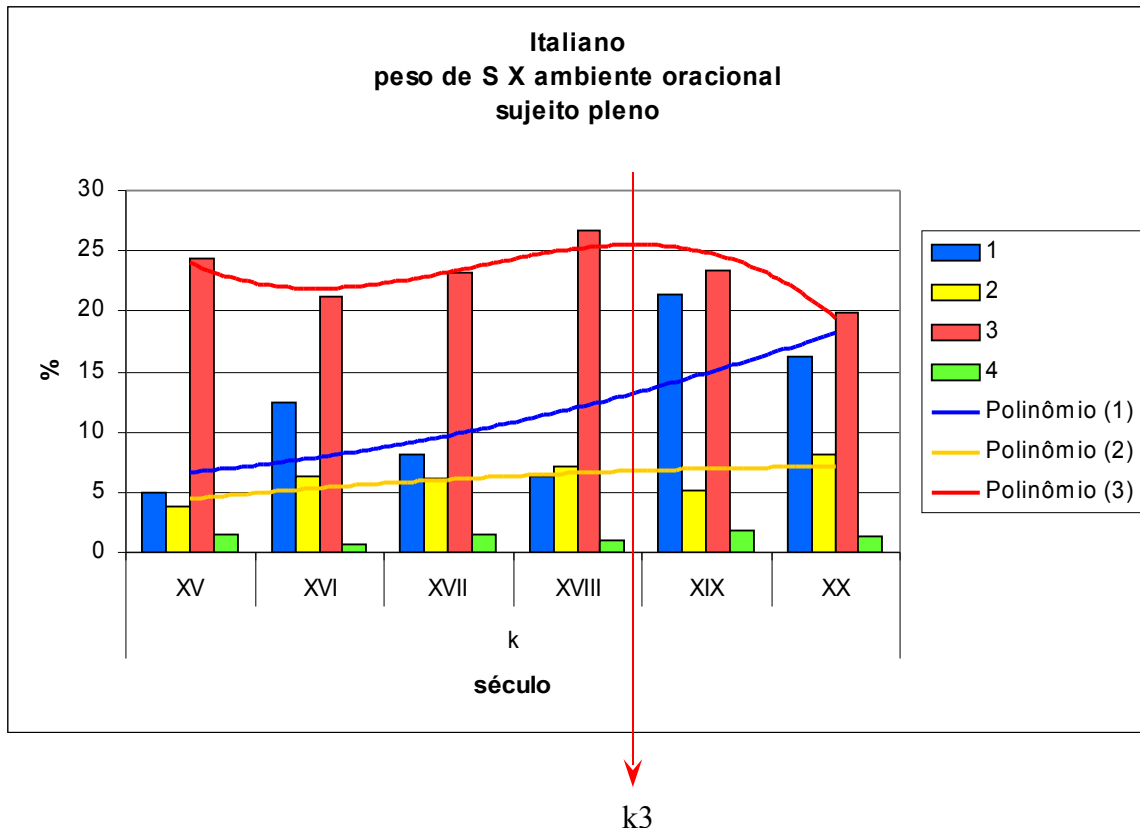


Gráfico I20a – freqüências absolutas de sujeito pleno diante do traço *ambiente oracional*.

Para o sujeito pronominal (classe I), que ocorre em proporções inferiores às da classe k (sujeito pleno), os perfis dos subconjuntos de dados mais freqüentes indicam uma tendência ao decréscimo; nota-se, além disso, que os máximos para as orações coordenadas (2) e as subordinadas (3) estão situados no século XVIII, mas as duas curvas convergem no final da amostragem. As freqüências de orações principais (1) associadas a sujeitos pronominais no *corpus* são excessivamente baixas (invariavelmente menores do que 4%), e não serão abordadas aqui. Apresentam-se, abaixo, exemplos de ocorrências associadas a esse traço no *corpus* italiano:

- 1 – Questo che io tho dicto certame[n]te se puo prouare:che - Dyalogo XV
- 2 – Cioè, non aspetasse, ne chiamasse lui, pur del medemo - LingTosc XVI
- 3 – hauendo io riceuuto risposta di molte altre mie, che portò il medesimo - ViagPellegr XVII
- 4 – la mia," disse Guglielmo, "e non so neppure se sia quella buona - Eco XX

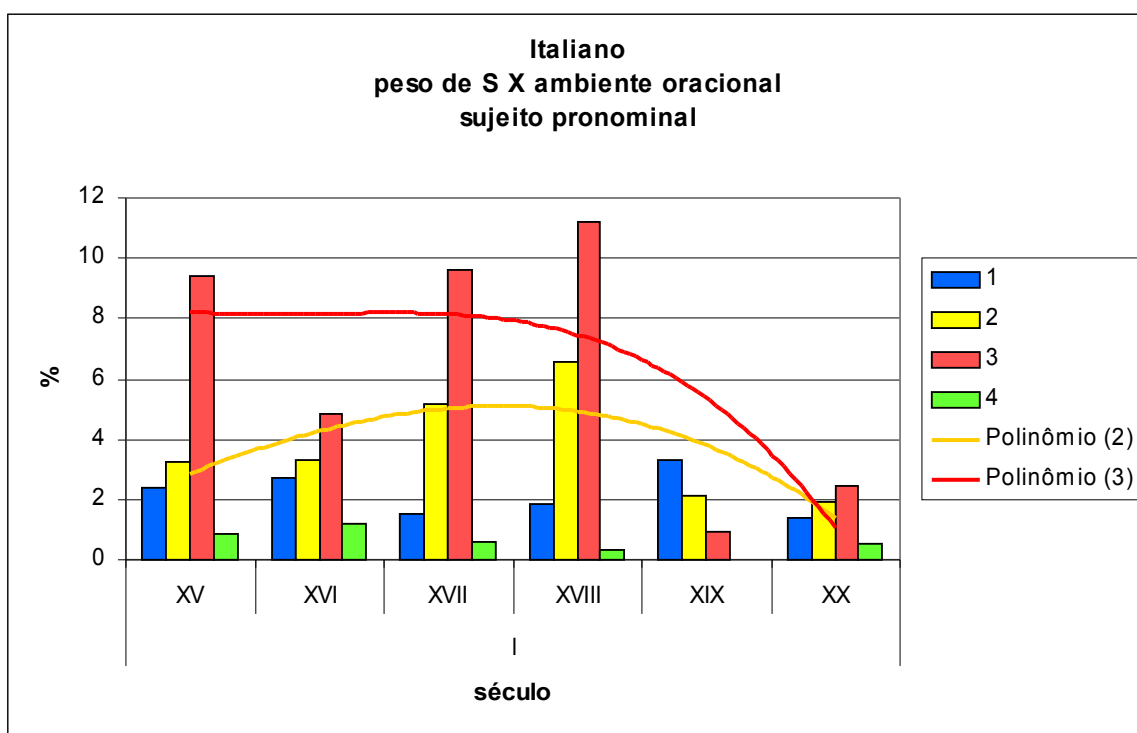


Gráfico I20b – freqüências absolutas de sujeito pronominal diante do traço *ambiente oracional*.

A classe **n**, que corresponde ao conjunto das ocorrências de complemento verbal pleno, atinge as frequências absolutas mais altas do *corpus* italiano. O perfil das classes de *ambiente oracional* é interessante: além da já esperada curva ascendente para o caráter **n** associado a orações principais (**1**), observa-se a inversão do traçado de **k3** no de **n3** (Gráfico I20d). Dito de outro modo, é possível que exista uma relação entre as duas classes; entretanto, é necessário expandir o *corpus* a fim de se poder investigar essa hipótese com mais propriedade estatística. Vejam-se os exemplos:

- 1 – AVE. Questa e una p{ar}ola salutariora et puossi transferire i{n}uulgare - ExposRev XV
- 2 – Il detto primo Dugale fu, & è chiamato Dugal della Masera, ò sia - CompReg XVI
- 3 – per verificare insieme il luogo,e'l tempo della nascita di lui - Christoforo XVII
- 4 – che dai più degli uomini... si ricercano , e vogliono acquistarsi molti superflui comodi - PrincLettere XVIII

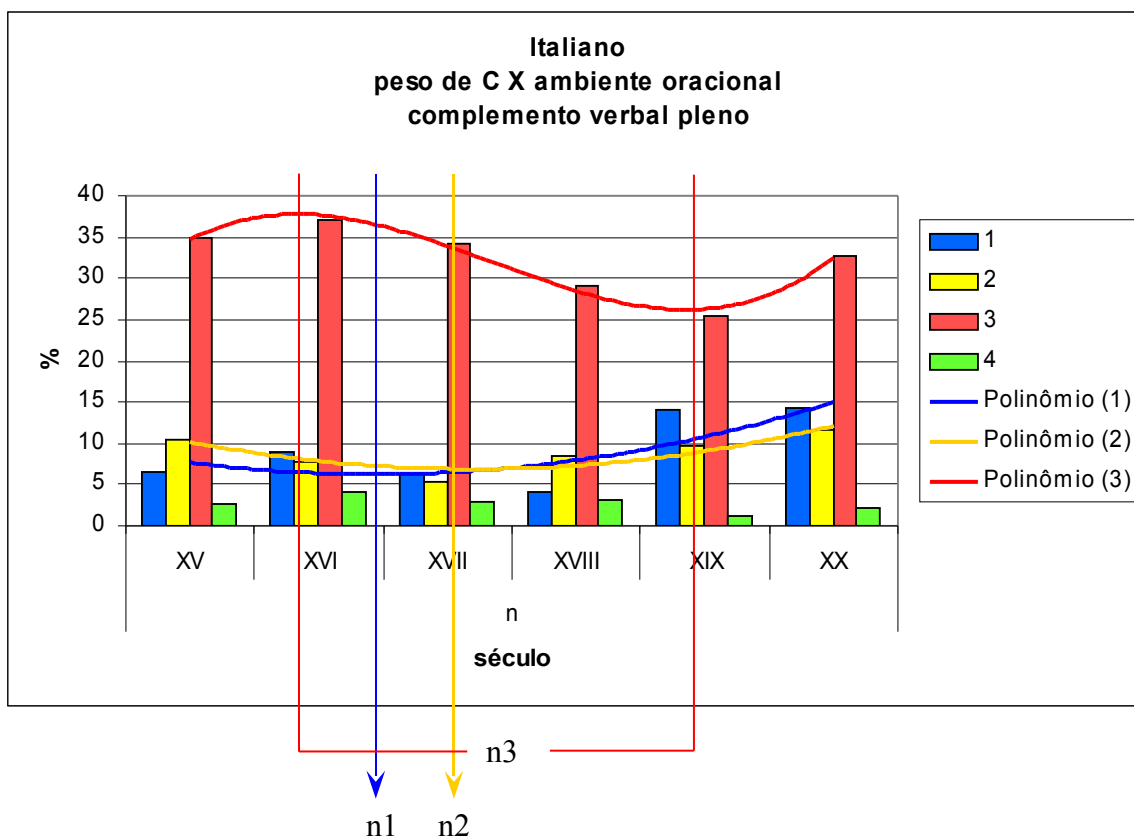


Gráfico I20c – frequências absolutas de complemento verbal pleno diante do traço *ambiente oracional*.

No gráfico abaixo, apresentam-se apenas as curvas correspondentes às frequências absolutas de **k3** e de **n3** para comparação. Notam-se imediatamente a) que as subordinadas são mais comuns que as principais em toda a amostra, e b) que as duas curvas são defasadas de $\frac{1}{2}$ período – ou seja, o máximo de uma praticamente coincide com o mínimo da outra. Entretanto, os pontos críticos de **k3** ocorrem num intervalo mais curto do que os de **n3** – ou seja, a taxa de mudança das frequências de **k3** é mais alta que a de **n3**.

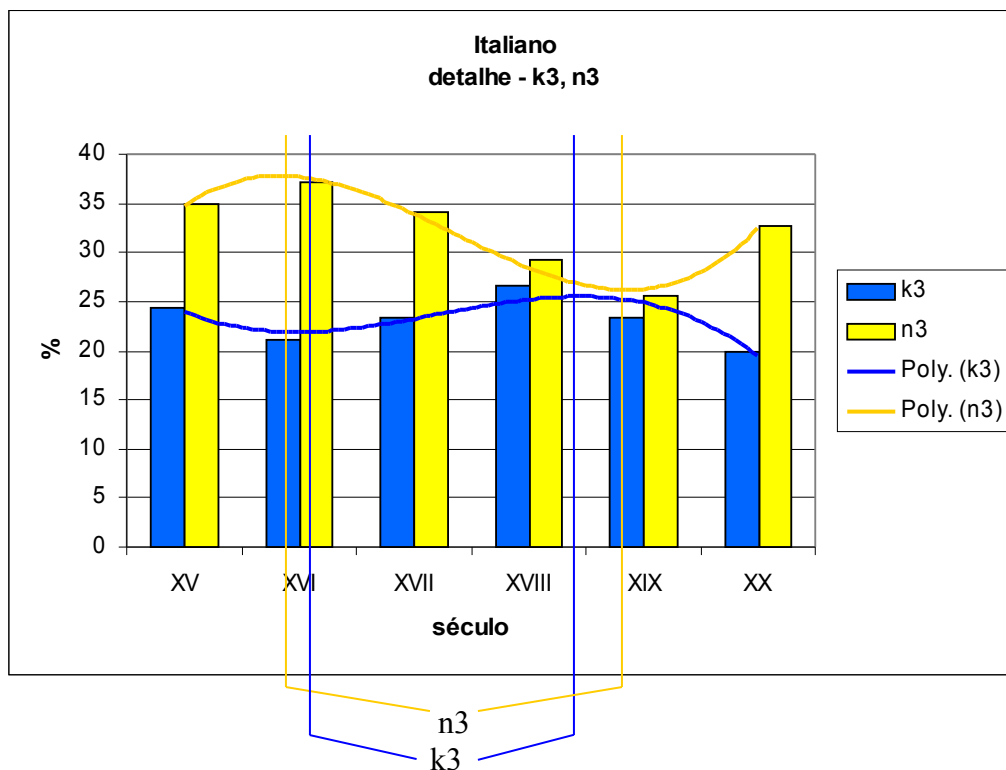


Gráfico I20d – contraposição dos perfis das classes **k** e **n** para orações subordinadas.

Os mesmos fatores podem, naturalmente, ser combinados de outra forma: fixando-se o ambiente oracional, as classes de *peso de S e de C* são comparadas em termos de frequência absoluta e de comportamento ao longo do intervalo de amostragem. Observam-se, então, para as orações principais, a predominância das classes **k** e **n**, cujos perfis são crescentes ($iRRL(k1) = 2,32$; $iRRL(n1) = 1,47$) – remetendo ao padrão de comportamento já definido para outras categorias. As outras classes ocorrem em proporções muito baixas em orações principais, esboçando-se apenas fracamente tendências ao aumento ou ao decréscimo.

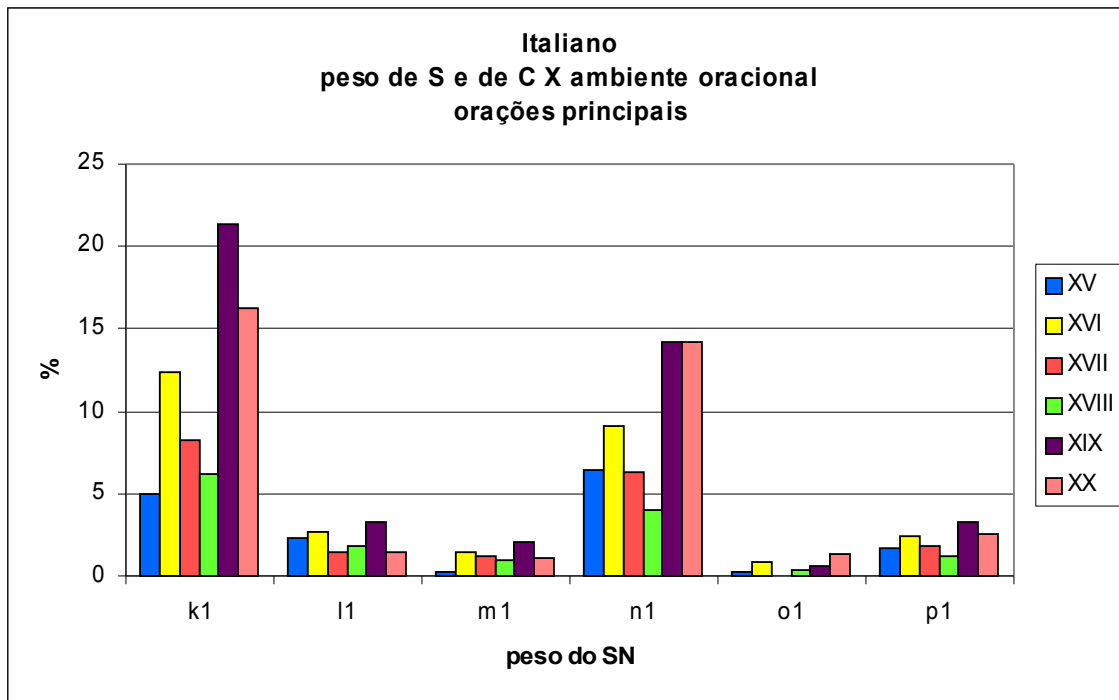


Gráfico I20e – frequências absolutas das classes de *peso de S e de C* para orações principais.

Também para as coordenadas, as classes de sujeito (**k**) e de complemento verbal pleno (**n**) apresentam perfis crescentes ($iRRL(\mathbf{k}) = .51$; $iRRL(\mathbf{n}) = .32$). Apesar de essas serem as duas classes mais freqüentes em orações coordenadas, nesse ambiente oracional há maior contribuição dos demais tipos de S (especialmente o pronominal, **l**) e de C (tanto pronominal, **o**, quanto oracional, **p**). Em outras palavras, as orações coordenadas são menos especializadas que as principais no que respeita aos tipos de argumento do V.

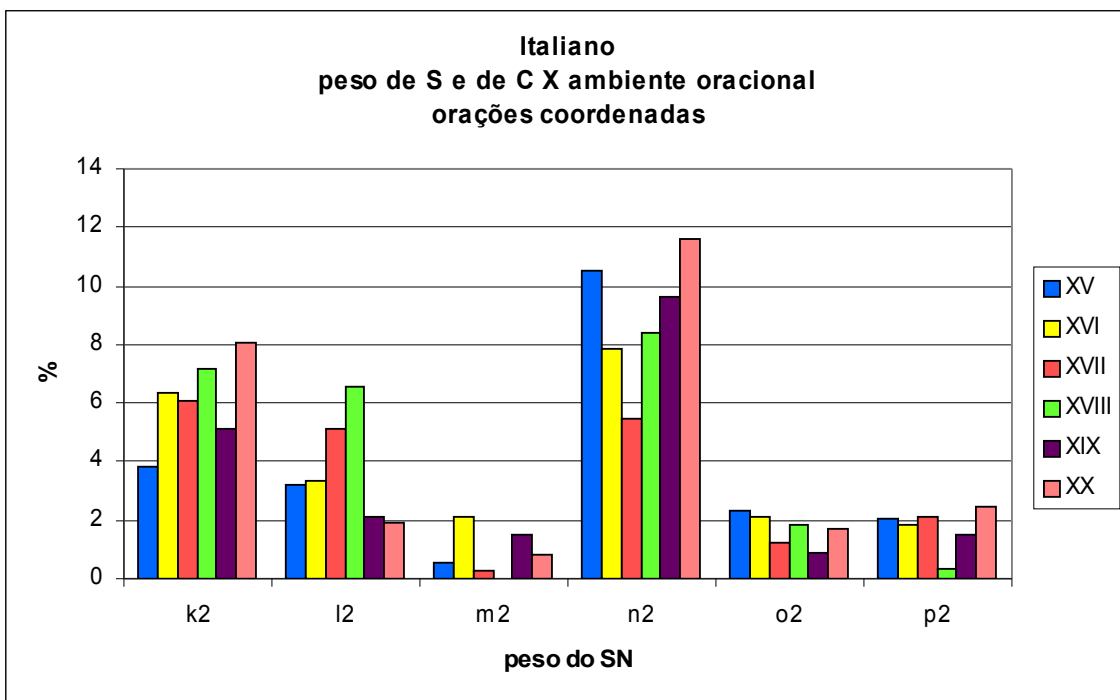


Gráfico I20f – freqüências absolutas das classes de *peso de S e de C* para orações coordenadas.

Para as subordinadas, por fim, verifica-se uma distribuição preferencial em torno das classes de sujeito pleno, **k**, e de complemento verbal pleno, **n**. Entretanto, as tendências dos dois conjuntos são decrescentes ($iRRL(\mathbf{k3}) = -1.18$; $iRRL(\mathbf{n3}) = -1.45$), a exemplo do que já foi observado para outras categorias acima, e dá suporte à afirmativa de que o fator *peso de S e de C* também está subordinado ao *ambiente oracional*.

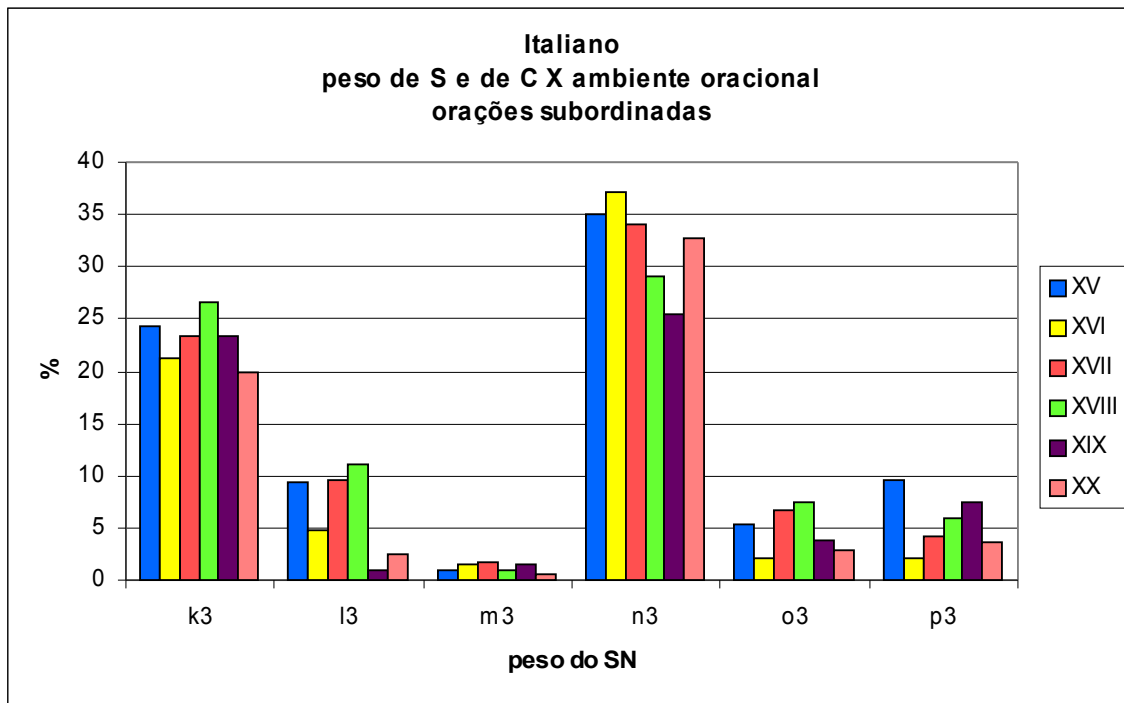


Gráfico I20g – frequências absolutas das classes de *peso de S e de C* para orações subordinadas.

O cruzamento dos fatores *regência do verbo* e *ordem* mostra que as ocorrências tendem a se concentrar nas classes **a** (SVC), **i** (VC), **g** (SV) e **h** (VS), enquanto as demais se apresentam com menor frequência. A expressividade relativa dessas classes no *corpus* é ilustrada no gráfico I23a; note-se a predominância absoluta, em todo o intervalo considerado, da classe **u** (transitivos diretos), ao passo que **v** (transitivos indiretos) ocorrem segundo proporções invariavelmente baixas.

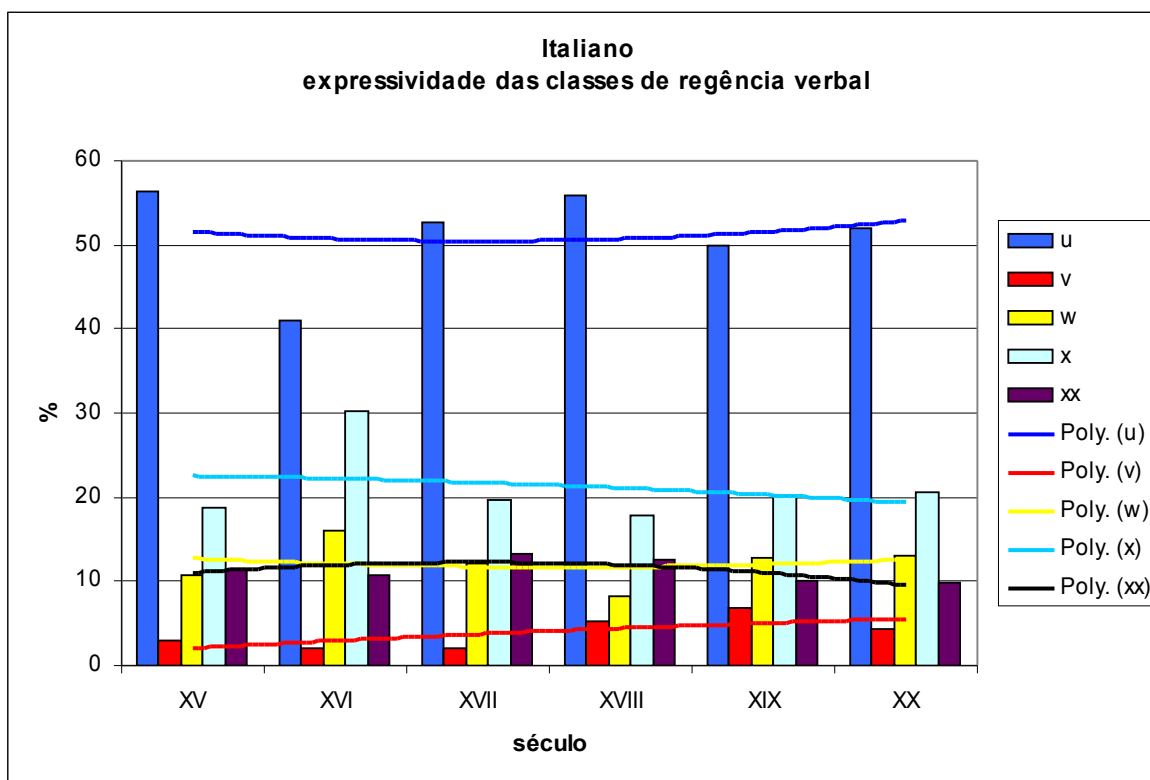


Gráfico I21a - proporções absolutas de cada classe de regência verbal no *corpus* italiano.

Os verbos de transitividade direta (**u**) estão preferencialmente associados às ordens **i** (VC) e **a** (SVC), mas os perfis dessas duas ordens são diferentes: **i** (VC) diminui em frequência (iRRL = -0,68), enquanto **a** (SVC) aumenta (iRRL = 2,37).

São exemplos de ocorrências associadas a esse traço:

ExposRev (XV)	lasua signoria sidebbe i{n}ogniluogo dimandare	g (SV)
LingTosc (XVI)	con le sue lagrime la lauò / Lauò essa testa di Lorenzo	i (VC)
Christoforo (XVII)	nō fù sì facile ad altri il saperlo, nè egli manifestare il volle giamai	i (VC)
Principj (XVIII)	con promuovere fuori d' ordine huomini di straordinario merito agli onori ;	i (VC)
Campanella (XIX)	che il Campanella ed il Ponzio negavano la ribellione ma confessavano l' eresia, per	a (SVC)
Eco (XX)	I valdesi predicavano una riforma dei costumi all'interno della chiesa, i catari	a (SVC)

Tabela I19 – exemplos de ocorrências com verbos transitivos diretos.

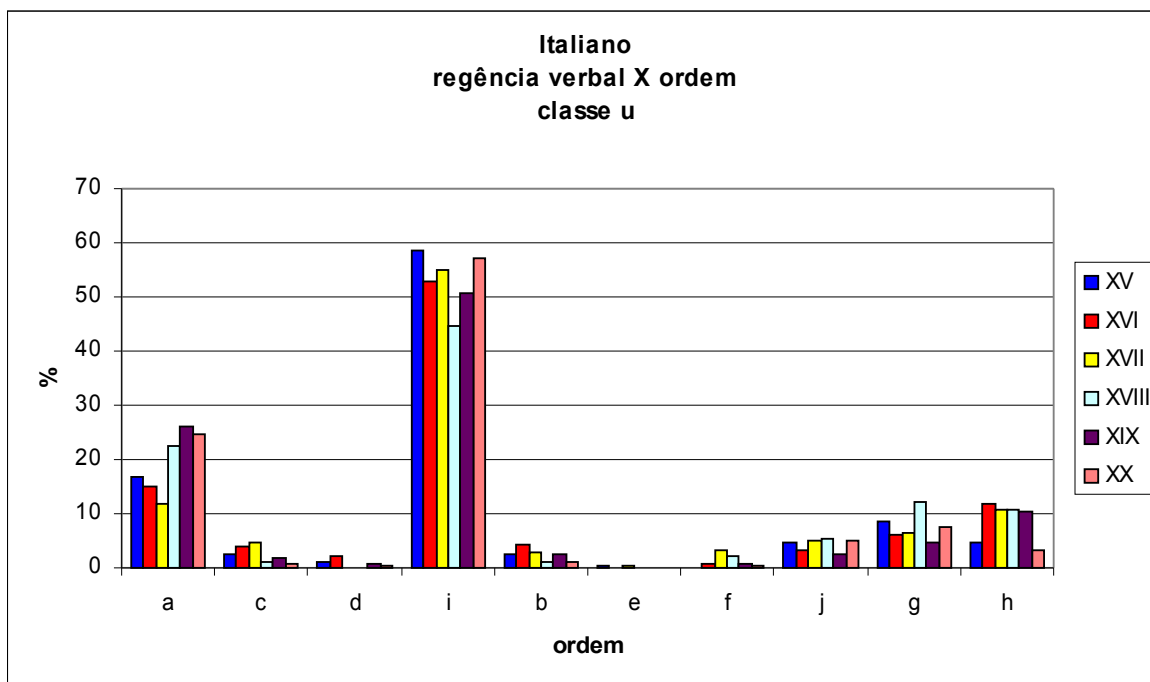


Gráfico I21b - proporções absolutas de verbos transitivos diretos no *corpus* italiano.

Para a classe **v** (verbos transitivos indiretos), as tendências indicam aumento na frequência das ordens diretas, **a** (iRRL(SVC) = 2,72), **i** (iRRL(VC) = 2,13) e **g** (iRRL(SV) = 5,83), e diminuição do número relativo de ocorrências de ordem **h** (iRRL(VS) = -1,63). É interessante notar que a ordem **g** (SV) é mais expressiva que **a** (SVC) e **i** (VC) juntas, indicando uma especialização desse tipo de verbo nessa mesma classe de *ordem de constituintes*.

São exemplos de ocorrências associadas a esse traço:

Dyologo (XV)	lo qual desiderando di piu piacere a dio : si fece monacho in quello luochu	i (VC)
LingTosc (XVI)	perche se sil Pronome non serue ad altro, che a'rappresentar'alcun nome	a (SVC)
Christoforo (XVII)	si esposse il primo ad vna nuoua insolita nauigatione, solcò	c (VSC)
Lettere (XVIII)	Ma , per Catone , un mezzo verso basta a Virgilio ; tre soli per Giunio Bruto ;	a (SVC)
Geografia (XIX)	essi appartenevano probabilmente alla stessa famiglia	a (SVC)
Graal (XX)	L'evoluzione ultima torna alla raffigurazione più semplice con sole tre figure	a (SVC)

Tabela I.20 – exemplos de ocorrências com verbos transitivos indiretos.

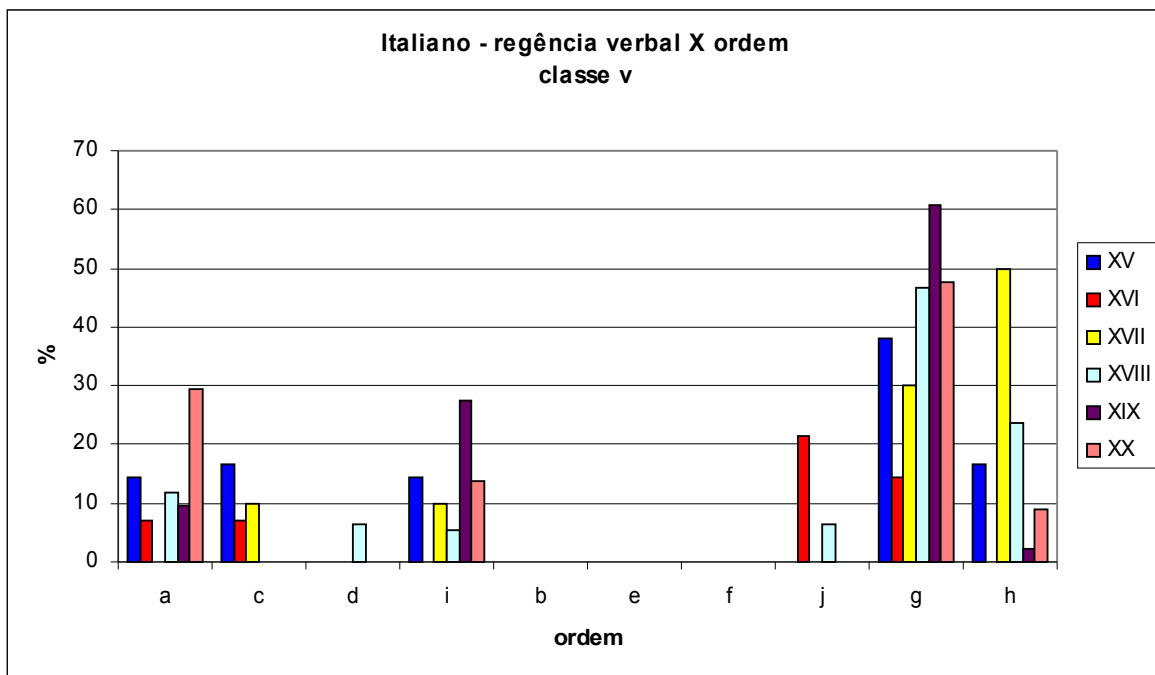


Gráfico I21c - proporções absolutas de verbos transitivos indiretos no *corpus* italiano.

O comportamento da classe **w** (verbos bitransitivos) assemelha-se, em vários aspectos, ao da classe **u** (transitivos diretos). Vejam-se a predominância da ordem **i** (VC), que diminui em frequência relativa (iRRL = -0,44), ao passo que as proporções das ordens que têm sujeito expresso, **a** (SVC), **g** (SV) e **h** (VS), tendem a aumentar, mesmo que discretamente (iRRL(**a**) = 1,53; iRRL(**g**) = 0,43; iRRL(**h**) = 0,86).

São exemplos de ocorrências associadas a esse traço:

Expositione (XV)	che qualche volta per me la offerischino alla madre del mio signore	j (CV)
Compendio (XVI)	si parte rendendone gratie al Sõmo Iddio donatore di ogni gratia	i (VC)
Principj (XVIII)	Repubbliche libere , con proporre a'sedotti o libertà , o potenza , o ricchezze :	i (VC)
Campanella (XIX)	Egli trasmise questa notizia al suo Governo, e contemporaneamente partecipò	a (SVC)
Eco (XX)	Puoi anche aggiungere alla mia pessima allegoria l'immagine di qualcuno che tenta	i (VC)

Tabela I.21 – ocorrências associadas ao traço **w** (verbo bitransitivo) no *corpus* italiano.

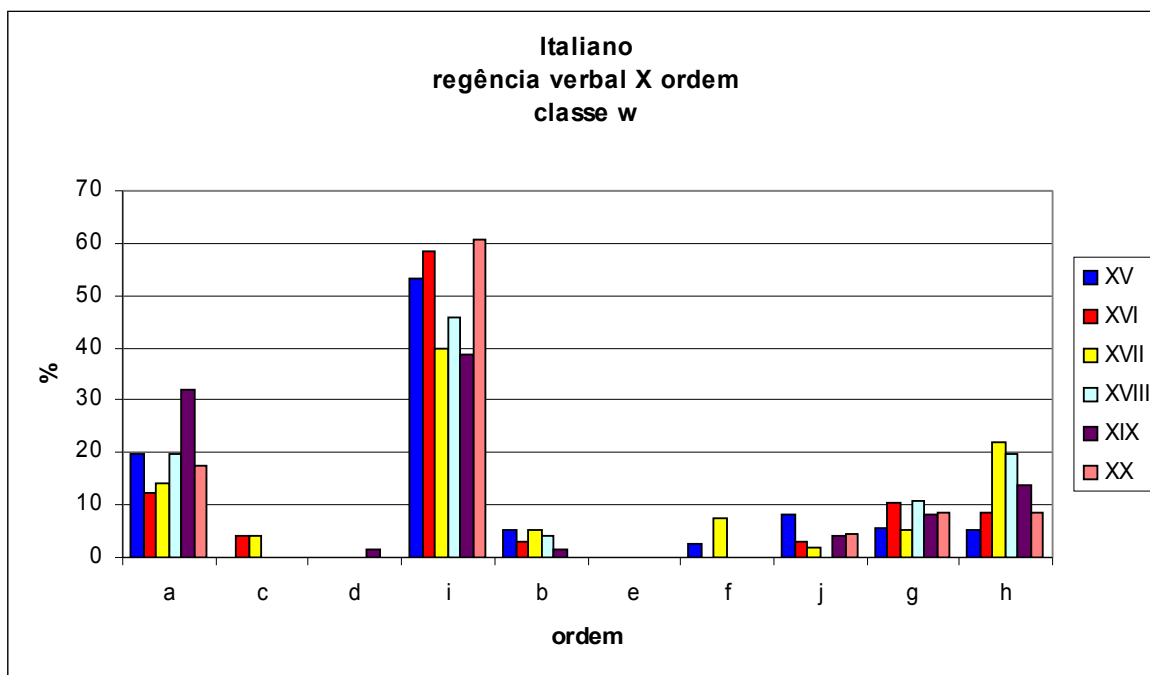


Gráfico I21d - proporções absolutas de verbos bitransitivos no *corpus* italiano.

Os verbos de ligação (classe **x**) mostram, em relação às outras classes para o mesmo fator, maiores proporções relativas das ordens **c** (VSC) e **d** (VCS), além de a ordem **a** (SVC) ser mais comum que **i** (VC). Feitas essas considerações, observa-se que, dentre essas ordens, as que têm expressos os dois argumentos do verbo tendem a aumentar em frequência (iRRL(**a**) = 3,61; iRRL(**c**) = 0,04; iRRL(**d**) = 1,47), enquanto as outras tendem a diminuir (iRRL(**i**) = -3,12; iRRL(**g**) = -0,74; iRRL(**h**) = -0,05).

São exemplos de ocorrências associadas a esse traço:

Expositione (XV)	p{er}che ilpadre et ilfigliuolo et ilspirito sancto sono uno dio benedecto in eterno	a (SVC)
LingToscana (XVI)	Indeclinabili sono Egli, Chi... e tutti gli altri di questa natura	f (CVS)
Christoforo (XVII)	non esser ella legitima,e vera,mà finta,&inuentata dal Gallo	c (VSC)
Lettere (XVIII)	di mano ? più certo era il pane ; non era infame il mezzo ; e non avreste così dovuto	d (VCS)
Geografia (XIX)	Il De Rougè pensa che questa parola sia la designazione d'una grande divisione geografica	a (SVC)

Tabela I22 – ocorrências associadas ao traço **x** (verbo de ligação) no *corpus* italiano.

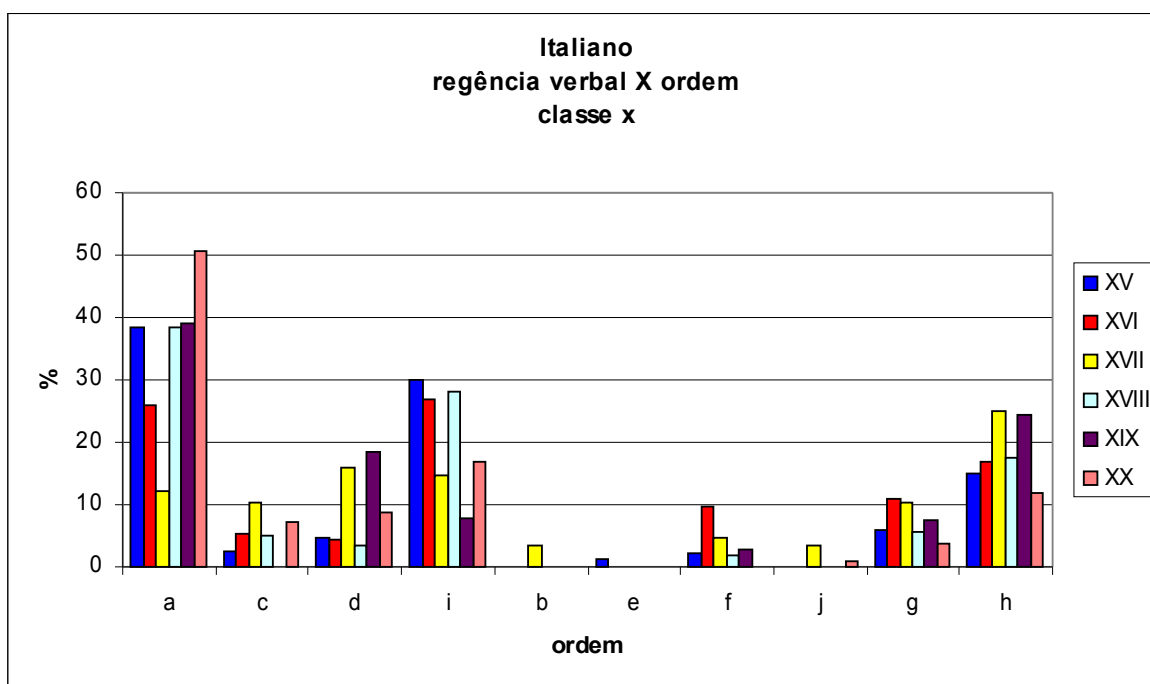


Gráfico I21e - proporções absolutas de verbos de ligação no *corpus* italiano.

Para a classe dos verbos intransitivos, **xx**, as duas ordens segundo as quais as ocorrências se distribuem, **g** (SV) e **h** (VS), apresentam-se em variação estável: $iRRL(\mathbf{g}) = 0,03$, $iRRL(\mathbf{h}) = -0,03$. Deve-se ressaltar o fato de que as frequências absolutas giram em torno dos 10% (q.v. Gráfico I.23).

São exemplos de ocorrências associadas a esse traço:

Expositione (XV)	p{er}che ilpadre et ilfigliuolo et ilspirito sancto sono uno dio benedecto in eterno	a (SVC)
LingToscana (XVI)	Indeclinabili sono Egli, Chi... e tutti gli altri di questa natura	f (CVS)
Christoforo (XVII)	non esser ella legitima,e vera,mà finta,&inuentata dal Gallo	c (VSC)
Lettere (XVIII)	di mano ? più certo era il pane ; non era infame il mezzo ; e non avreste così dovuto	d (VCS)
Geografia (XIX)	Il De Rougè pensa che questa parola sia la designazione d'una grande divisione geografica	a (SVC)
Graal (XX)	È difficile collegare questo personaggio a una qualsiasi narrazione	d (VCS)

Tabela I23 – ocorrências associadas ao traço **xx** (verbo intransitivo) no *corpus* italiano.

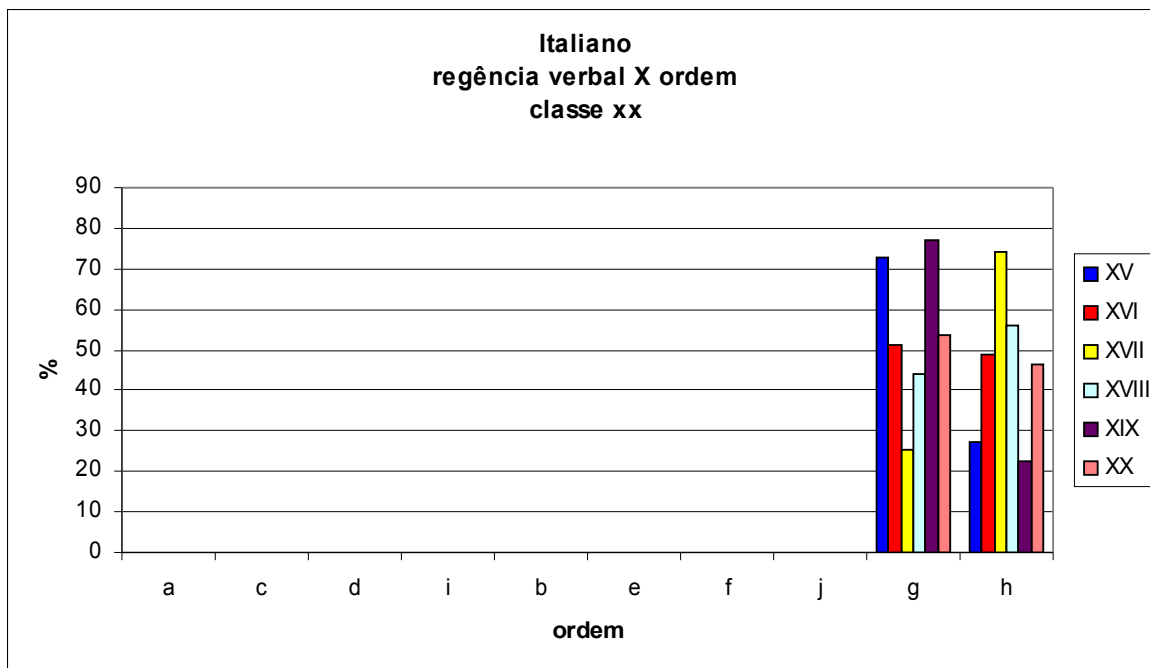


Gráfico I21f - proporções absolutas de verbos intransitivos no *corpus* italiano.

A Tabela I29 sumariza os comportamentos das ordens mais comuns no *corpus*; observa-se que, para todas as classes de regência, a ordem **a** (SVC) aumenta em frequência (cf. Gráfico I.2).

Em alguns casos, há comportamentos que parecem ser determinados pela presença e pela posição dos argumentos do verbo: a classe **v** (trans. indiretos), por exemplo, diminui de frequência quando o S é pós-verbal; **w** (bitransitivos) o faz quando o S não é expresso, e **xx** (intransitivos) só aumenta na ordem direta **g** (SV). É razoável afirmar, portanto, que há um condicionamento – ainda que fraco – da distribuição das classes de *regência verbal* relacionado ao fator *ordem de constituintes*.

<i>regência do verbo</i>	a (SVC)	i (VC)	g (SV)	h (VS)
<i>u</i> (trans. diretos)	↗	↘	→	↘
<i>v</i> (trans. indiretos)	↗	↗	↑	↘
<i>w</i> (bitransitivos)	↗	↘	↗	↗
<i>x</i> (ligação)	↗	↘	↘	→
<i>xx</i> (intransitivos)			↗	↘

Tabela I.24 – comportamento das ordens mais frequentes no *corpus* italiano segundo a regência do verbo.

Graças ao cruzamento dos fatores *regência do verbo* e *ambiente oracional*, podem-se observar em detalhe as relações entre as diversas categorias de dados. A classe **u** (verbos transitivos diretos) exhibe a mesma concentração em orações subordinadas (**3**) e os mesmos perfis que compõem o Gráfico I8. Há uma tendência ao aumento das frequências absolutas de verbos dessa classe em orações principais (**1**), seguida de perto por uma diminuição compensatória na frequência da classe **u** (trans. diretos) em orações subordinadas (**3**); as frequências de orações coordenadas (**2**) apresentam um perfil estável. Aos exemplos:

- 1 – Gabriel fu mandato di Cielo i terra a fare lambasciata della - ExposRev XV
- 2 – vi si sta sentēdone cōtentezza, & con gaudio si parte - CompRegole XVI
- 3 – che dell'importante impresa, ch'ei fece, non ha fin'hora il Mondo - Christoforo XVII
- 4 – e che all'entrata del porto ne appiccassero quattro alle antenne - Campanella XIX

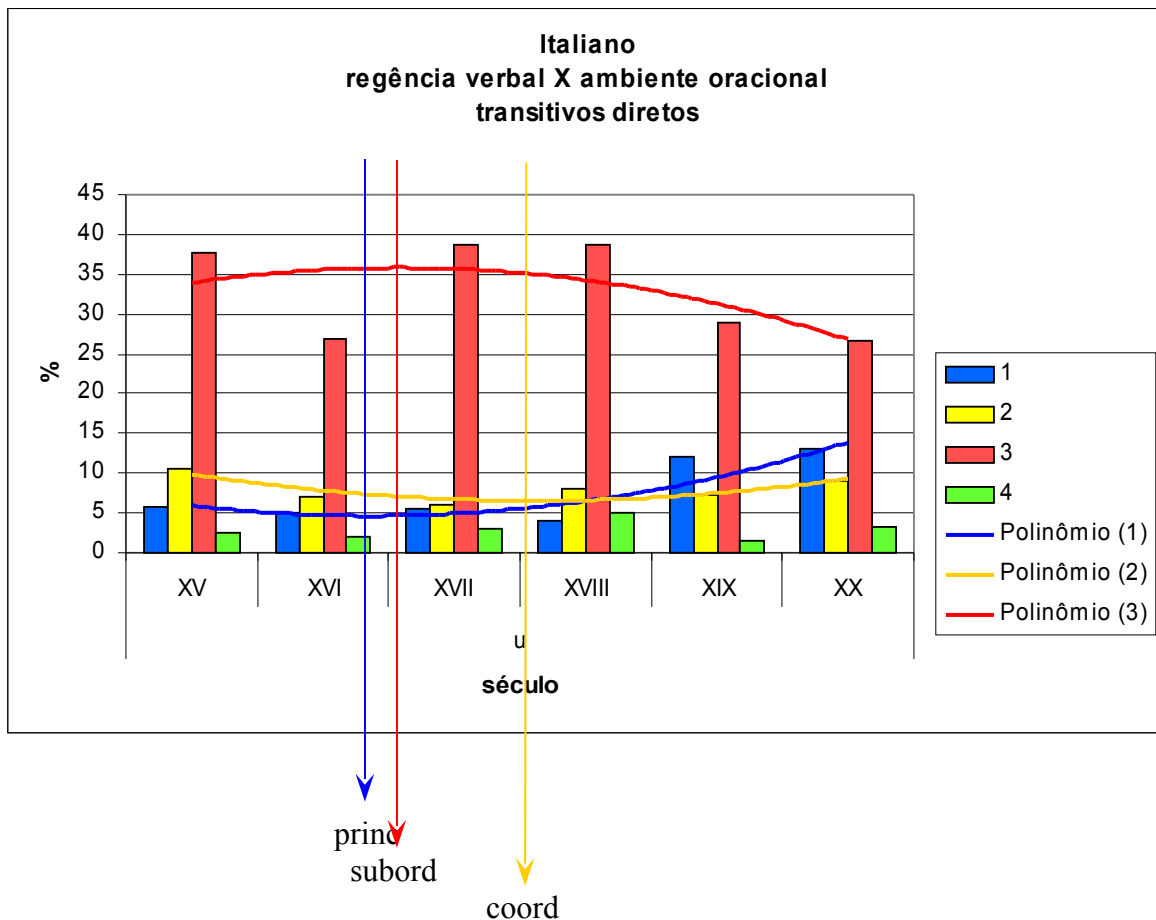


Gráfico I22a – frequências absolutas de verbos transitivos diretos segundo o ambiente oracional.

Os verbos de ligação (classe **x**), por sua vez, têm um comportamento peculiar no que toca aos diferentes ambiente oracionais em que ocorrem. Apesar de as freqüências serem sensivelmente mais baixas do que as observadas para os verbos transitivos diretos, os pontos críticos das curvas senoidais que representam as subconjuntos de orações principais, coordenadas e subordinadas são próximos uns dos outros. Além disso, nota-se que as curvas obedecem a uma seqüência **1 > 2 > 3**: máximos e mínimos seguem a ordem principais > coordenadas > subordinadas, como se as primeiras liderassem um movimento em direção ao aumento ou à diminuição das freqüências absolutas dessa classe **x** na amostra.

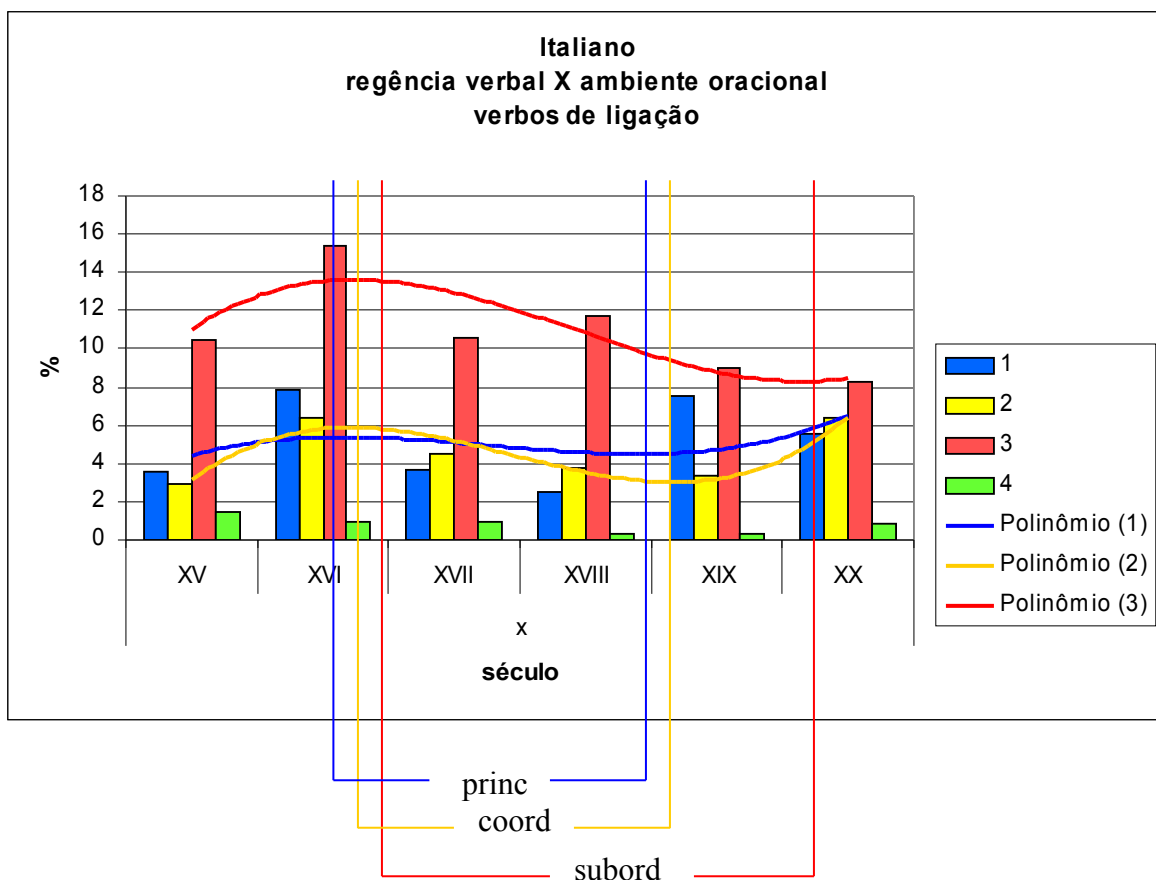


Gráfico I22b – freqüências absolutas de verbos de ligação segundo o ambiente oracional.

O mesmo par de critérios pode ser utilizado para se abordar o comportamento da amostra de outro ângulo: para o mesmo *ambiente oracional*, podem-se observar as tendências das classes de *regência verbal*. Os perfis das curvas associadas a verbos transitivos diretos (classe **u**) e a verbos de ligação (classe **x**) são diferentes, mas os pontos críticos das mesmas praticamente coincidem no século XVI. A curva da classe **u** (trans. diretos) segue o traçado parabólico (com concavidade para cima) previsto para as orações principais (1) no Gráfico I8.

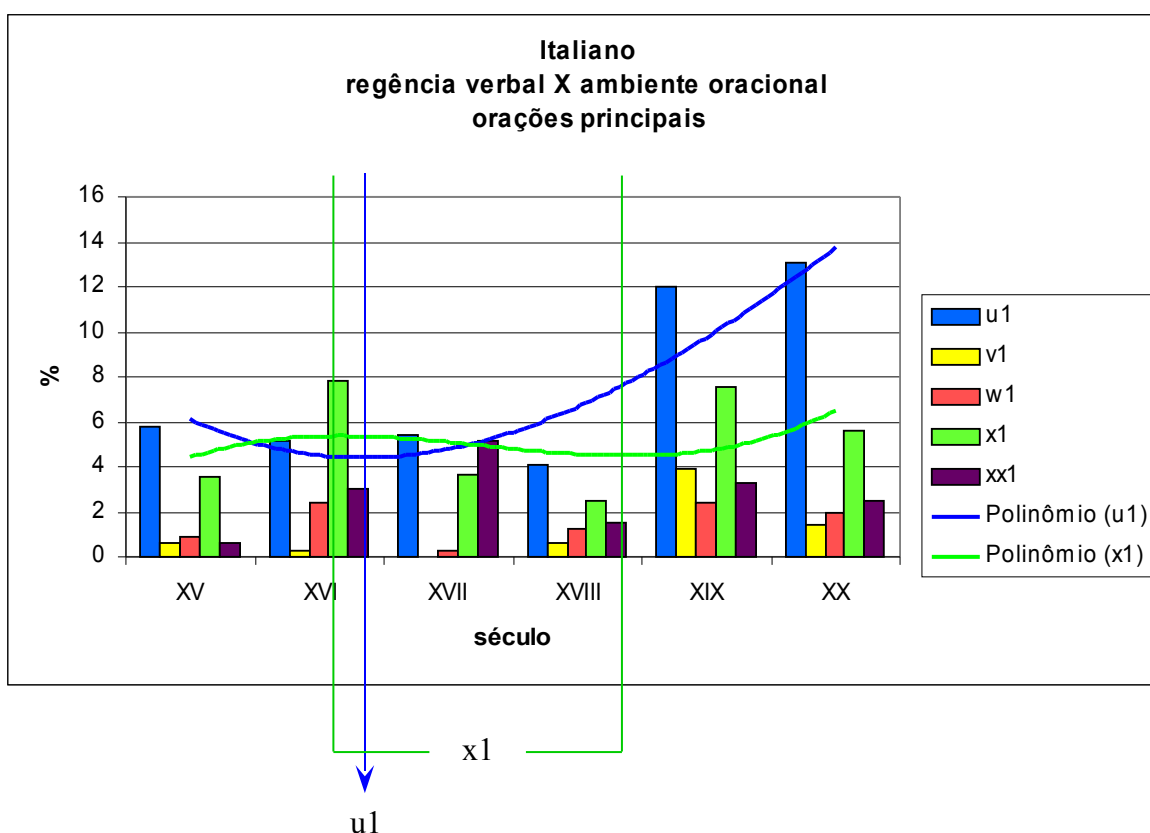


Gráfico I22c – freqüências absolutas das classes de *regência verbal* em orações principais.

As freqüências absolutas das classes de *regência verbal* associadas a orações coordenadas estão quase sempre abaixo dos 10% e sua variação máxima não passa dos 4%, motivo pelo qual não serão tratadas aqui.

As classes de *regência verbal* associadas a orações subordinadas exibem frequências absolutas elevadas, e seus perfis são semelhantes a outros já apresentados para esse mesmo ambiente oracional. As tendências gerais, resolvidas por regressão linear, apontam decréscimo generalizado das duas categorias mais frequentes ($iRRL(u3) = -1,42$; $iRRL(x3) = -0,83$) no Gráfico I22.

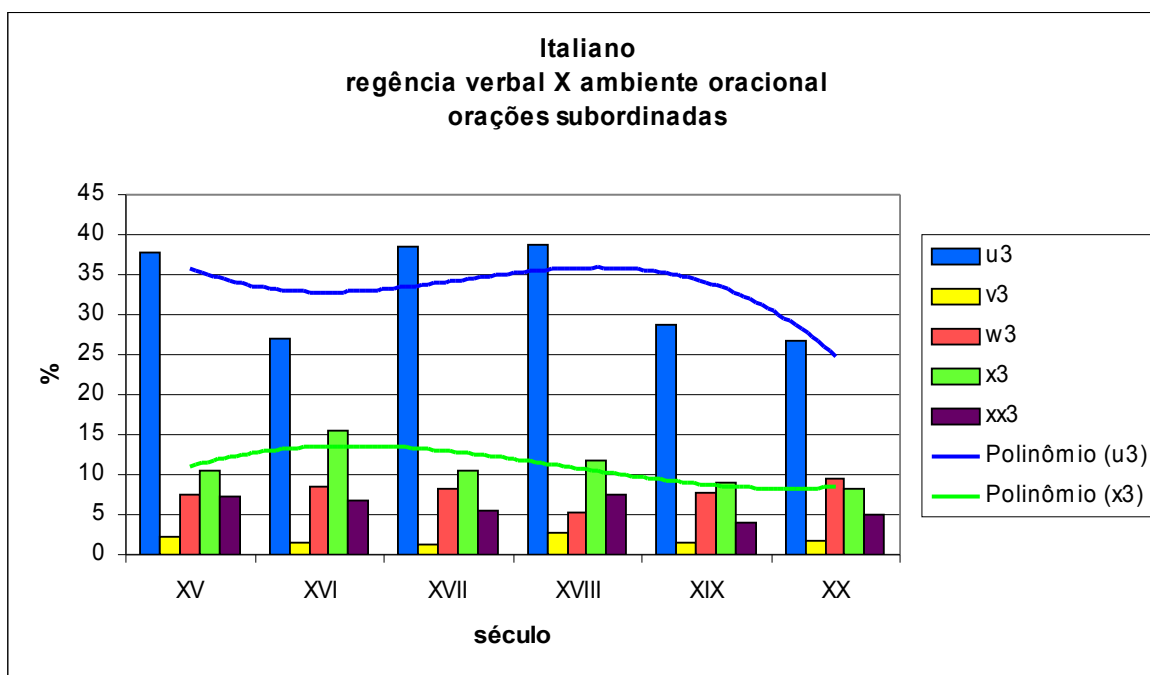


Gráfico I22d – frequências absolutas das classes de *regência verbal* em orações principais.

As orações coordenadas em subordinação ocorrem em frequências sempre abaixo dos 5%, e não serão discutidas aqui.

Com respeito à classificação das ocorrências segundo a vogal temática do verbo, os subconjuntos mais frequentes são aqueles associados às ordens **a** (SVC), **i** (VC), **g** (SV) e **h** (VS). VC destaca-se diante das demais ordens, apesar de exibir uma tendência a diminuir em frequência absoluta (iRRL = -1,33). Dentre os outros subconjuntos mais expressivos de dados, VS flutua, mas também diminui levemente (iRRL = -0,39); por outro lado, as frequências de SVC e de SV aumentam no intervalo de amostragem (iRRL(**a**) = 1,94; iRRL(**g**) = 0,91). Em suma, tendem a aumentar as frequências absolutas das ordens que têm sujeito pré-verbal; as demais, que trazem apenas um argumento (S ou C) pós-verbal, vêm suas frequências absolutas diminuir.

São exemplos de ocorrências a que se atribui esse traço:

Del qle Gaudentio prete:lo qle e anchora viuo narra molti miraculi [Dyalogo XV]

& de' proprij cõtadini che laورانo detti luoghi [Compendio XVI]

Il Weisgerber pensa pure che questo movimento di migrazione dei popoli [Geografia XIX]
voi vi sforzavate di spiegargli la differenza tra eretico e eretico, e tra eretico e ortodosso.
[Eco XX]

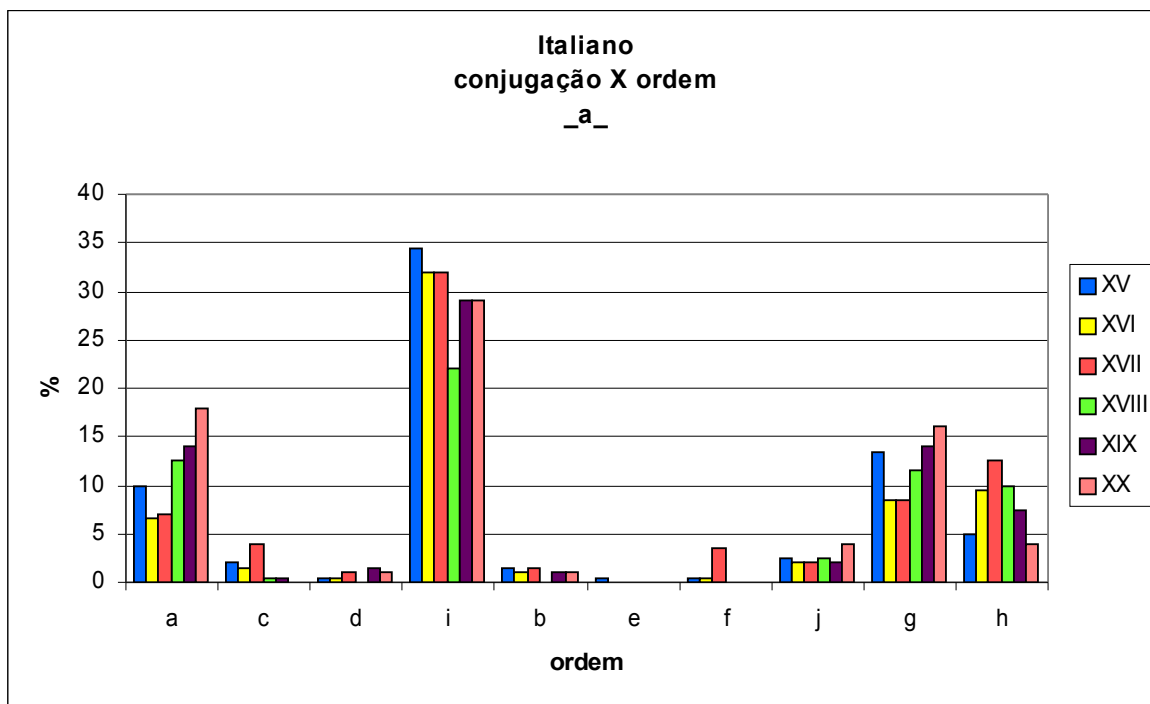


Gráfico I23a – frequências absolutas de verbo de primeira conjugação no *corpus* italiano.

As freqüências absolutas de verbos da segunda conjugação guardam uma desproporção menos acentuada entre as ordens **a** (SVC) e **i** (VC), pois há mais verbos dessa conjugação na ordem SVC do que verbos de vogal temática **_a_**. Não obstante a flutuação dos valores relativos a cada século, as tendências gerais de cada subconjunto, definidas por regressão linear, são semelhantes às encontradas para os verbos de primeira conjugação: $iRRL(\mathbf{a}) = 0,69$; $iRRL(\mathbf{i}) = -0,36$; e $iRRL(\mathbf{h}) = -0,47$. Apenas a ordem **g** (SV) tende à estabilidade, além de ser menos comum do que **h** (VS).

São exemplos de ocorrências classificadas sob esse traço:

e / desideroso che lhavesse maggior gratia dadio che epsa non haueua [Expositione XV]
 Occorre per tão a chi parla bene spesso tacere il nome d'alcuna cosa [LingToscana XVI]
 perche non pare da credersi,che a'Genouesi legato alcun facesse il Colombo [Christoforo XVII]

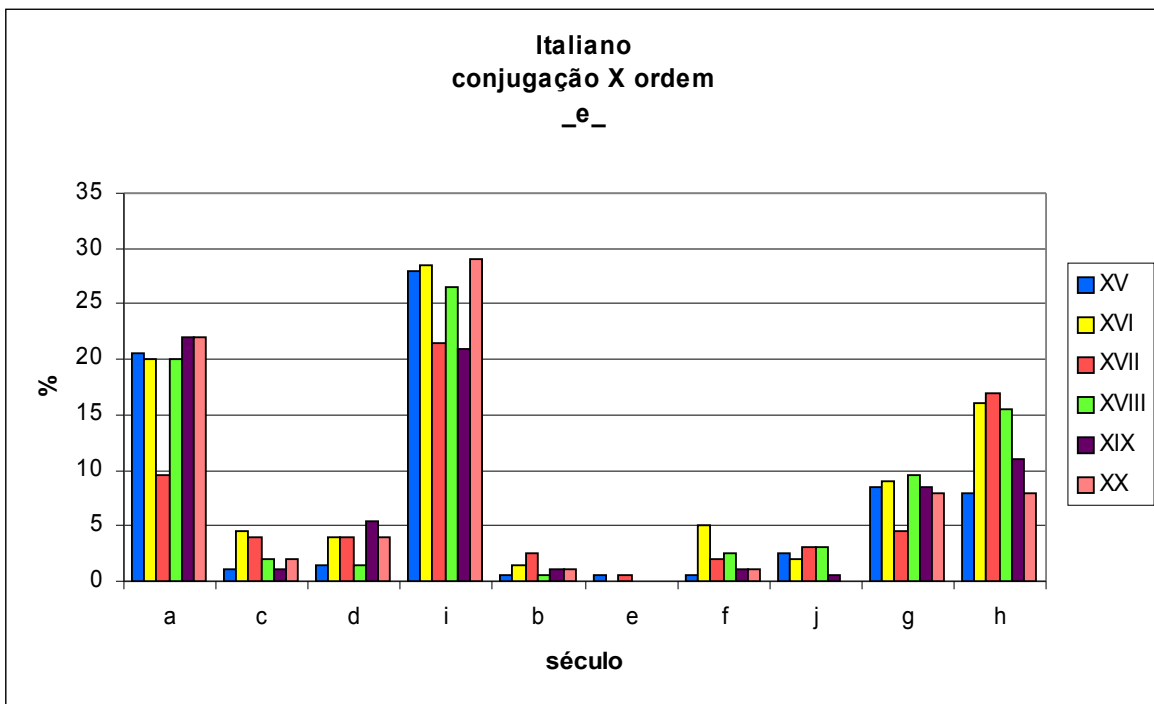


Gráfico I23b – frequências absolutas de verbo de segunda conjugação no *corpus* italiano.

A terceira conjugação, por sua vez, frequenta o *corpus* italiano em proporções mais baixas do que as duas anteriores. Ainda assim, esboçam-se tendências de aumento para as ordens **a** (SVC) e **g** (SV) ($iRRL(\mathbf{a}) = 0,76$; $iRRL(\mathbf{g}) = 0,54$) – i.e., nas duas em que o S é pré-verbal – e de decréscimo para **i** (VC) ($iRRL(\mathbf{i}) = -0,63$); **h**, que tem sujeito pós-verbal, flutua em torno de uma RRL subhorizontal ($iRRL(\mathbf{h}) = 0,17$).

São exemplos de ocorrências associadas a esse traço:

Ora ti voglio dire alchuna cosa piu di presso : la quale [Dyologo XV]

Le darò dunque ragguaglio, per seguir lo stile incominciato [Viaggi XVII]

Giuseppe di Arimatea a seppellire il corpo di Gesù secondo l'usanza ebraica [Graal XX]

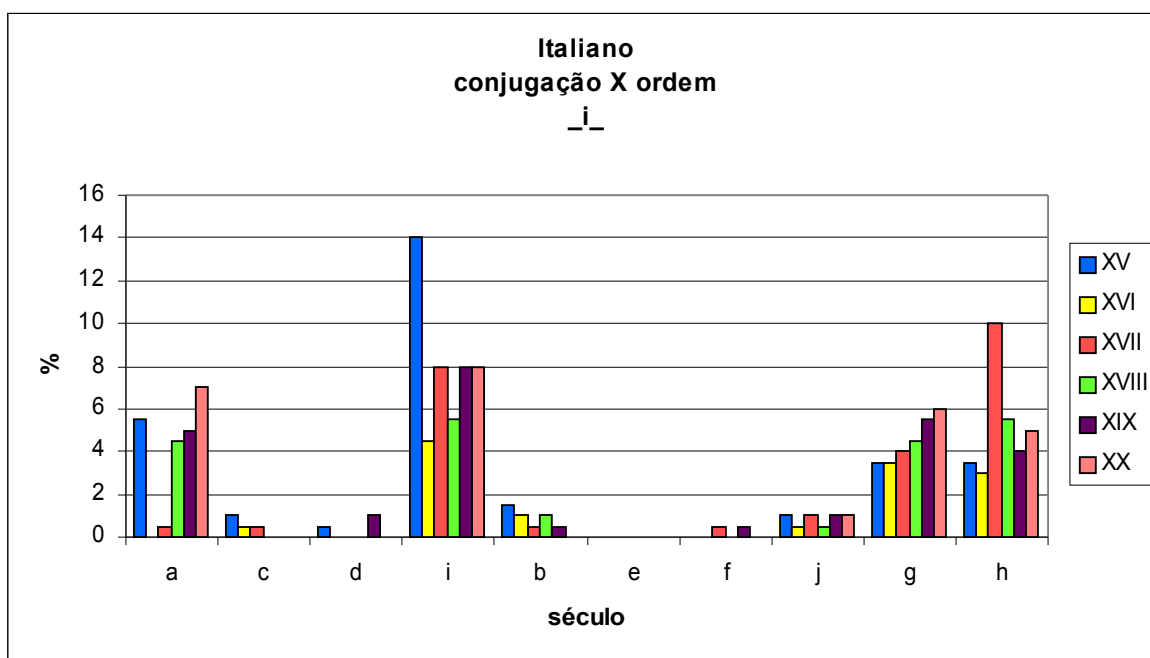


Gráfico I23c – frequências absolutas de verbo de terceira conjugação no *corpus* italiano.

A tabela abaixo resume as observações feitas sobre essas ordens mais comuns; nota-se um padrão também orientado pela *ordem de constituintes*, e segundo o qual a) a classe a (SVC) aumenta em frequência para todas as conjugações, e b) as ordens em que o sujeito é pré-verbal são cada vez mais comuns, ao passo que há diminuição das proporções daquelas cuja posição pós-verbal é a única preenchida:

<i>ordem</i>	<i>I</i>	<i>II</i>	<i>III</i>
<i>a (SVC)</i>	↗	↗	↗
<i>i (VC)</i>	↘	→	↘
<i>g (SV)</i>	↗	→	↗
<i>h (VS)</i>	↘	→	→

Tabela I25 – tendências gerais de comportamento das ordens mais freqüentes segundo a vogal temática do verbo.

O cruzamento dos fatores *conjugação verbal* e *ambiente oracional* mostra que, tanto para a primeira quanto para a segunda conjugações, as curvas que representam orações principais e subordinadas têm semelhanças e diferenças: ambas correspondem a arcos parabólicos, mas a curva das subordinadas exhibe concavidade para baixo, e as principais, para cima. Em outras palavras, o aumento das proporções de uma classe corresponde ao decréscimo das freqüências absolutas da outra.

É interessante notar, nos dois gráficos seguintes, que o mínimo da curva **1** (para I24a, inferido no século XIV) precede o máximo da curva **3**. Isso significa que os comportamentos dos dois subconjuntos de ocorrências estão relacionados, de maneira que a mudança na tendência de **1** é acompanhada de movimento contrário em **3** – ou seja, as freqüências absolutas de *_a_* e de *_e_* em orações principais começam a aumentar antes que as freqüências em subordinadas comecem a cair. Apesar de se terem curvas definidas¹⁶, a existência de picos esparsos (como o de subordinadas nos séculos XVII e XVIII nos dois gráficos) mostra que há influência das freqüências do ambiente oracional **1** sobre as freqüências de **3**, mas que elas não são compensadas uma pela outra; ou seja, o aumento das proporções de orações principais não se dá às custas de uma diminuição de subordinadas.

Também merece destaque o fato de o máximo da curva de subordinadas para verbos de primeira conjugação estar localizado antes do máximo correspondente para a segunda conjugação. Se se aceita que as orações principais em I26a têm um mínimo fora do intervalo de amostragem, a mesma observação pode ser estendida para essas orações. O cruzamento dos dois critérios sugere, portanto, que as tendências à mudança nas freqüências absolutas das ocorrências esteja associada a dois traços preferenciais:

1. em termos de *ambiente oracional*, os pontos críticos nas curvas de freqüência tendem a surgir primeiro nas orações principais, e só então nas subordinadas.
2. em termos de *conjugação*, as mudanças tendem a surgir em verbos de vogal temática *_a_* antes de chegar à segunda conjugação.

¹⁶ Como se dá, aliás, nos demais gráficos. As freqüências dos dados não evoluem segundo padrões irregulares e não-orientados (Kuramochi & Karipis 2004) – e, se assim fosse, isso colocaria em xeque não apenas a pertinência e a aplicabilidade dos critérios de classificação, como também a própria codificação das ocorrências.

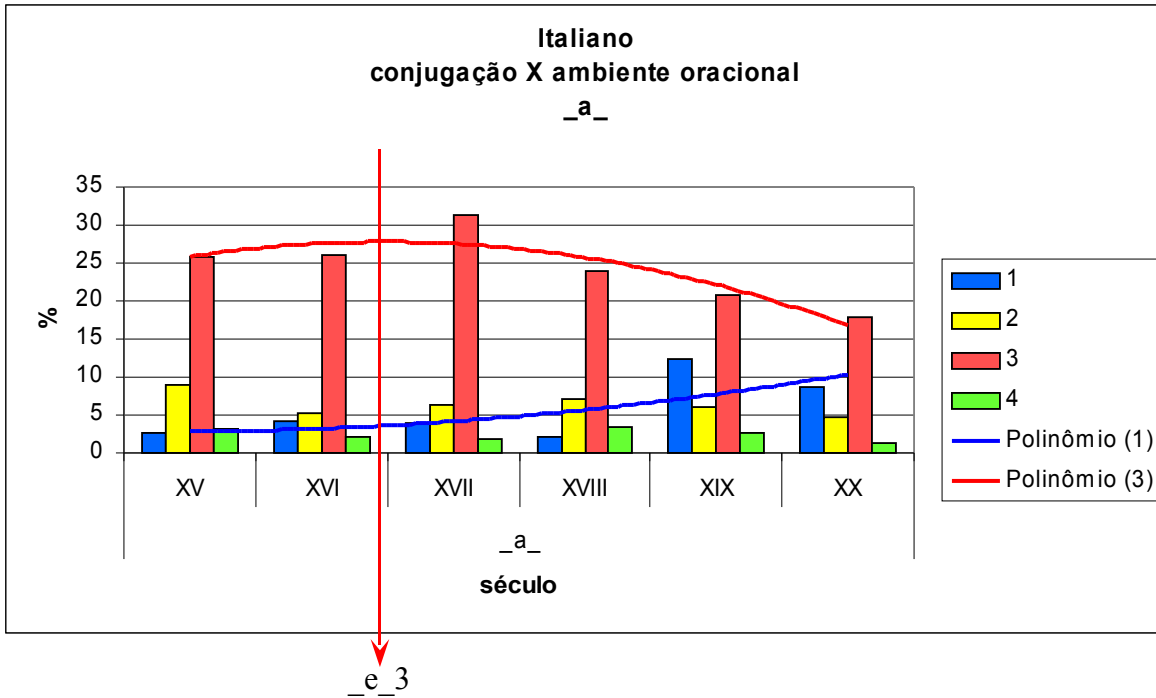


Gráfico I24a – frequências absolutas de ocorrência de verbos de vogal temática **_a_** segundo o *ambiente oracional*.

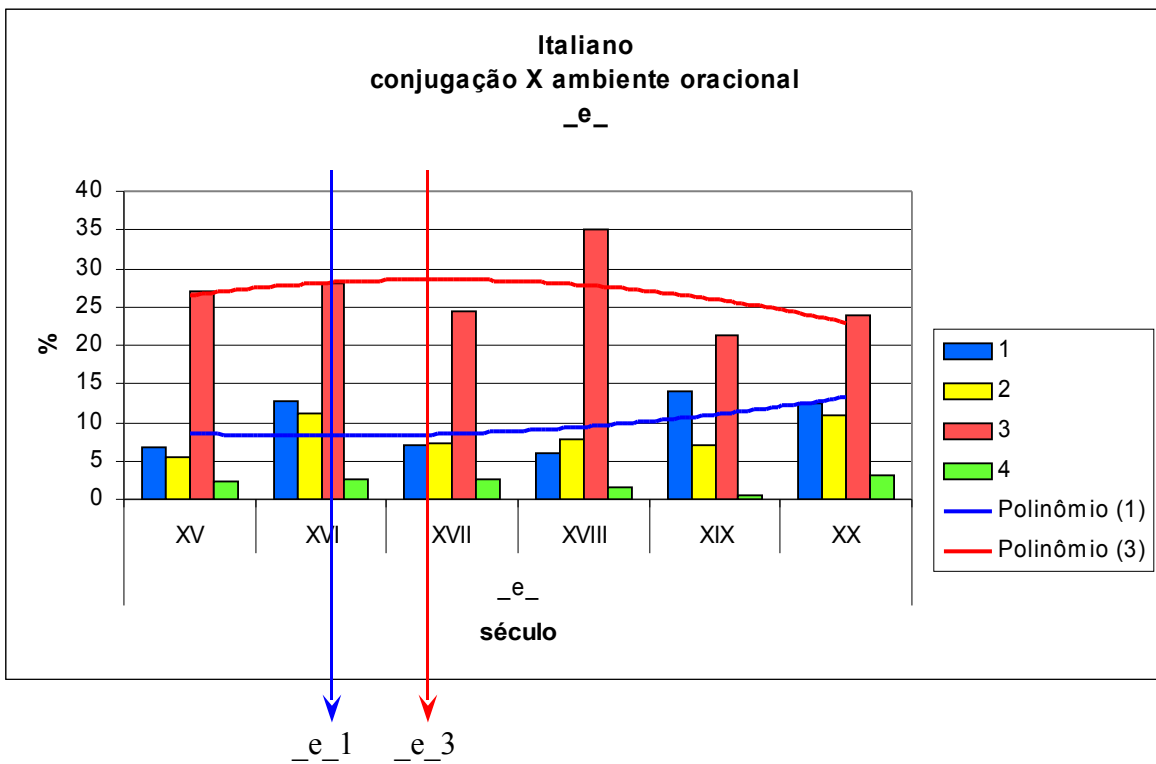


Gráfico I24b – frequências absolutas de ocorrência de verbos de vogal temática **_e_** segundo o *ambiente oracional*.

Chamam também a atenção os resultados obtidos graças à fixação do fator *ambiente oracional*, permitindo-se o grau de liberdade *conjugação*. Para as orações principais (1), as classes mais freqüentes correspondem, como já seria de se esperar, aos verbos de vogais temáticas *_e_* e *_a_*, a exemplo do que já foi observado para os *corpora* português e espanhol acima. Contudo, os máximos das curvas correspondentes aos dados de italiano são praticamente coincidentes, enquanto os mínimos mostram alguma defasagem. Isso quer dizer que, no intervalo de amostragem, a diminuição das freqüências absolutas de verbos de primeira conjugação é imediatamente acompanhada pela segunda conjugação, mas esta última leva algum tempo para responder ao aumento da primeira. As RRLs de inclinação positiva mostram, por fim, que as freqüências de verbos dessas duas conjugações tendem a aumentar em orações principais.

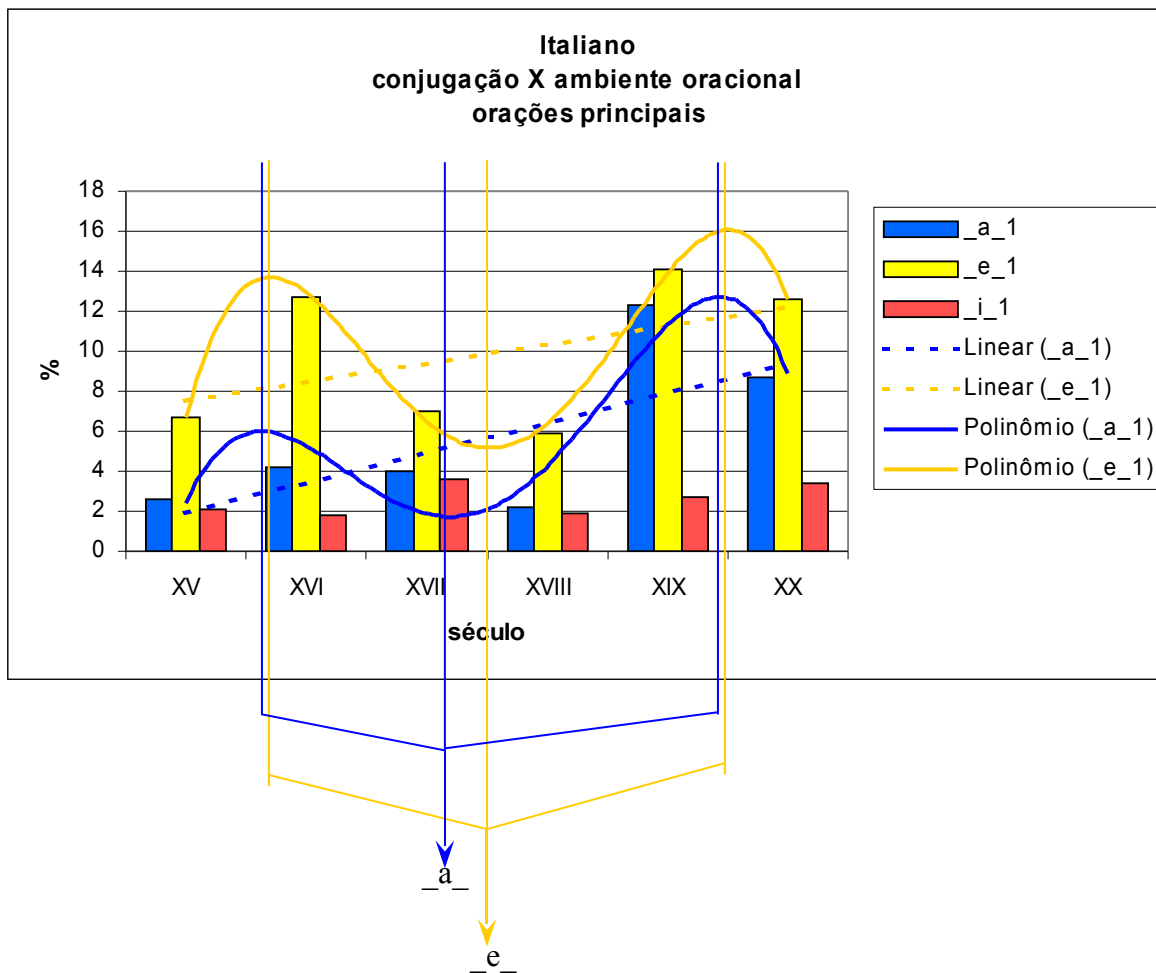


Gráfico I24c – freqüências absolutas de ocorrência das classes de *conjugação* para orações principais.

O mesmo tratamento para as orações subordinadas, o ambiente mais comum no *corpus* italiano, leva a perfis um pouco diferentes para os desdobramentos do fator *conjugação*. Nota-se uma defasagem entre os máximos das curvas de *_a_* e de *_e_*, mas também que a curva correspondente aos verbos de terceira conjugação tem um período menor que o das duas primeiras; o mínimo local dessa mesma curva *_i_* está situado entre os máximos das duas outras curvas. Mais uma vez, entretanto, não há relação de dependência exclusiva entre as freqüências dessas classes: o aumento de uma delas não é função direta da freqüência de outra, mas também, certamente, das freqüências dessas mesmas classes em outros ambientes e segundo outros critérios de classificação das ocorrências em seus contextos.

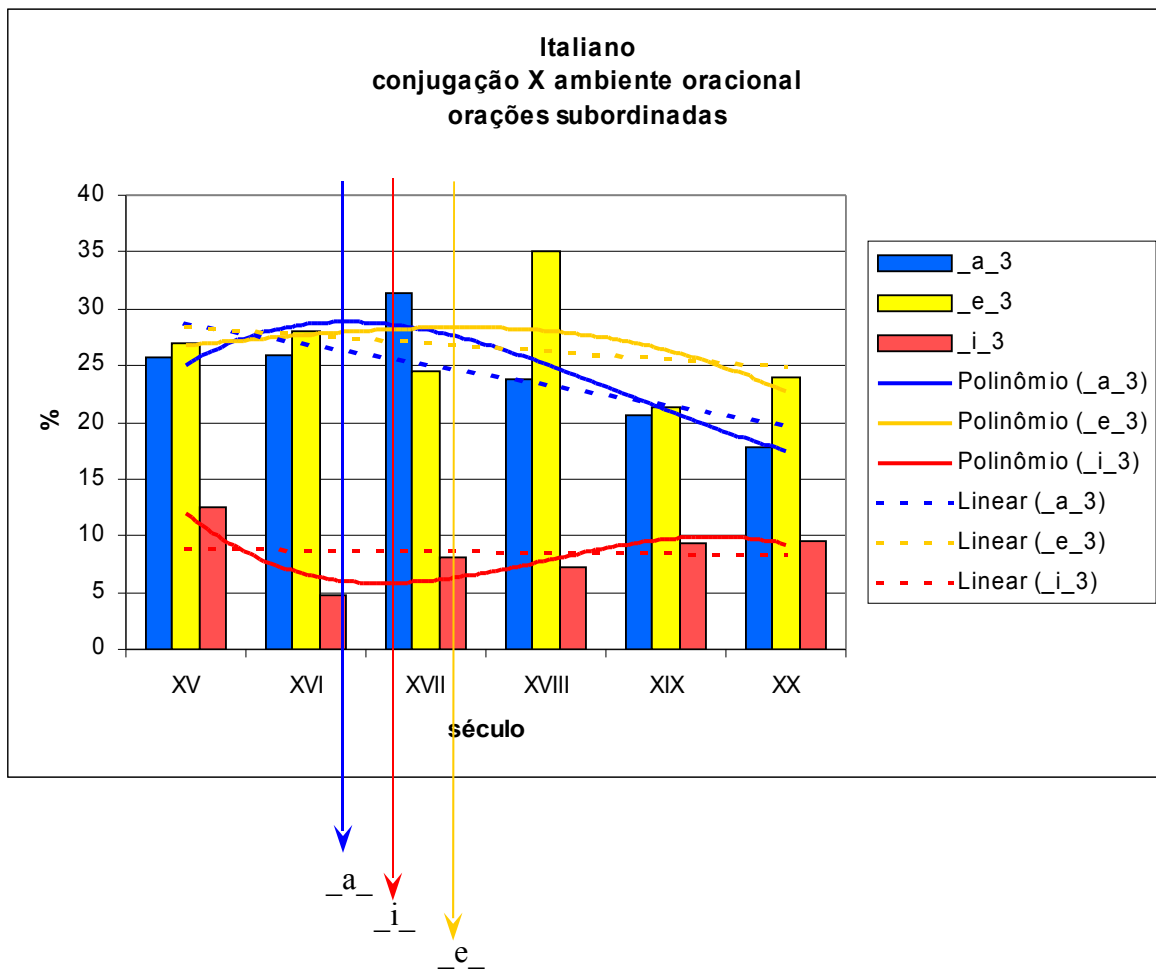


Gráfico I24d – freqüências absolutas de ocorrência das classes de *conjugação* para orações subordinadas.

O cruzamento dos fatores *tempo do verbo* e *ordem de constituintes* oferece outras informações sobre a evolução da distribuição das ocorrências. Assim é que, para a classe α (formas verbais finitas), observa-se a especialização dos dados em torno das ordens **a** (SVC, iRRL = 3.94), **i** (VC, iRRL = -1.30), **g** (SV, iRRL = .06) e **h** (VS, iRRL = -.03). Seguem exemplos dessa classe e o gráfico que sumariza o comportamento das ocorrências no *corpus* italiano.

a (SVC) – ma etia{m} desiderosi che lei habbia tâto triumpho / et - ExposRever XV

i (VC) – non sò : ma m'imagino le insegne della Giustitia, e - ViagPellegr XVII

g (SV) – Ma , per Catone , un mezzo verso basta a Virgilio - PrincLettere XVIII

h (VS) – di ciascuna galera si vide spenzolare un uomo appiccato, e - Campanella XIX

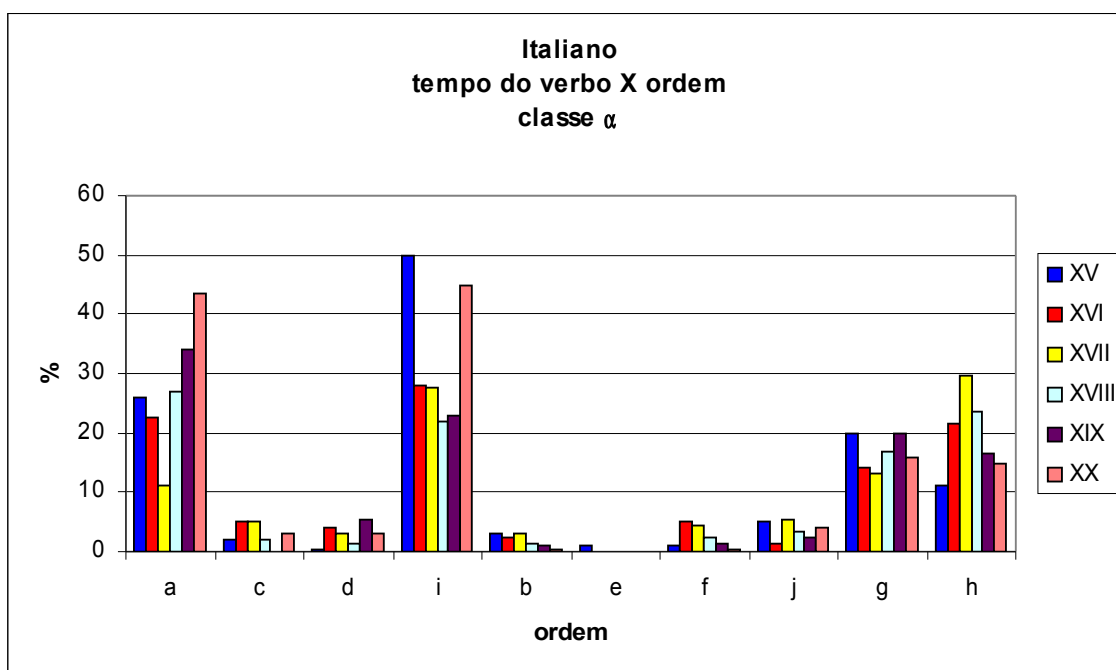


Gráfico I25a – freqüências absolutas de ocorrência das classes de *ordem de constituintes* para formas verbais finitas.

As formas verbais reduzidas (classe β) mostram uma franca especialização em torno da ordem **i** (VC), mas a inclinação da reta de regressão linear desse conjunto de dados é negativa (-.37), o que indica tendência à diminuição das freqüências absolutas do conjunto. A ordem **h** (VS), que mal chega aos 5% de freqüência absoluta, também tem iRRL negativa (-.81). Para ambas, a posição pré-verbal é vazia, mas β se realiza preferencialmente associada à posição pós-verbal preenchida pelo complemento do verbo. Passemos aos exemplos e ao gráfico que sumariza o comportamento desse subconjunto de dados:

i (VC) – sedurre ad assoggettire la sua pubblica libertà all' ambizione - Vico XVIII

h (VS) – di S. Giacomo, venendo poi anche il De Lellis pel servizio - Campanella XIX

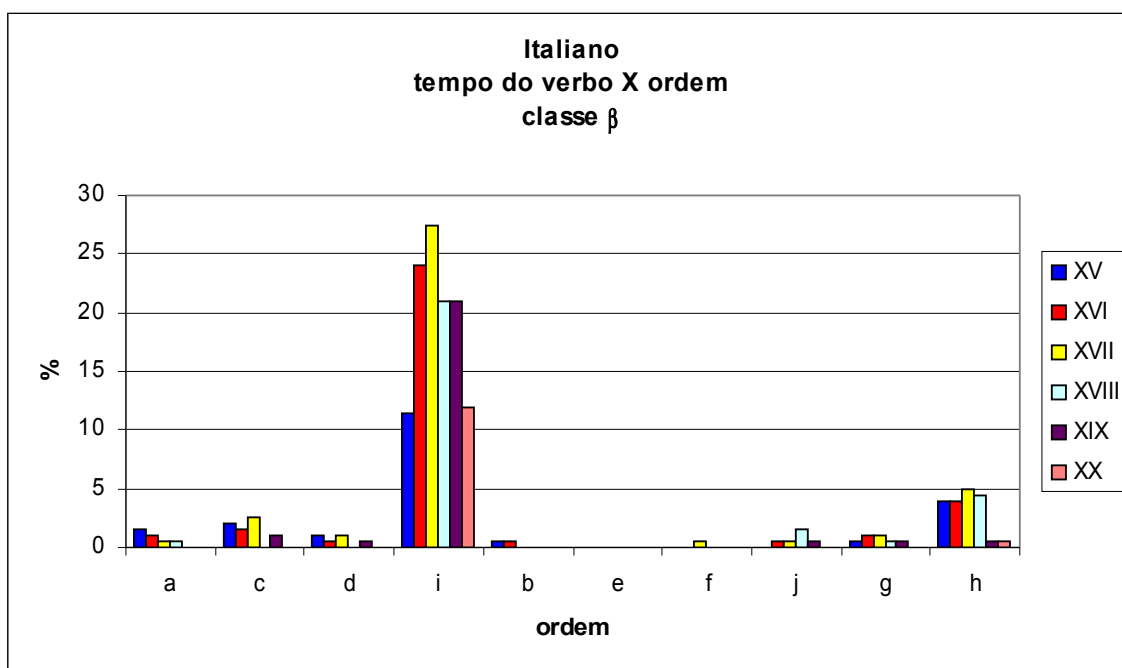


Gráfico I25b – frequências absolutas de ocorrência das classes de *ordem de constituintes* para formas verbais reduzidas.

No caso da frequência das locuções verbais (classe γ), as projeções máximas são mais reduzidas, mas as ocorrências são mais distribuídas – notadamente pelas classes **a** (SVC, iRRL = -.14), **i** (VC, iRRL = -.17), **g** (SV, iRRL = -.10) e **h** (VS, iRRL = .22), que já marcaram outros gráficos nesta análise (q.v. Gráficos I21d e I23c). As ocorrências associadas a esse traço γ são exemplificadas abaixo:

a (SVC) - essendo la Virgine andata aduisitare sancta Helisabeth - ExposRever XV

i (VC) - Nome ; viene a significare in vn certe modo alcuna cosa - LingTosc XVI

g (SV) - in Ispagna, e nell'Indie, e gli altri due potero... accertarsi - Christoforo XVII

h (VS) - "Oh povero Adso," rise Guglielmo dandomi un affettuoso sciaffetto - Eco XX

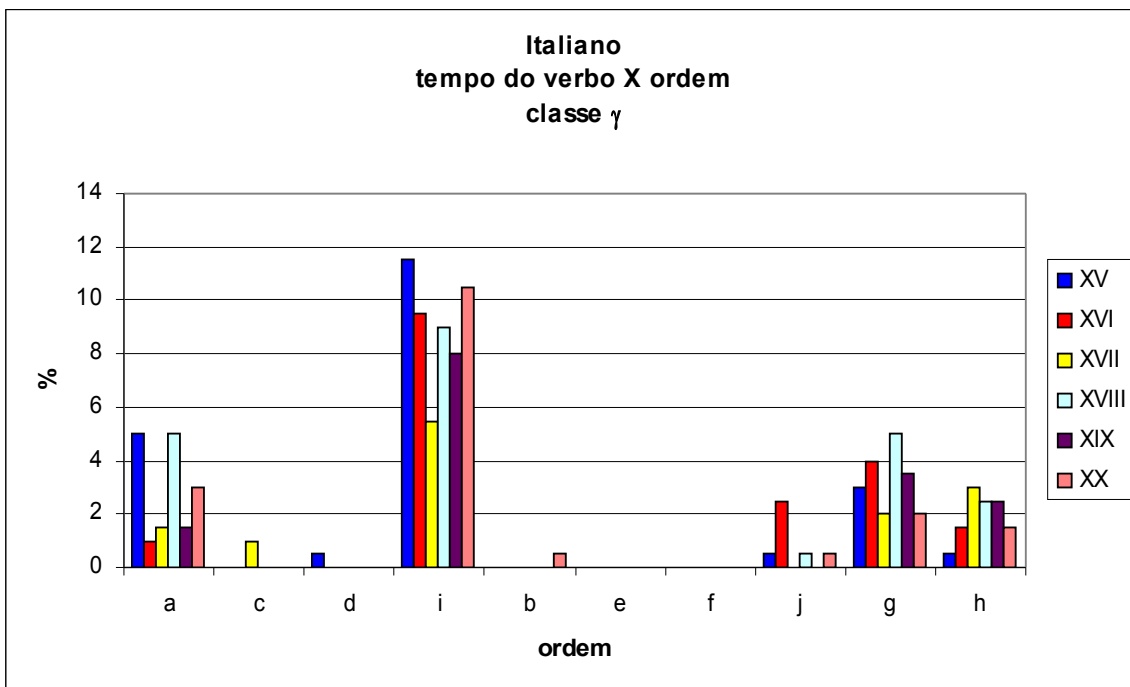


Gráfico I25c – frequências absolutas de ocorrência das classes de *ordem de constituintes* para locuções verbais.

A classe δ (tempos verbais compostos), por fim, é a menos freqüente no *corpus* italiano, e a extensão máxima da escala do gráfico I25d não chega aos 10% de expressividade. Nota-se, contudo, uma distribuição semelhante àquela apresentada pelo subconjunto γ (Gráf. I25c), e segundo a qual as ordens **a** (SVC, iRRL = .24), **i** (VC, iRRL = .96), **g** (SV, iRRL = .59) e **h** (VS, iRRL = .02) são as mais comuns.

a (SVC) - se i settatori di un movimento hanno commesso un delitto - Eco XX

i (VC) - E p lo suo mutame[nto] hauia lassato a frati largo spacio p fare - Dyalogo XV

g (SV) - bãchette per sedere,& il pauimento fu fatto dell'anno 1524 - CompRegole XVI

h (VS) - segno è ancora... l'hauer omesso la Republica di ergere - Christoforo XVII

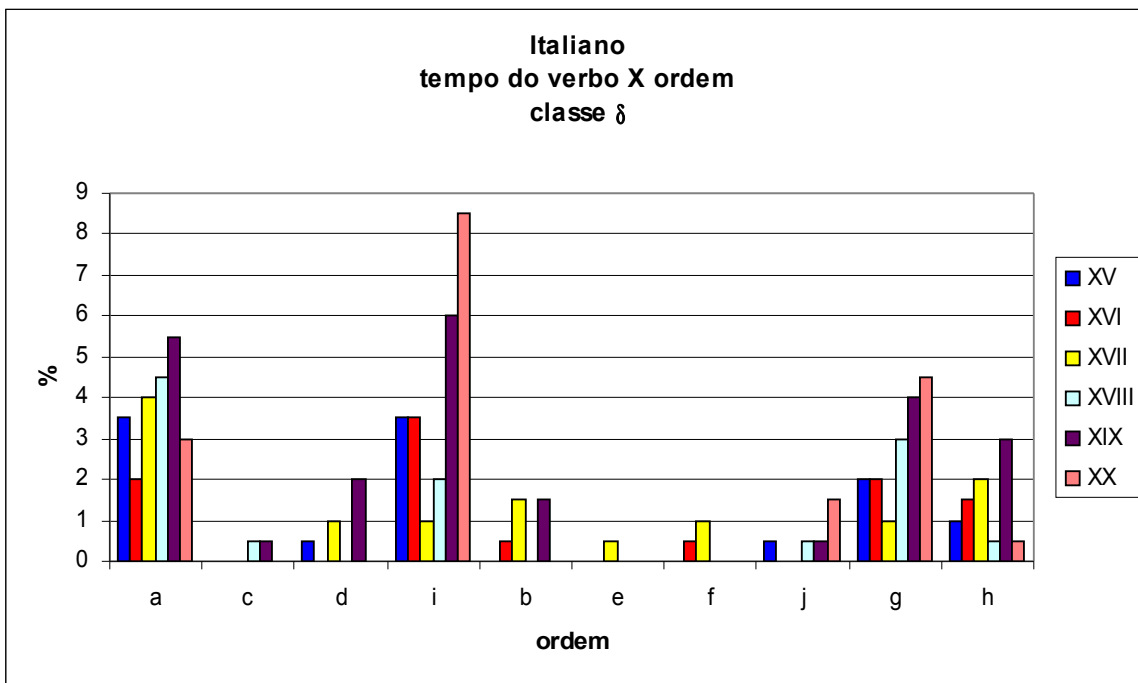


Gráfico I25d – freqüências absolutas de ocorrência das classes de *ordem de constituintes* para tempos verbais compostos.

Outra possibilidade combinatória é dada pela associação entre os fatores *tempo do verbo* e *ambiente oracional*; os perfis gerais indicam, entretanto e uma vez mais, a diminuição das freqüências absolutas de ocorrência de todos os traços ligados ao ambiente 3 (orações subordinadas, iRRL = -1.11). É o que se passa, inicialmente, com o subconjunto α (formas verbais finitas), destacando-se também o aumento das proporções relacionadas ao ambiente 1 (orações principais, iRRL = 1.82); os exemplos são os seguintes:

- 1 - Disse Larchangelo il Signore : pero che lui e / Signore dogni cosa - ExposRever XV
- 2 - Ma se vn'accusativo richiede; Accusatiuo sarà la perticella senz'altro - LingTosc XVI
- 3 - quando vide i Giannizzeri con quelle accette e zappe dietro alla - ViagPellegr XVII
- 4 - che prima furono monarchici , ... e finalmente vennero gli aristocratici - Vico XVIII

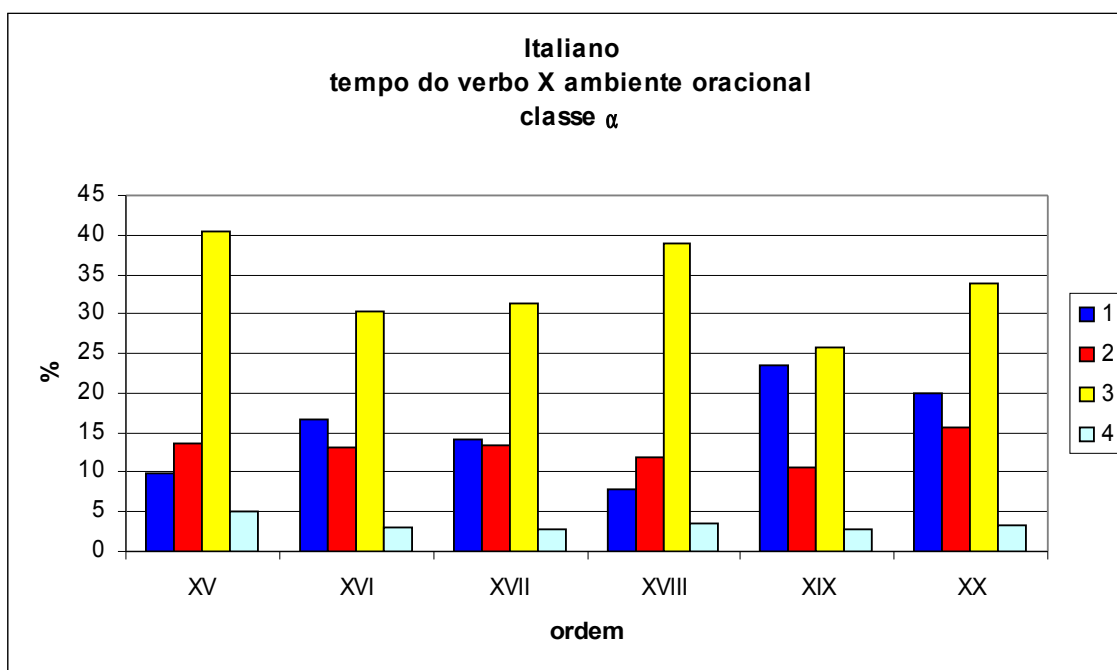


Gráfico I26a – frequências absolutas de ocorrência das classes de *ambiente oracional* para formas verbais finitas.

As formas verbais reduzidas (β), apesar de ocorrerem segundo proporções mais modestas que as formas finitas, mostram a mesma tendência ao decréscimo das freqüências associadas às orações subordinadas (3, iRRL = -1.21). Observa-se no gráfico abaixo que essa é praticamente a única ordem em que se realiza a classe β , de maneira que as demais freqüentem o *corpus* italiano em proporções extremamente baixas. Tomem-se os exemplos associados a esse traço:

- 1 – I Latini, auêdo riguardo a questo suo naturale vficio lo dissero pronome - LingTosc XVI
- 2 – & a far viaggi lunghissimi, e pericolosissimi insieme - Christoforo XVII
- 3 – se, ad illuminare le fitte tenebre in cui trovasi ravalta la storia - GeogrEtnStor XIX
- 4 – che voleva andare à casa del diavolo, et ivi aspettare don loyse sciarava - Campanella XIX

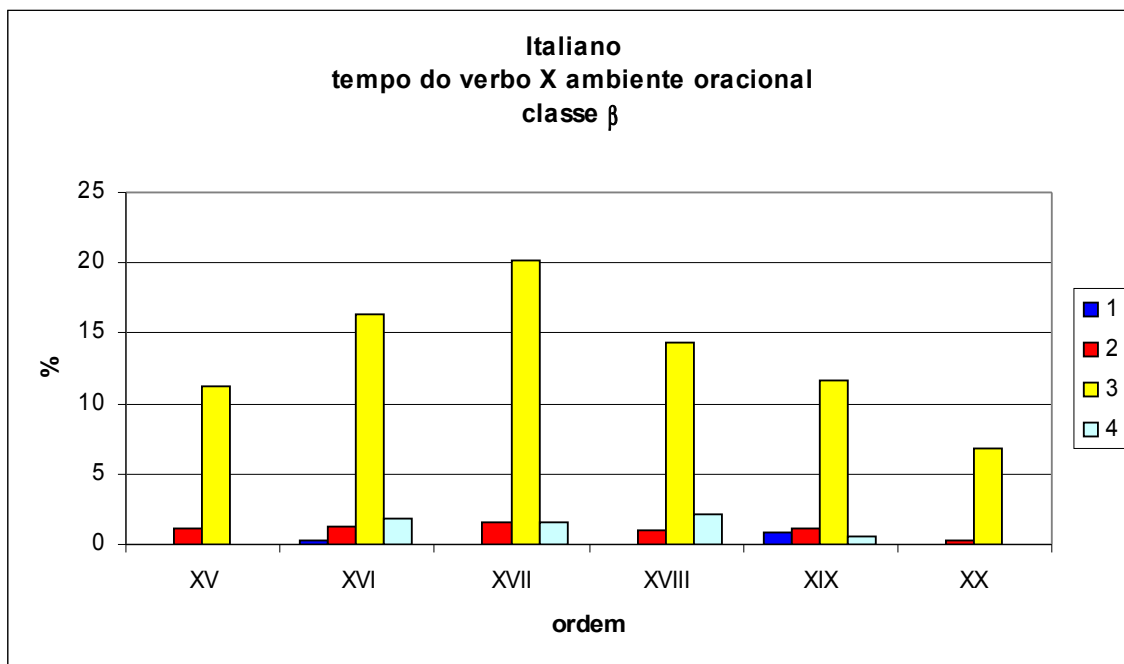


Gráfico I26b – freqüências absolutas de ocorrência das classes de *ambiente oracional* para formas verbais reduzidas.

As ocorrências associadas ao traço γ (locuções verbais) apresentam igualmente tendência a se organizarem em orações subordinadas, motivo pelo qual as projeções da classe **3** são mais elevadas que as demais. Não obstante, esse conjunto apresenta o mesmo caráter descendente (iRRL = -.62) já assinalado para outros grupos de dados ligados a esse ambiente oracional. Tomem-se os exemplos:

- 1 – Ora ti voglio dire alchuna cosa piu di presso : la quale - Dyalogo XV
- 2 – al suo paese;non se ne puo hauere notitia cosi facilmente - CompRegole XVI
- 3 – della natura , più forse quale ella potrebbe essere, che - PrincLettere XVIII
- 4 – era color smeraldo ed egli, diceva, non voleva vedere le pergamene - Eco XX

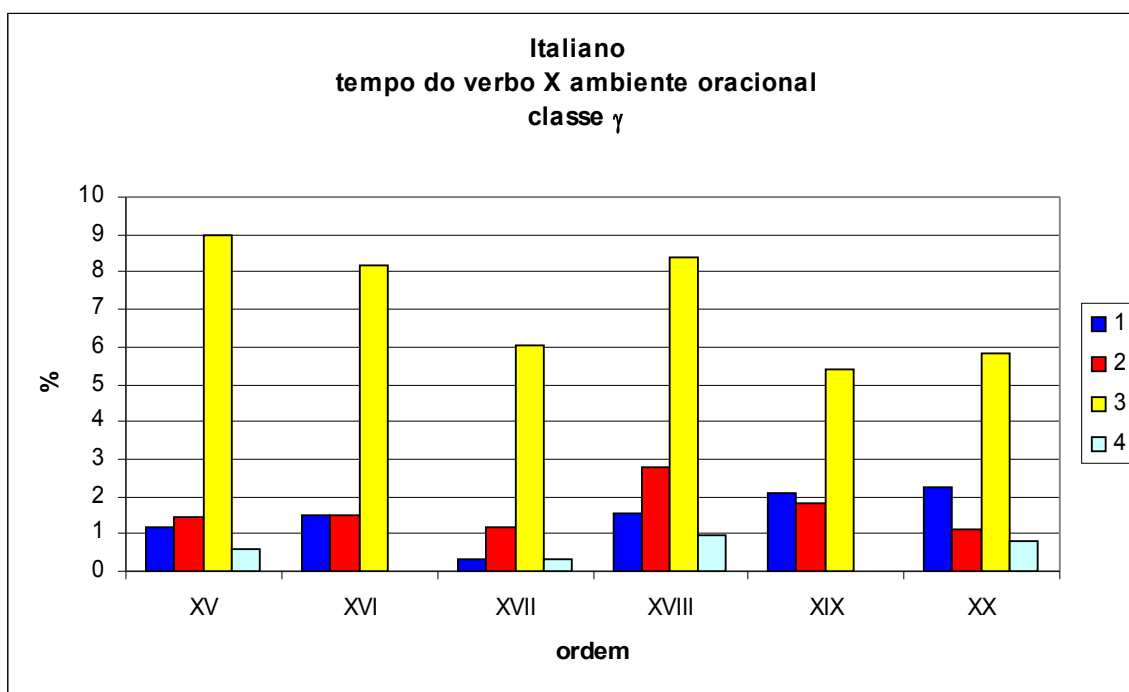


Gráfico I26c – frequências absolutas de ocorrência das classes de *ambiente oracional* para locuções verbais.

Os tempos compostos (classe δ de *tempo verbal*) ocorrem em proporções muito reduzidas no *corpus* italiano, o que torna difícil tecer comentários seguros sobre o comportamento desse subconjunto de dados. Por essa razão, não tratarei desses dados.

A distribuição das ocorrências segundo o critério *classe semântica do verbo* mostra uma concentração dos dados em torno de quatro categorias principais, ATTR, COGN, DISP e SEM. Em decorrência da dispersão encontrada para a amostra, as características individuais dos textos podem resultar em desvio-padrão mais elevado para uma classe qualquer. Tomem-se como exemplos os picos apresentados para as classes ATTR, DISP e, possivelmente, SEM nos séculos XV e XX: se os valores obtidos para as médias de frequências de DISP(VC) estiverem exagerados por causa do efeito provocado por um dos textos do século XV, a iRRL corrigida para esse subconjunto teria um valor positivo; se o valor exagerado for o do século XX, a iRRL tem valor francamente negativo. Não obstante, perfis como o do subconjunto COGN(VC) e, em menor escala, o de SEM(CV) mostram que há dados de cujo comportamento se pode falar com mais propriedade, já que se tem um perfil claramente delineado que referenda o método. Esses são os motivos pelos quais a análise sugere apenas tendências de comportamento do *corpus* para esse fator; conclusões mais seguras podem ser extraídas a partir do tratamento de uma base de dados mais extensa – da ordem mínima equivalente ao quádruplo do total de ocorrências de italiano utilizado neste trabalho.

Pois bem, no Gráfico I29, às classes mais comuns no *corpus* analisado podem-se atribuir tendências ao decréscimo (iRRL(DISP) = -0,29; iRRL(SEM) = -1,73), à manutenção (iRRL(ATTR) = 0,05 ou ao aumento (iRRL(COGN) = 1,72) nas frequências absolutas. As demais classes ocorrem em proporções muito baixas, além de a maioria delas mostrar iRRLs subhorizontais. Vejam-se alguns exemplos:

ATTR – gratia / lagratia e / uno grandissimo thexoro | una pietra pretiosissima -
ExposRever XV

COGN – Chi vuol conoscere quelche operi il Pronome, cōsideri quãto inelegante -
LingTosc XVI

DISP – fatica, nè spesa,per conseguire ogn'anno vna sì grande, e perpetua rendita -
Christoforo XVII

SEM – per inciso racconta come ha decifrato i segni negromantici lasciati - Eco XX

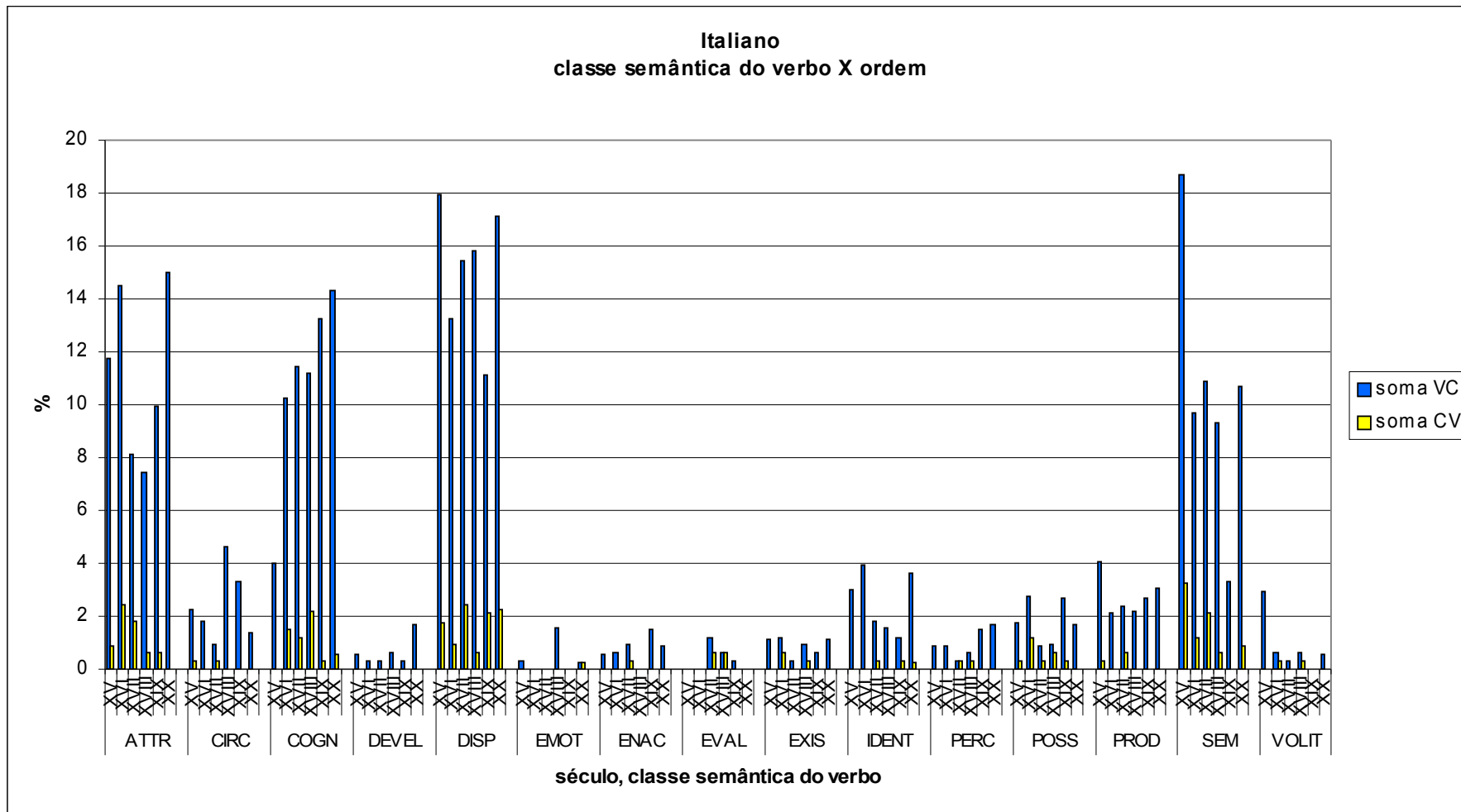


Gráfico I27 – frequências absolutas de ocorrência das classes semânticas do verbo.

Com respeito ao fator *item lexical*, cabe observar que distribuição das frequências em relação ao número de *tokens* para cada entrada lexical acompanha o perfil previsto por Chitashvili & Baayen 1993 (a partir dos estudos de Zipf 1935), como apresentado no Gráf. I28a; nota-se o mesmo perfil obtido para os dados de português (Gráf. P141) e de espanhol (Gráf. E35). Os quatro itens mais frequentes estão entre os cinco mais comuns também no grupo ibérico: *essere* (20% do total bruto de ocorrências), *avere* (6,1%), *dire* (4,3%) e *fare* (3,2%). Para *essere*, a única ordem que apresenta iRRL positiva (0,75) é **a** (SVC), acompanhando a distribuição do Gráf. I2.

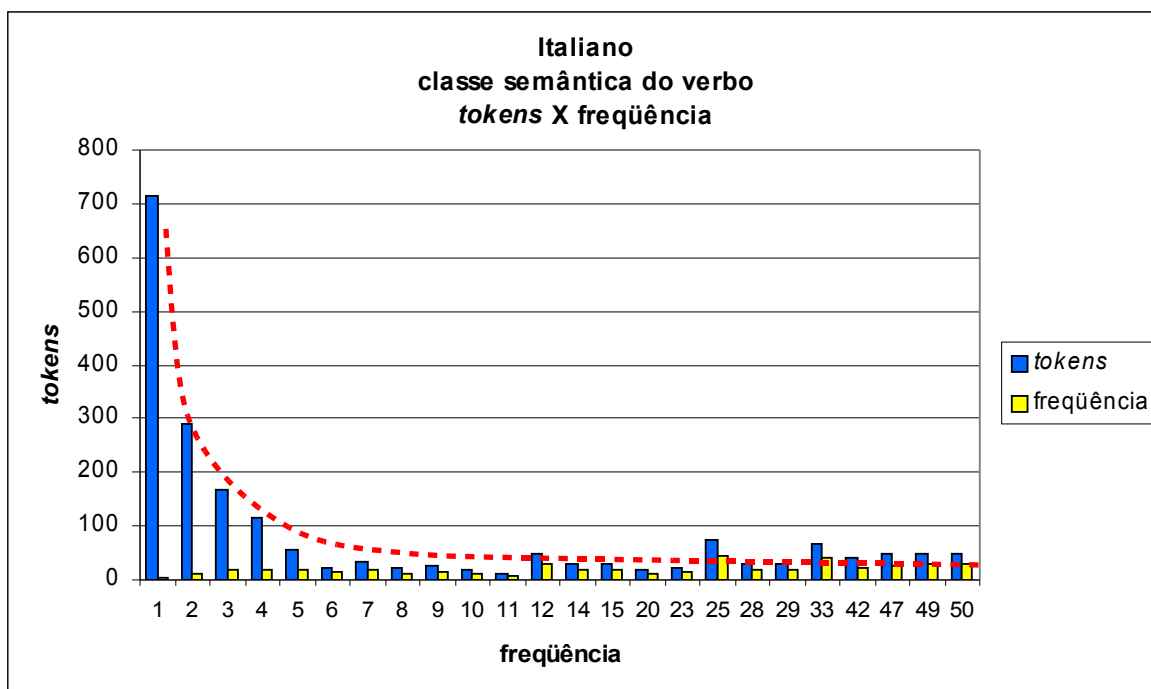


Gráfico I28a – distribuição dos itens lexicais por frequência absoluta.

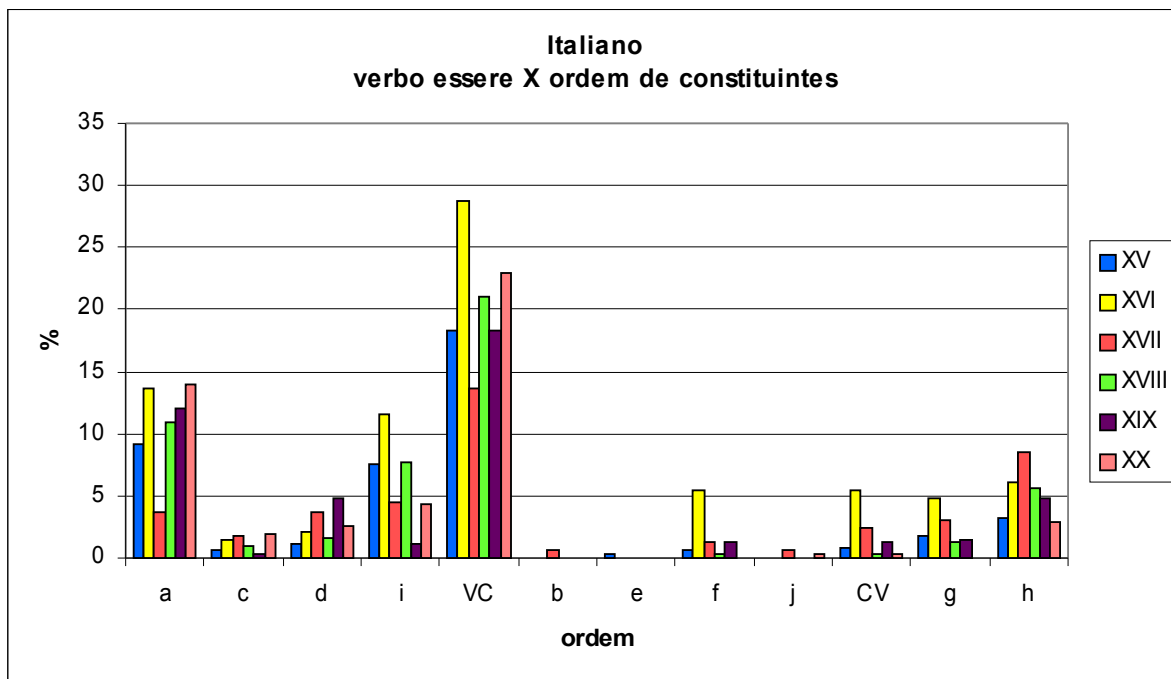


Gráfico I28b – distribuição do verbo *essere* segundo a ordem de constituintes.

Graças à sua elevada frequência no *corpus*, o verbo *essere* também permite observar de maneira mais clara a distribuição dos itens lexicais segundo o fator *ambiente oracional* (Gráf. I28c): repetem-se os padrões do Gráf. I8 acima para as orações principais e subordinadas, indicando que o *ambiente oracional* é determinante, também para o italiano e a exemplo do observado para o português (cf. Gráf. P14m) e para o espanhol (cf. Gráf. E36), da distribuição de frequência dos itens lexicais.

Tomem-se os exemplos associados a *essere*:

a (SVC) – Onde quelle cose che isancti impetrano sono predestinate in tal modo : che - Dyalogo XV

i (VC) – nõ saria stimato caro,essendo in q̄sto modo fertili,& buoni - CompRegole XVI

g (SV) – mà per più certo si tiene, che egli fosse di Cugureo, luogo pur presso la città - Christoforo XVII

h (VS) – ch'erano le due principali popolazioni dell'Africa settentrionale - GeogrEtnStor XIX

1 – Fue vno huomo di sancta vita:che hebbe nome Bonifacio:lo quale - Dyalogo XV

2 – Gli Aggiamoglani sono in gran numero ; & è vn'ordine di militia bassa - ViagPellegr XVII

3 – dalla natural' equità : onde le Monarchie sono le più conformi all' umana natura della più spiegata ragione - Vico XVIII

4 – e che questi Libii biondi sieno Celti, venuti dalla Spagna e dominanti gl'indigeni - GeogrEtnStor XIX

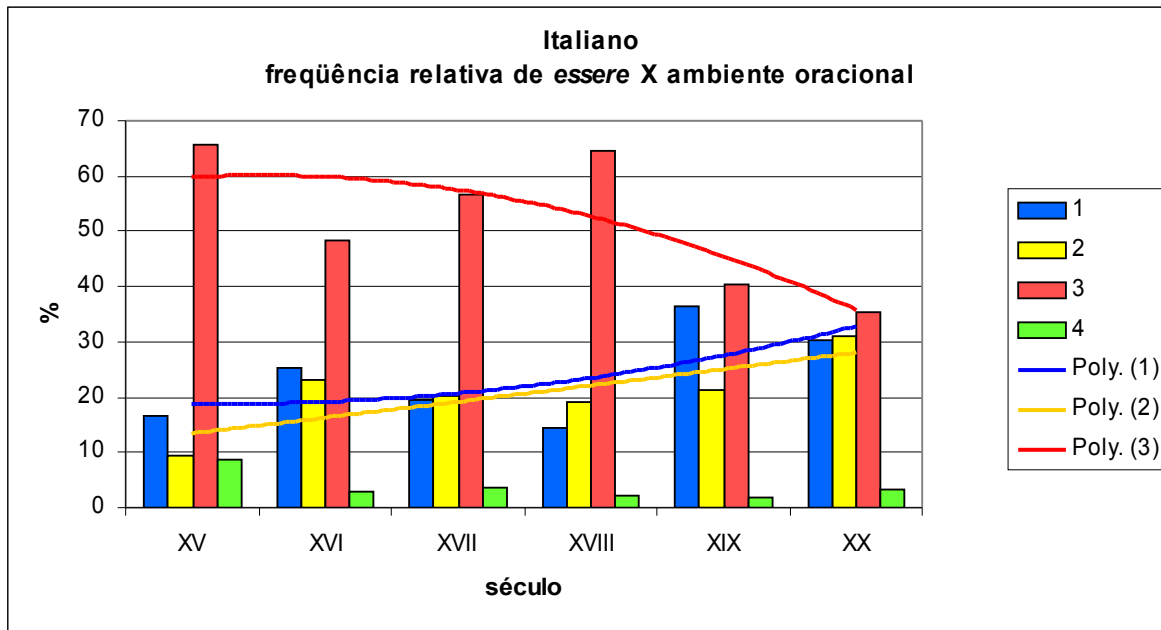


Gráfico I28c – distribuição do verbo *essere* segundo o fator *ambiente oracional*.

O tratamento dos dados do *corpus* italiano fornece, como apresentado acima, resultados consistentes com aqueles obtidos para os conjuntos de dados ibéricos. Observa-se, inicialmente, o paralelismo nos movimentos de AN para NA entre os conjuntos de dados italiano, espanhol (Totaro 1998) e português (Cohen 1986/1990). Em seguida, o exame dos resultados numéricos sumarizados nos gráficos e nas tabelas mostra, também para o italiano e a exemplo do observado acima para o português e para o espanhol, que vários fatores são funções de *ordem de constituintes*: *agentividade* (cf. gráficos I3-I7 e tabelas I2-I6), *peso de S e de C* (cf. gráficos I19a-e e tabela I18), os traços associados ao verbo – *regência* (cf. gráficos I21a-f e tabela I24), *conjugação* (cf. gráficos I23a-c e tabela I25), *tempo* (cf. gráficos I25a-d) e *classe semântica* (cf. gráfico I27) – e *item lexical* (cf. gráfico I28b, traçado a partir das frequências de ocorrência de *essere*). Por sua vez, o fator *ambiente oracional* é função apenas do tempo (cf. gráfico I8), mas controla claramente os comportamentos de *ordem de constituintes* (cf. gráficos I9-I11 e tabelas I7-I11), *peso de S e de C* (cf. gráficos I20a-g), *regência verbal* (cf. gráficos I22a-d), *conjugação verbal* (cf. gráficos I24a-d), *tempo do verbo* (cf. gráficos I26a-c) e *item lexical* (cf. gráfico I28c, também construído segundo as distribuições das frequências de *essere*). Portanto, o fator determinante do comportamento geral dos dados que compõem os *corpora* é *ambiente oracional*.

Vejamos, na próxima seção, a comparação fina entre os resultados obtidos para os três conjuntos de dados.

Comparações entre os *corpora*

Podem-se contrapor os resultados obtidos para os *corpora* individuais a fim de se ressaltarem as semelhanças e as diferenças entre os sistemas ao longo do tempo coberto pela amostragem. Os Gráficos C1 a C13 sumarizam, portanto, as informações apresentadas nas seções de tratamento dos dados, agrupando os valores de iRRL para cada traço nos três sistemas; os Gráficos C14 a C34 comparam os *corpora* situando os pontos críticos das curvas de comportamento dos dados.

Para construir os Gráficos apresentados nesta Seção, tomaram-se os valores de iRRL referentes a determinado fator em cada base de dados, colocando-se-os lado a lado para efeito de comparação. Os valores positivos indicam uma tendência crescente daquele traço no *corpus*, e os negativos apontam a tendência à diminuição da frequência absoluta de um traço; quanto maior a altura de uma barra, maior a inclinação da reta de regressão linear correspondente – e, conseqüentemente, maior a taxa de mudança. Os valores baixos para a iRRL, entretanto, devem ser avaliados diante do gráfico relativo a cada traço, uma vez que as nuances das curvas de comportamento são diluídas pelas próprias características do método de regressão linear.

O primeiro gráfico comparativo é composto das ordens de constituintes mais comuns – **a** (SVC), **i** (VC), **g** (SV) e **h** (VS). Para as duas primeiras, que trazem o complemento verbal expresso, o sistema português tem iRRLs negativas: em **a** (SVC), o espanhol tem uma inclinação muito baixa, como se fosse intermediária entre o português e o italiano; em **i** (VC), a maior perda acontece no português, mas nota-se que os três sistemas mostram decréscimo dessa ordem. É interessante notar os resultados de **g** (SV) e de **h** (VS), que são opostos entre si em todos os sistemas, e que as tendências do espanhol são, em ambos, contrárias às de português e de italiano. Por fim, observe-se que, excetuando-se as ordens [VC], o italiano tem as iRRLs globais mais baixas dentre os *corpora* da amostragem, em decorrência do fato de os valores de iRRL relativos a dessas ordens praticamente se cancelarem diante do fator *ambiente oracional* (cf. Gráfs. I9 – I11).

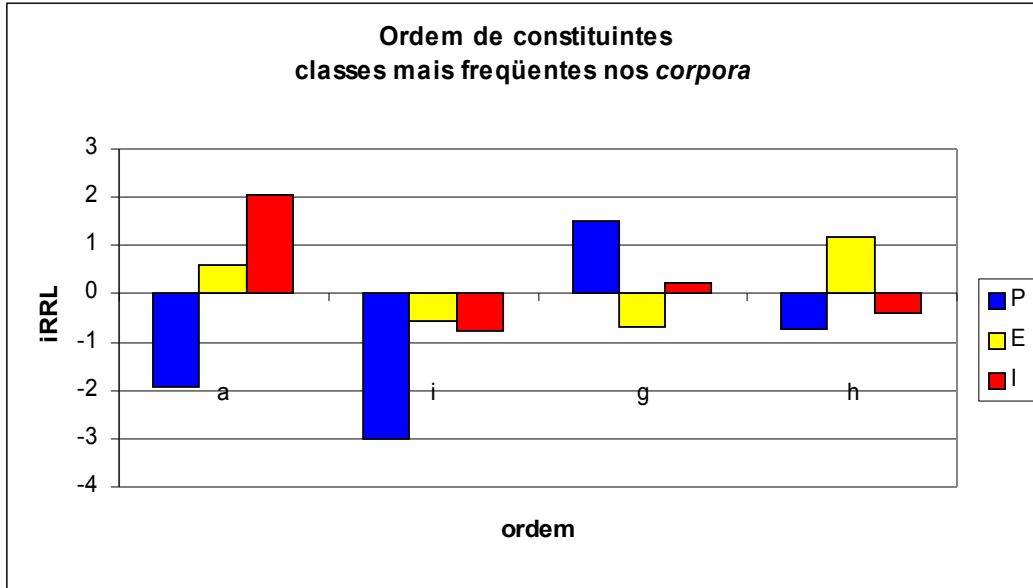


Gráfico C1 – comparação entre as iRRLs das ordens de constituintes mais comuns nos três *corpora*.

Observando-se em detalhe, no Gráfico C2, o comportamento dessas classes de ordenamento de constituintes segundo o ambiente oracional, percebe-se que a iRRL negativa da ordem *a* (SVC) em português é devida principalmente à queda na freqüência absoluta de orações subordinadas; a mesma observação se aplica à ordem *i* (VC). Para o italiano, o valor positivo encontrado para a ordem *a* (SVC) no Gráfico C1 é atribuído ao valor de iRRL em orações principais, cuja freqüência tende a aumentar no *corpus*. Nas demais ordens e ambientes oracionais (e, obviamente, tanto para C2 quanto para C3), os valores relativos ao subconjunto italiano são superados, no mínimo, por um dos subconjuntos ibéricos – o que pode indicar o fato de que estes últimos ainda estão a processar mudanças que já são residuais no primeiro. Essa observação reflete a afirmação de von Wartburg 1966 sobre a mudança dos constituintes intra-sintagmáticos A – N, a qual teria atingido, no italiano, um ponto de equilíbrio em torno de um modelo NA antes do português e do espanhol. Além disso, apesar de o intervalo de amostragem não permitir observar a maior parte da mudança ocorrida no SV antes da que se deu no SN, os dados que levaram à identificação do final do século XV para a mudança AN > NA no italiano (cf. Gráf. II acima) remetem à proposição de que as mudanças se processariam no SV antes de atingirem o interior do sintagma nominal (Greenberg 1966 e Canale 1976).

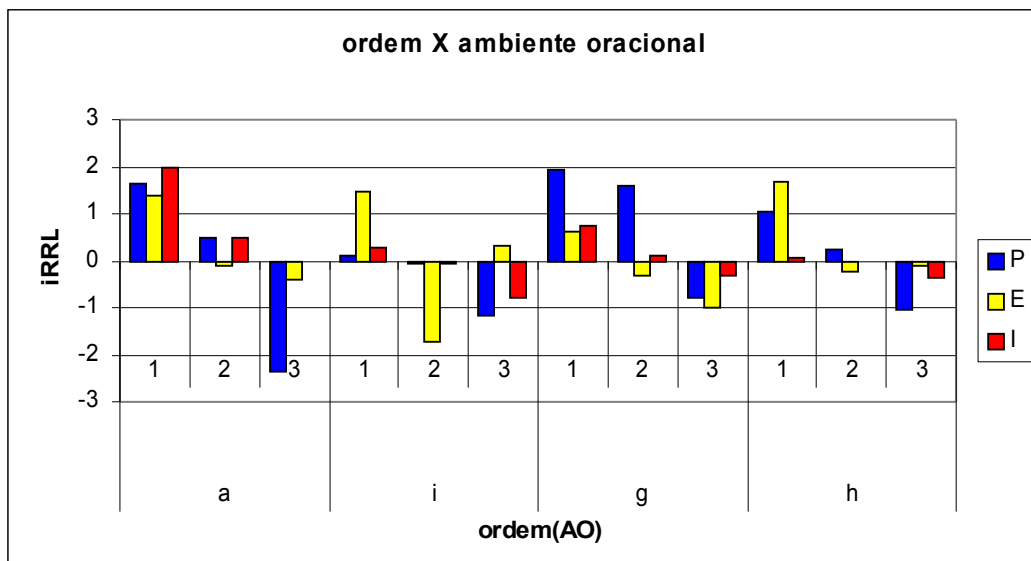


Gráfico C2 – comparação entre as iRRLs das ordens de constituintes mais comuns nos três *corpora* como função do ambiente oracional.

Graças à combinação inversa dos mesmos fatores, pode-se apreciar a relevância de cada ordem de constituintes na composição dos ambientes oracionais nos três *corpora*. Dessa forma, percebe-se facilmente que os três sistemas tendem, como já ressaltado, a sofrer aumento nas proporções de orações principais – mas a) o *corpus* italiano privilegia o acréscimo da ordem **a** (SVC), b) o espanhol lidera o aumento em **i** (VC) e em **h** (VS), e c) o português supera os outros dois sistemas no incremento da ordem **g** (SV). Para as coordenadas (2), veja-se como o espanhol se destaca em **i** (VC) (de valor negativo) e o português, mais uma vez, no aumento de **g** (SV); por fim, para as subordinadas (3), o português tem as taxas mais elevadas de perda nas frequências absolutas das ordens **a** (SVC), **i** (VC) e **h** (VS).

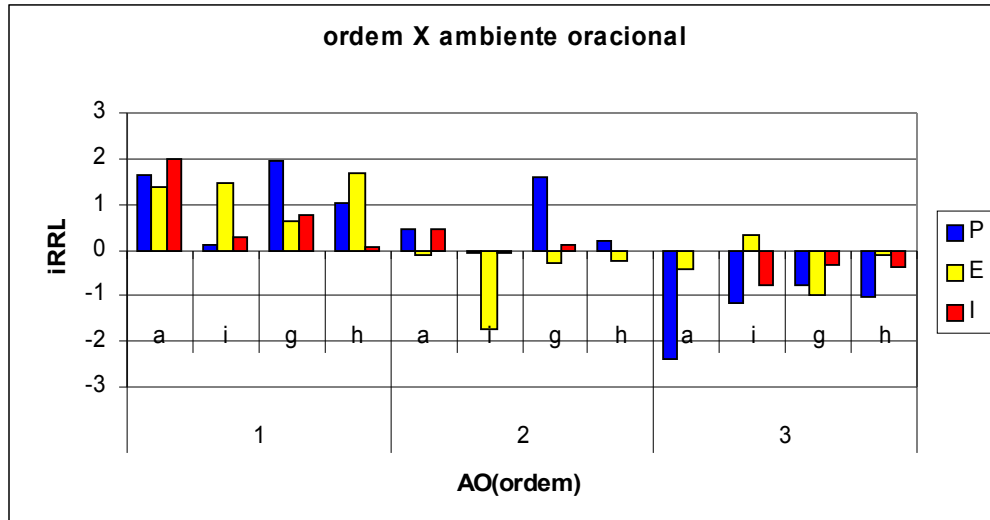


Gráfico C3 – comparação entre as iRRLs dos ambientes oracionais mais comuns nos três *corpora* como função da ordem de constituintes.

Com respeito à agentividade, notam-se comportamentos distintos para os três sistemas: o português é o único sistema que mostra decréscimo mais marcado nas frequências das duas classes mais comuns, Ah e Pn; o espanhol sofre diminuição semelhante apenas nas proporções de Ah, e o italiano apresenta aumento em ambas. Por outro lado, nota-se que o espanhol diminui em Ah, mas aumenta sensivelmente em An.

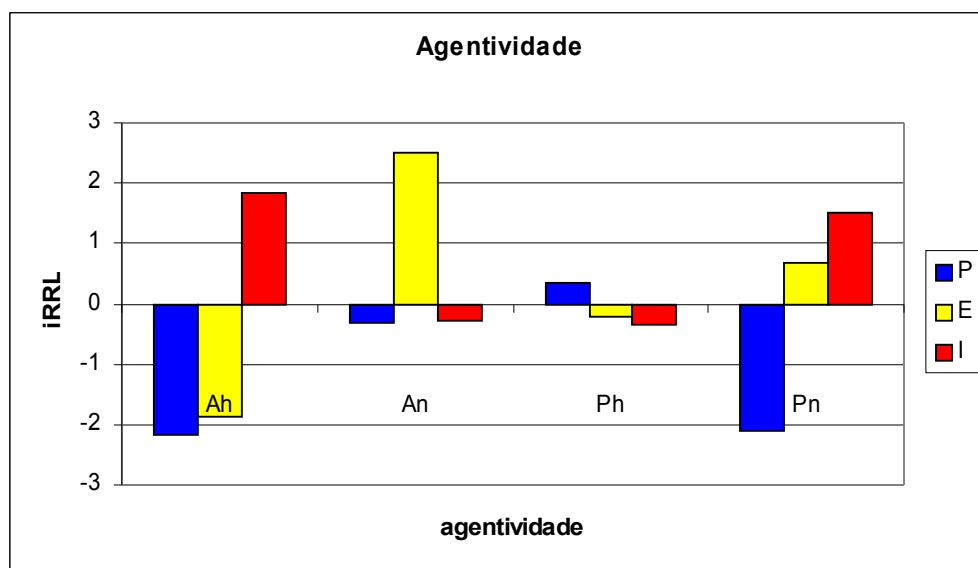


Gráfico C4 – comparação entre as iRRLs das classes de agentividade nos três *corpora*.

Uma observação mais detalhada do comportamento das classes mais comuns de agentividade nos três sistemas mostra, em primeiro lugar, que a diminuição das proporções de Ah em português é devida à contribuição das ocorrências associadas a orações subordinadas (3) – ao passo que, em espanhol, tanto as orações coordenadas (2) quanto as subordinadas (3) são responsáveis por esse decréscimo. No *corpus* espanhol, aliás, essas duas classes mais freqüentes só aumentam em orações principais (1), diminuindo nos demais ambientes; para o italiano, os movimentos de maior amplitude estão relacionados às orações principais (1); por fim, nota-se, também aqui, a tendência negativa dos três subconjuntos de dados no que respeita às orações subordinadas (3).

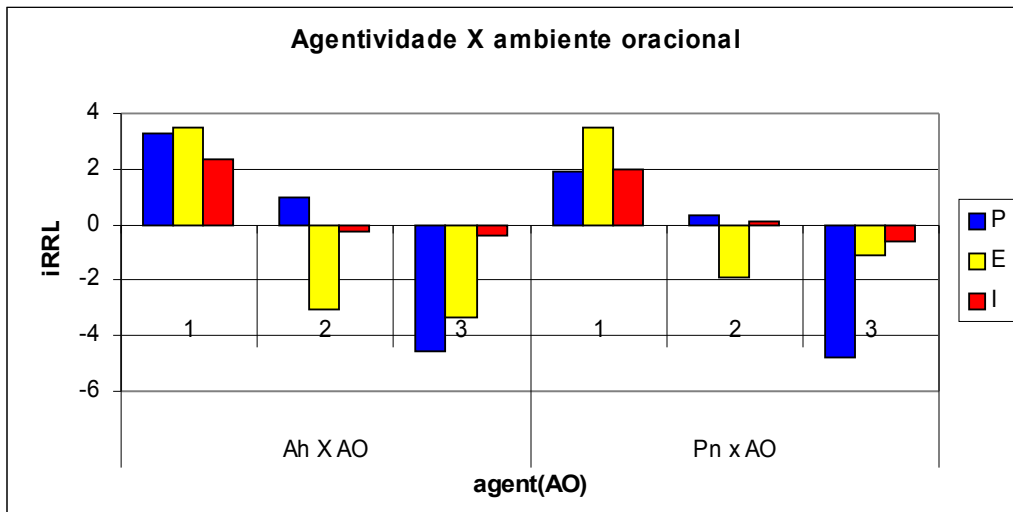


Gráfico C5 – comparação entre as iRRLs das classes de agentividade nos três *corpora* como função do ambiente oracional.

O arranjo dos resultados expostos no gráfico C5 segundo uma disposição diferente mostra que, de fato, o ambiente oracional é determinante das tendências de distribuição das ocorrências nas três bases de dados: as duas classes mais freqüentes de agentividade mostram aumento para P, E e I em orações principais (1); é também flagrante a diminuição em orações subordinadas (3), apesar de, neste caso, o italiano mostrar as inclinações mais baixas e o português, as mais elevadas. Para as orações coordenadas (2), destacam-se os módulos mais altos das iRRLs de Ah e de Pn em espanhol.

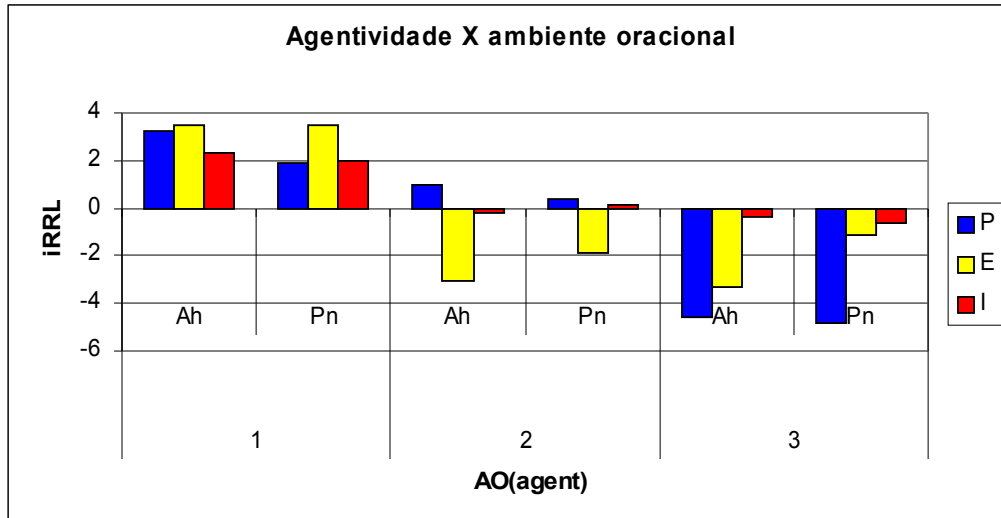


Gráfico C6 – comparação entre as iRRLs dos ambientes oracionais mais comuns nos três *corpora* como função da agentividade.

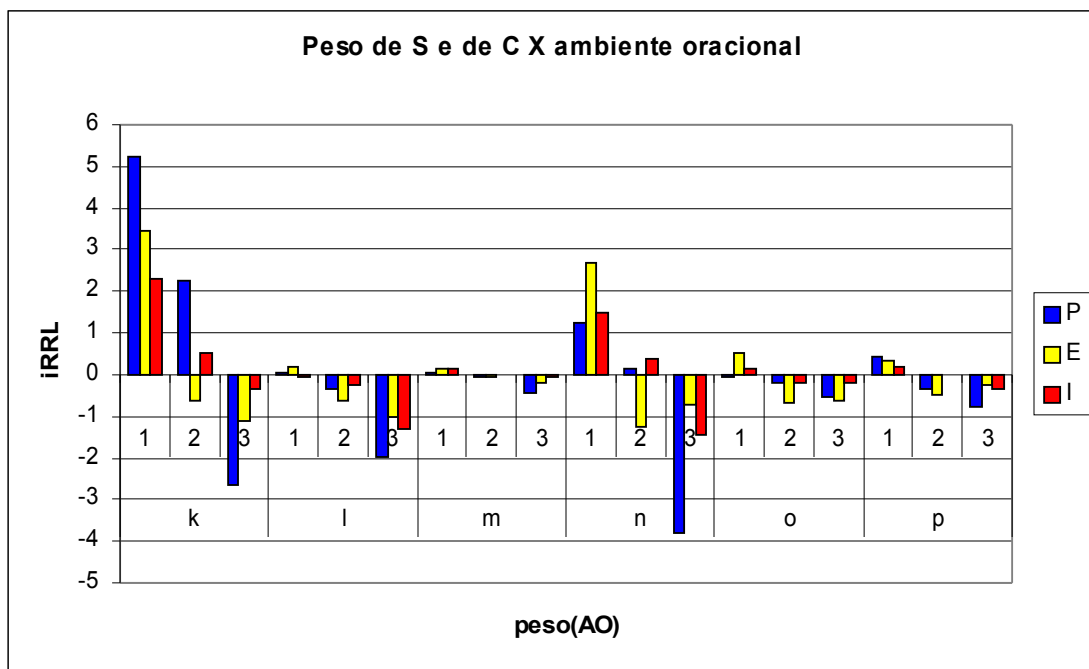


Gráfico C7 – comparação entre as iRRLs das classes de *peso de S e de C* nos três *corpora* como função do ambiente oracional.

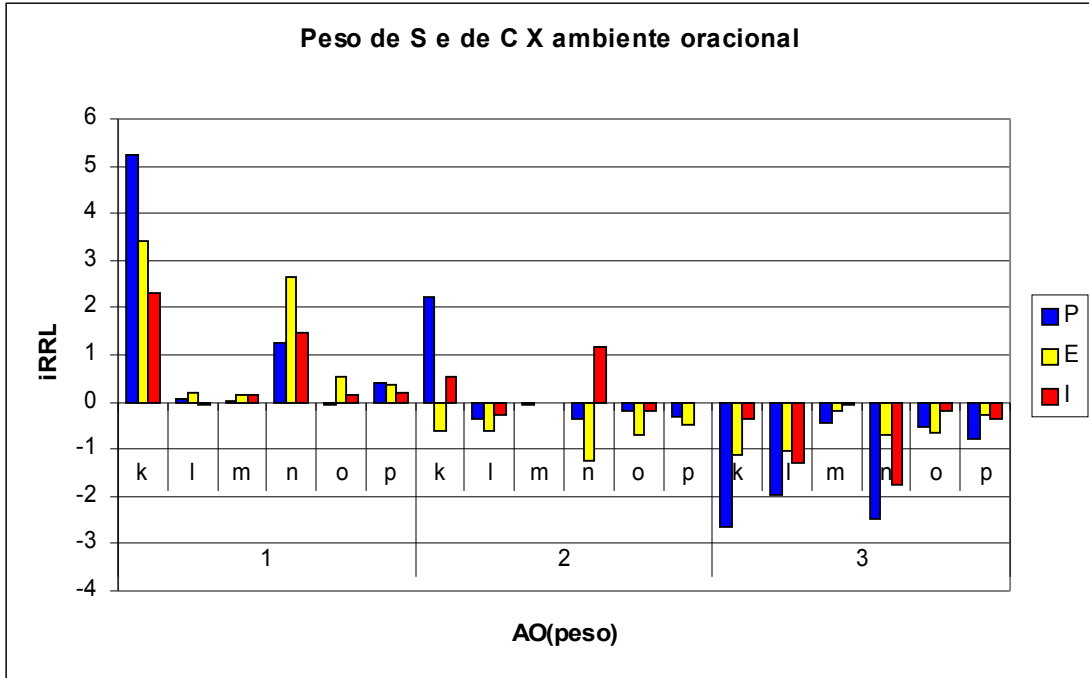


Gráfico C8 – comparação entre as iRRLs das classes de ambiente oracional mais comuns nos três corpora como função do peso de S e de C.

Para o fator *regência verbal*, as categorias mais comuns são, como já exposto, *transitivos diretos (u)*, *verbos de ligação (x)* e *verbos intransitivos (xx)*. Nesses três casos e para os três sistemas, os resultados mostram que os subconjuntos ibéricos têm, em geral, as iRRLs mais elevadas – i.e., P e E apresentam, no intervalo de amostragem, maiores diferenças entre os valores inicial e final das freqüências absolutas desses traços, enquanto o italiano tem valores mais equilibrados para os mesmos. Tanto em C9 quanto em C10 (que traz os mesmos resultados em outro arranjo, favorecendo o ponto-de-vista do ambiente oracional), evidenciam-se as tendências a) de aumento das freqüências absolutas de **u** (trans. diretos), **x** (ligação) e **xx** (intransitivos) em orações principais (1), b) de diminuição das freqüências dessas mesmas classes em orações subordinadas (3), e c) um estado intermediário em orações coordenadas (2).

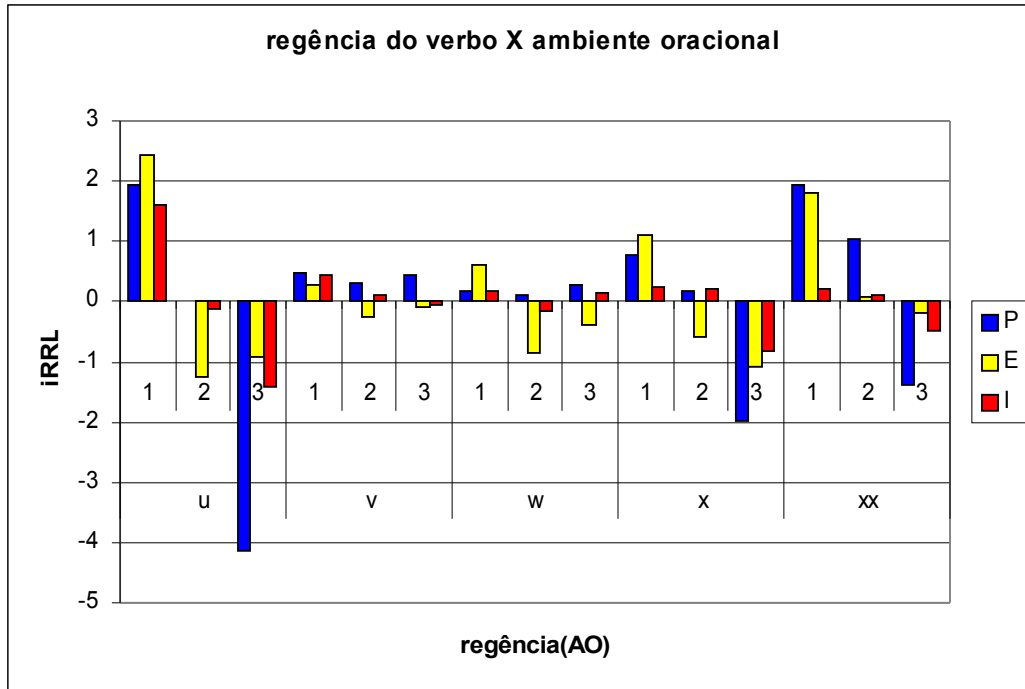


Gráfico C9 – comparação entre as iRRLs das classes de regência verbal nos três *corpora* como função do ambiente oracional.

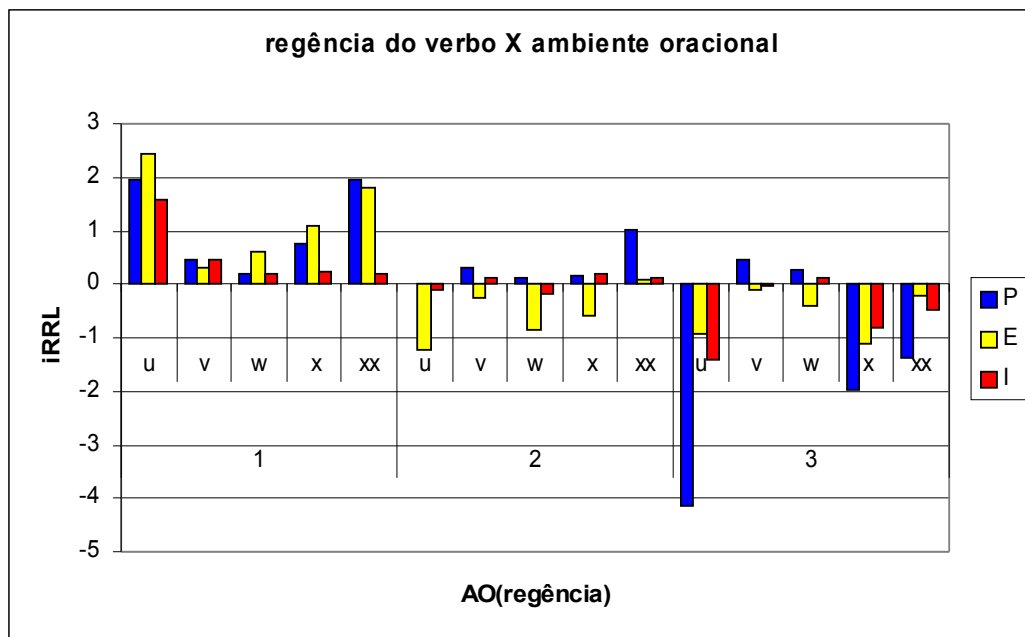


Gráfico C10 – comparação entre as iRRLs das classes de ambiente oracional mais comuns nos três *corpora* como função da regência verbal.

Com respeito ao fator *tempo do verbo*, nota-se, nos Gráficos C11 e C12, que as formas finitas (α) mostram as tendências mais sensíveis à mudança – seja no sentido de aumento das freqüências absolutas em orações principais (1) nos três *corpora*, ou de decréscimo dessas freqüências em orações subordinadas (3). As formas finitas em orações coordenadas ($\alpha 2$) sofrem pequeno aumento em P e diminuição em E, e têm distribuição de iRRL subhorizontal em italiano. Nessa mesma classe α , o par de sistemas ibéricos responde, mais uma vez, pelas mais elevadas taxas de mudança em qualquer ambiente oracional.

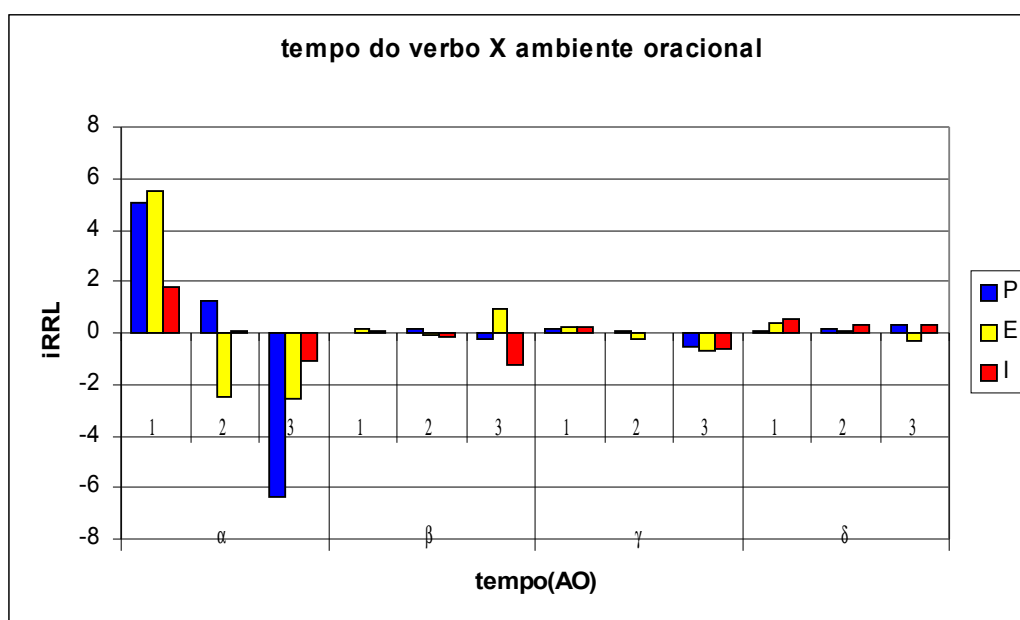


Gráfico C11 – comparação entre as iRRLs das classes de *tempo do verbo* nos três *corpora* como função do ambiente oracional.

Em virtude de as categorias do fator *tempo do verbo* se concentrarem já no Gráfico C11 e de os valores encontrados para as classes β , γ e δ serem pouco expressivos, limitaremos-nos à apresentação desse fator como função de *ambiente oracional*; a inversão dos fatores levaria apenas à dispersão das tendências mais salientes oferecidas no gráfico acima.

Os Gráficos C12 e C13, por fim, mostram que a classe flexional do verbo também é sensível ao ambiente oracional. Assim é que as tendências observadas para as três conjugações são de aumento das frequências absolutas das ocorrências associadas a orações principais (1) e de diminuição das frequências relacionadas a orações subordinadas (3); as coordenadas (2), mais uma vez, oscilam entre esses dois extremos.

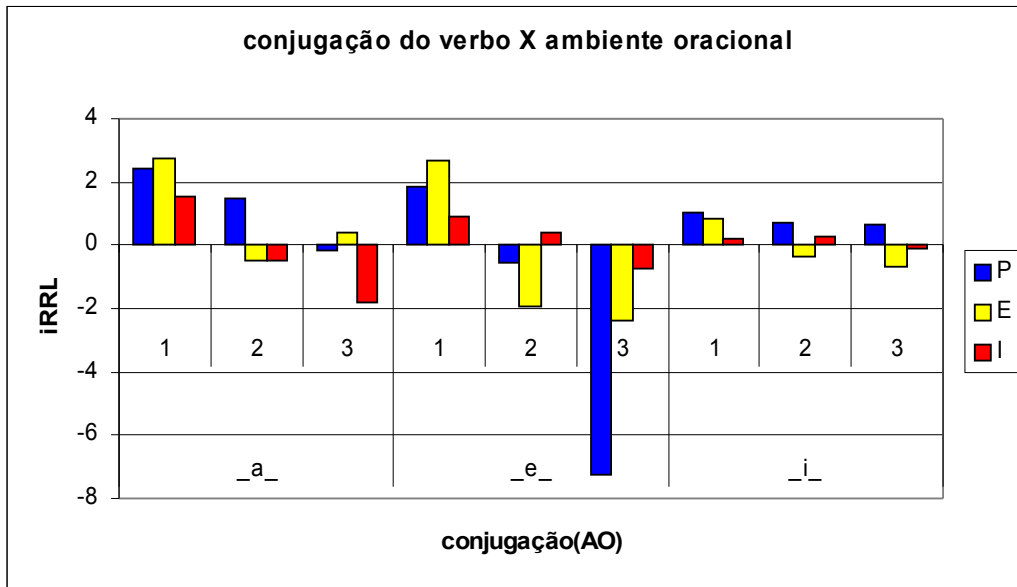


Gráfico C12 – comparação entre as iRRLs das classes flexionais do verbo nos três *corpora* como função do ambiente oracional.

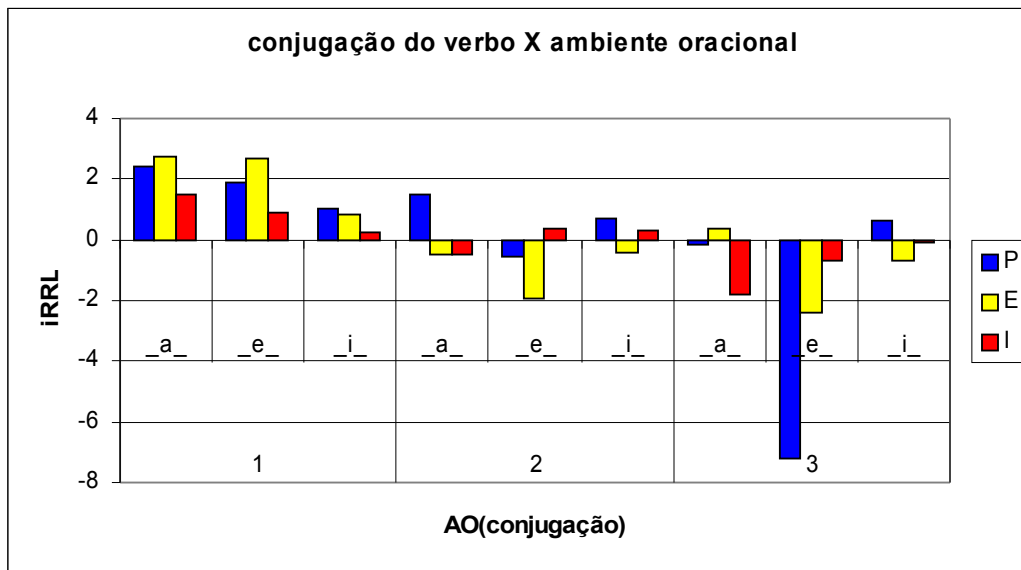
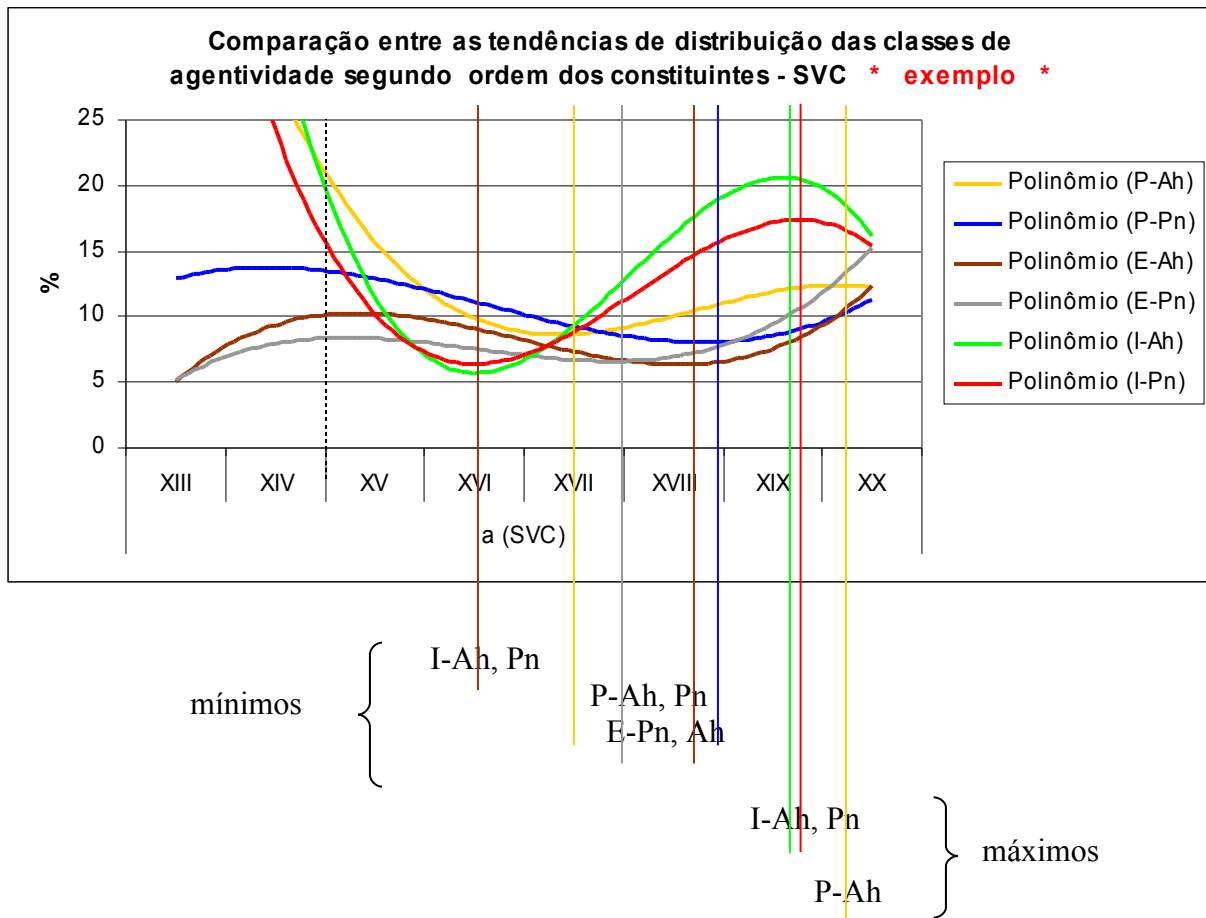


Gráfico C13 – comparação entre as iRRLs das classes de ambiente oracional mais comuns nos três *corpora* como função da classe flexional do verbo.

Outro tipo de comparação entre os *corpora* pode ser feito a partir da sobreposição das curvas de comportamento de cada sistema para um traço específico. Dessa forma, é possível verificar, à mesma escala, as proporções de distribuição e de mudança de frequência de um traço num sistema em relação ao que se passa nos outros *corpora*.

Nos gráficos seguintes, uma linha vertical tracejada no início do século XV indica o ponto a partir do qual devem ser consideradas as curvas de P e de I; à esquerda dessa linha, essas curvas são apenas projetadas pelo mesmo polinômio que se ajusta aos dados, uma vez que este estudo não comporta amostragem anterior ao século XV para esses dois sistemas. Em alguns casos, como ilustrado no exemplo abaixo, podem-se identificar pontos críticos para as curvas, mas, na maioria absoluta dos gráficos, não há uma seqüência recorrente desses pontos que indique com segurança um dos sistemas como líder de uma mudança em dado ambiente e difundindo-se para outros. Em outras palavras, as tendências dos três *corpora* são similares, mas não guardam relação de causalidade umas em relação às outras.



É imperativo observar, ainda, que os resultados apresentados nesta Seção devem ser analisados em conjunto com os oferecidos à Seção anterior: aqui, temos a comparação entre as freqüências absolutas dos dados; àquela Seção, as taxas de mudança.

Em C14, apresentamos as curvas referentes à distribuição de estruturas **a** (SVC), **i** (VC) e [VC]. 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 **a** (SVC) ficam na faixa dos 10 – 30%, as **i** (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 já expostos na Seção que trata da Metodologia.

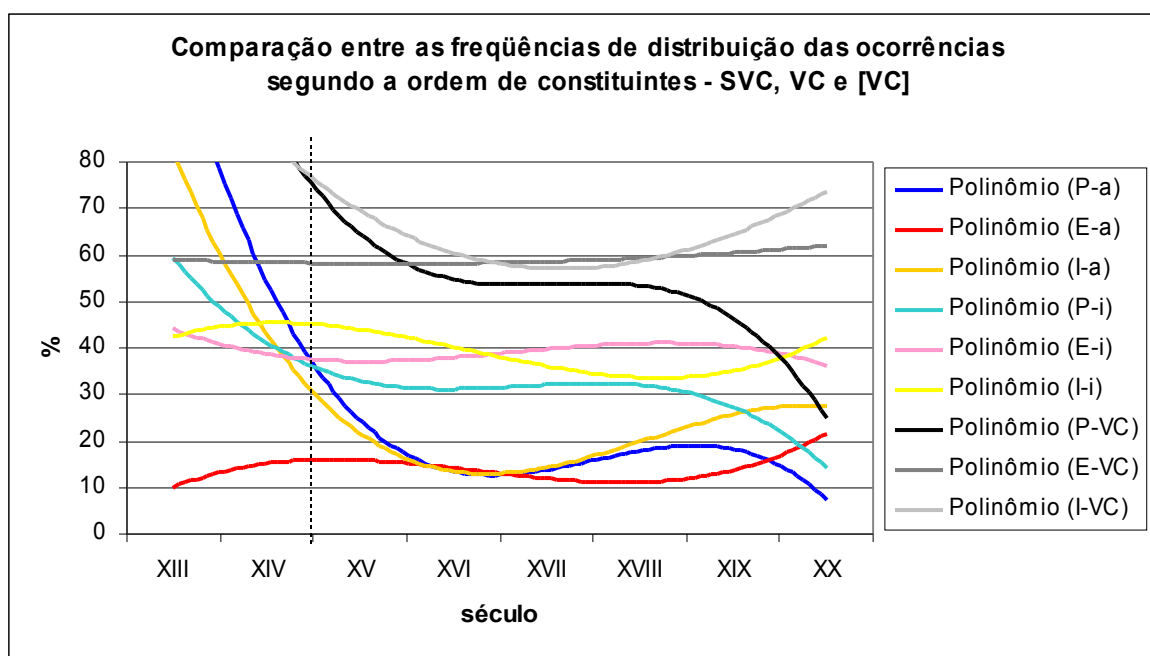


Gráfico C14 – sobreposição das curvas de frequência absoluta de ocorrência de estruturas SVC, VC e do grupo VC.

O Gráfico C15 reúne as curvas das classes **g** (SV) e **h** (VS), além de trazer, para efeito de aproximação com C14, as curvas de [VC]. Os dois conjuntos de curvas relativas às estruturas em que o complemento do verbo não é expresso estão situados abaixo dos 30% para os três sistemas, além de convergirem no final do intervalo de amostragem. As

classes **g** (SV) e **h** (VS), portanto, ocorrem segundo freqüências equilibradas em toda a amostra, equiparando-se à freqüência absoluta da classe **a** (SVC) do gráfico anterior. Observa-se, para os Gráficos C14 e C15, que as curvas não permitem inferir que uma mudança de freqüência de determinada ordem de constituintes seja liderada por qualquer dos três sistemas, e nem a identificação de um *corpus* em que esses movimentos sejam tardios em relação aos demais.

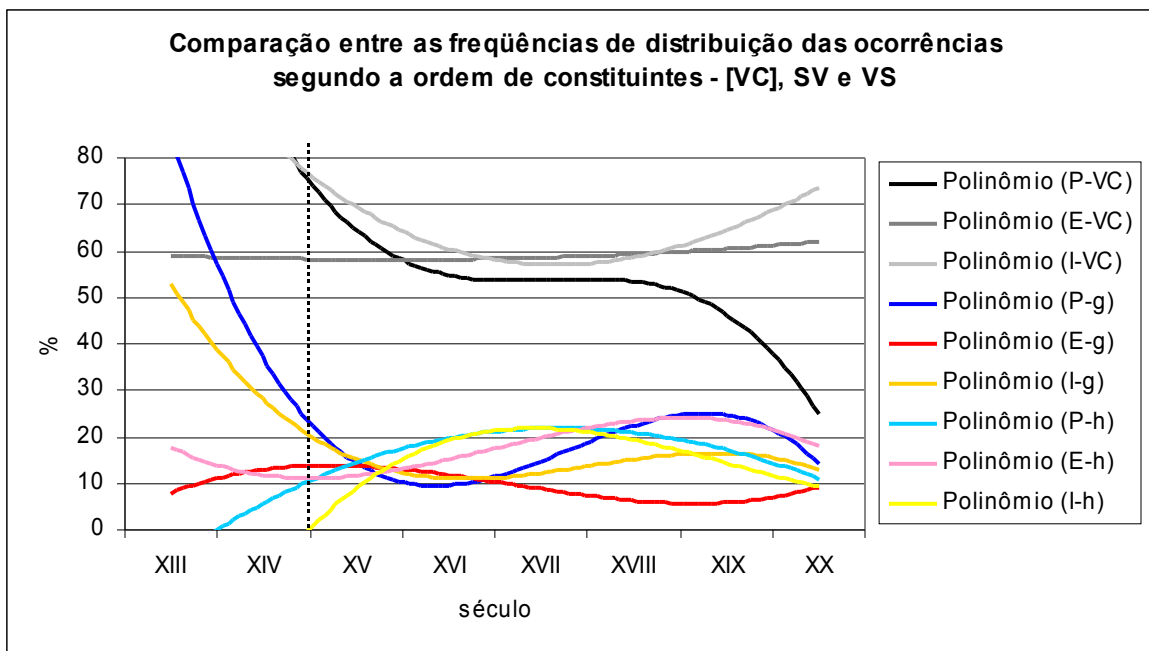


Gráfico C15 – sobreposição das curvas de freqüência absoluta de ocorrência de estruturas SVC, VC e do grupo VC.

O Gráfico C16, utilizado como exemplo acima para salientar a possibilidade de se identificarem pontos críticos para as curvas, abrange as classes de agentividade mais freqüentes, Ah e Pn, associadas a estruturas **a** (SVC). Há dois traçados básicos para os seis polinômios, mas de maneira a cobrir desigualmente as subclasses: as duas curvas de italiano se assemelham à de Ah do português, e a de Pn dessa mesma língua é similar às duas de espanhol. No que tange às classes de agentividade em orações **a** (SVC) e a partir dos dados, portanto, o sistema P se comporta em parte como o espanhol, e de outra parte como o italiano. Nota-se ainda que as curvas de italiano correm praticamente paralelas em todo o intervalo de amostragem, e que seus pontos de mínimo absoluto e de máximo local antecedem seus correspondentes na curva de Ah do *corpus* português.

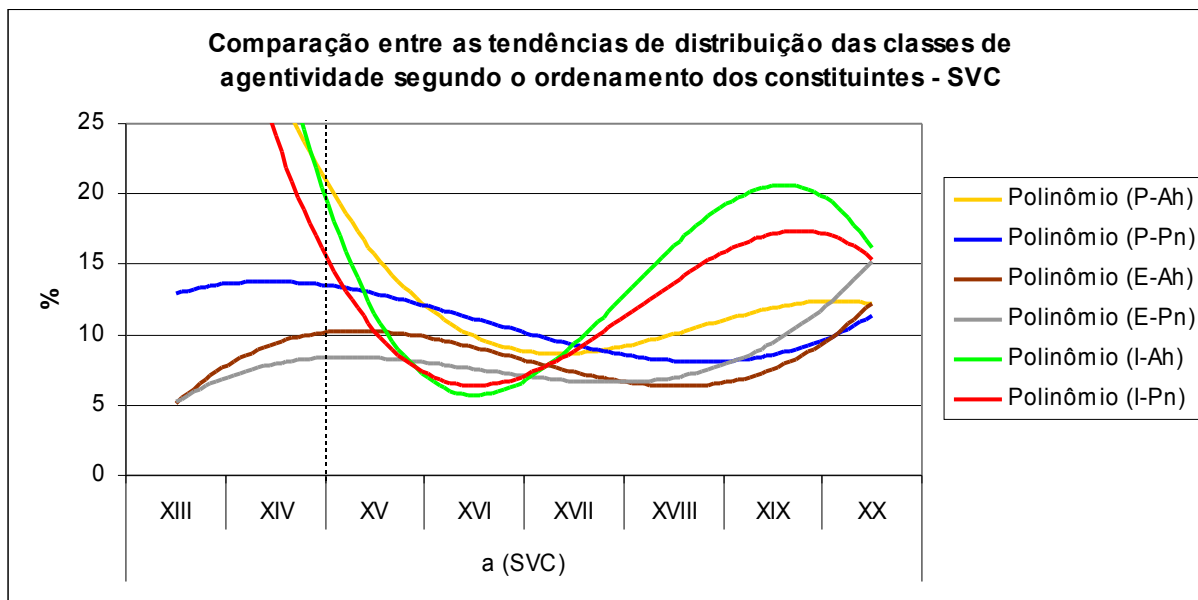


Gráfico C16 – sobreposição das curvas de freqüência absoluta de ocorrência das classes mais abundantes de agentividade em estruturas SVC.

As curvas compreendidas no Gráfico C17 constituem um feixe muito mais compacto do que o observado no gráfico anterior e distribuído numa faixa de freqüência superior à de *a* (SVC). Nota-se facilmente que as curvas de português e de italiano relativas ao mesmo traço seguem paralelas, enquanto as de espanhol parecem constituir um par isolado.

Já em C18, os polinômios têm um traçado diferente – e o próprio gráfico é plotado em escala menor do que a de C17 –, mas as curvas se associam umas às outras em pares da mesma forma que o observado para a ordem VC: P-Ah corre junto de I-Ah, P-Pn segue aproximadamente o mesmo contorno de I-Pn, e as curvas de espanhol. A propósito, os máximos absolutos das curvas de E-Ah e de E-Pn precedem os mínimos das curvas de P-Ah e de I-Ah em cerca de um século; também o máximo local de P-Pn no século XVI e o mínimo local no século XVIII antecedem, respectivamente, mínimos absolutos e máximos locais de P-Ah e de I-Ah.

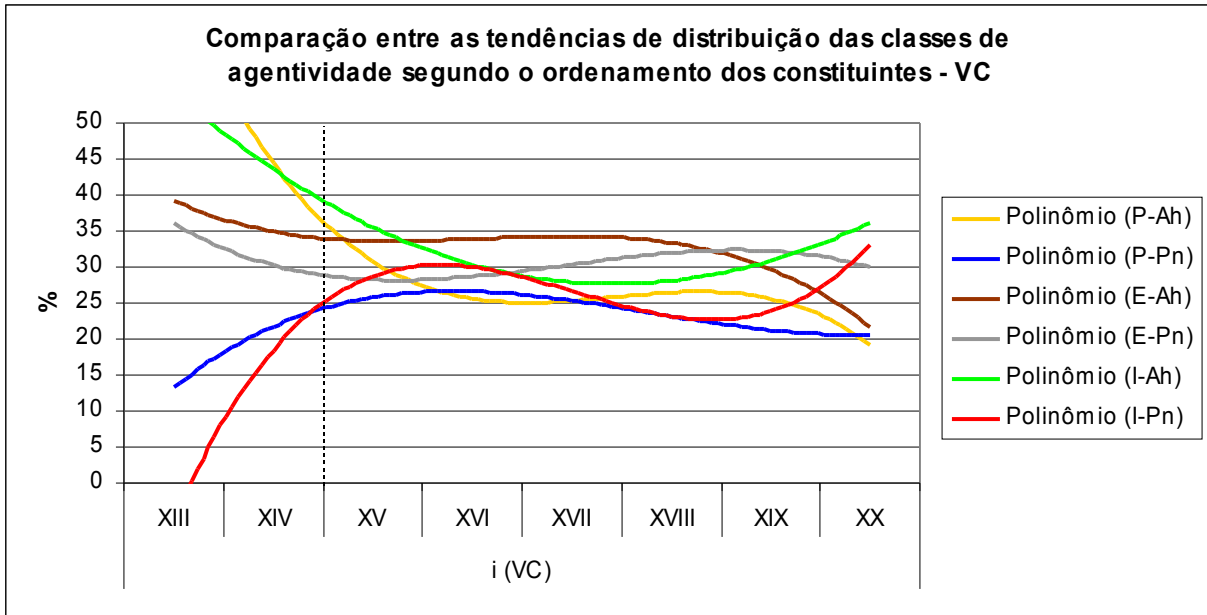


Gráfico C17 – sobreposição das curvas de frequência absoluta de ocorrência das classes mais abundantes de agentividade em estruturas VC.

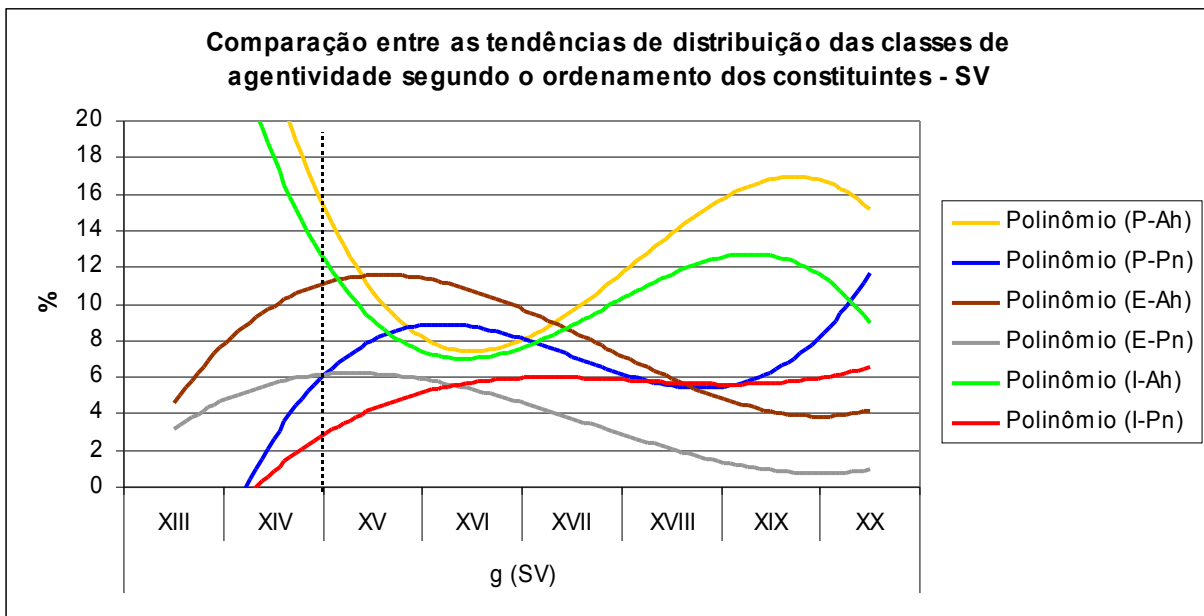


Gráfico C18 – sobreposição das curvas de frequência absoluta de ocorrência das classes mais abundantes de agentividade em estruturas SV.

A classe **h** (VS), a fechar as observações sobre as relações entre os fatores *agentividade* e *ordem de constituintes*, é representada por curvas que correm numa escala semelhante à de SV: as de português e as de italiano têm o mesmo traçado básico, enquanto as de espanhol são praticamente paralelas em todo o intervalo de amostragem. Veja-se que o período da curva de espanhol é aproximadamente o dobro do apresentado pelas curvas de italiano e de português, mas esse fato é atribuído à limitação dos *corpora* português e italiano a textos escritos a partir do século XV inclusive. De uma forma ou de outra, como a extensão e o perfil das curvas de P e de I são similares entre si, os pontos críticos das mesmas não estão distanciados cronologicamente uns dos outros.

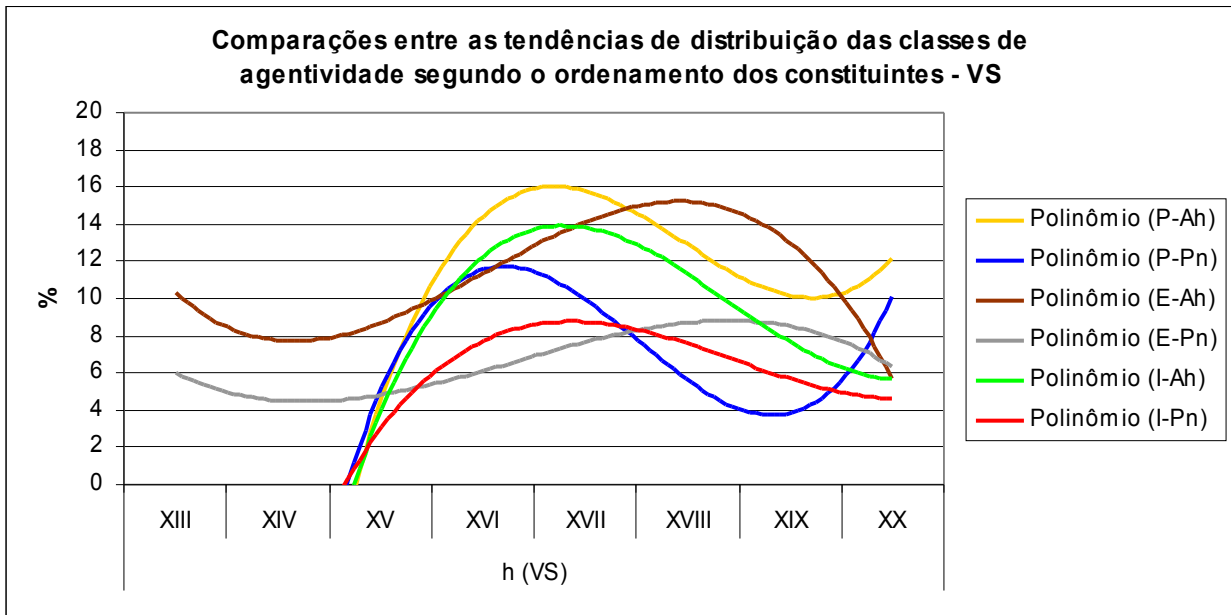


Gráfico C19 – sobreposição das curvas de frequência absoluta de ocorrência das classes mais abundantes de agentividade em estruturas VS.

Indicando, mais uma vez, as peculiaridades de cada sistema, nota-se, no Gráfico C20, que as curvas referentes às classes de *agentividade* seguem aos pares, nos *corpora* espanhol e italiano, os mesmos perfis; para o subconjunto português, os traçados das curvas de Ah e de Pn são diferentes dos demais e entre si. Já foi salientado que as frequências absolutas de orações principais tendem a aumentar nos três sistemas – por vezes, de maneira mais patente que em outras –, mas chama a atenção o fato de as curvas para os três *corpora* correrem dentro de limites relativamente estreitos de variação, constituindo um feixe nítido de perfis de mudança.

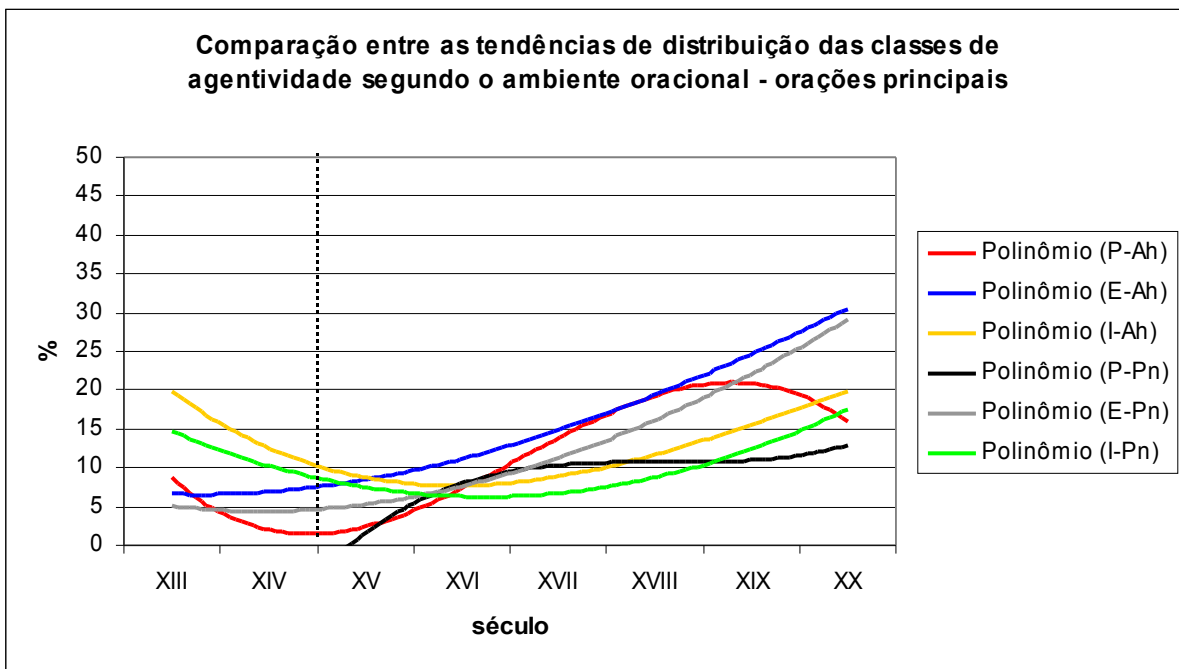


Gráfico C20 – sobreposição das curvas de frequência absoluta de ocorrência das classes mais abundantes de agentividade em orações principais.

Em C21, tomam-se as curvas relativas às classes mais comuns de *agentividade* associadas a orações coordenadas. Aplicam-se também aqui as observações feitas ao Gráfico C20 no que respeita à proximidade entre as curvas – que exibem perfis característicos para cada sistema, como já ventilado: os polinômios correspondentes a P-Ah e a P-Pn são representados por curvas praticamente paralelas, e o mesmo se verifica para o par E-Ah e E-Pn; I-Ah e I-Pn são divergentes no início do intervalo de amostragem (possivelmente em decorrência de o século XV ser o primeiro na amostra, o que causaria uma distorção no traçado inicial dessa curva), mas caminham paralelas a partir do século XVI.

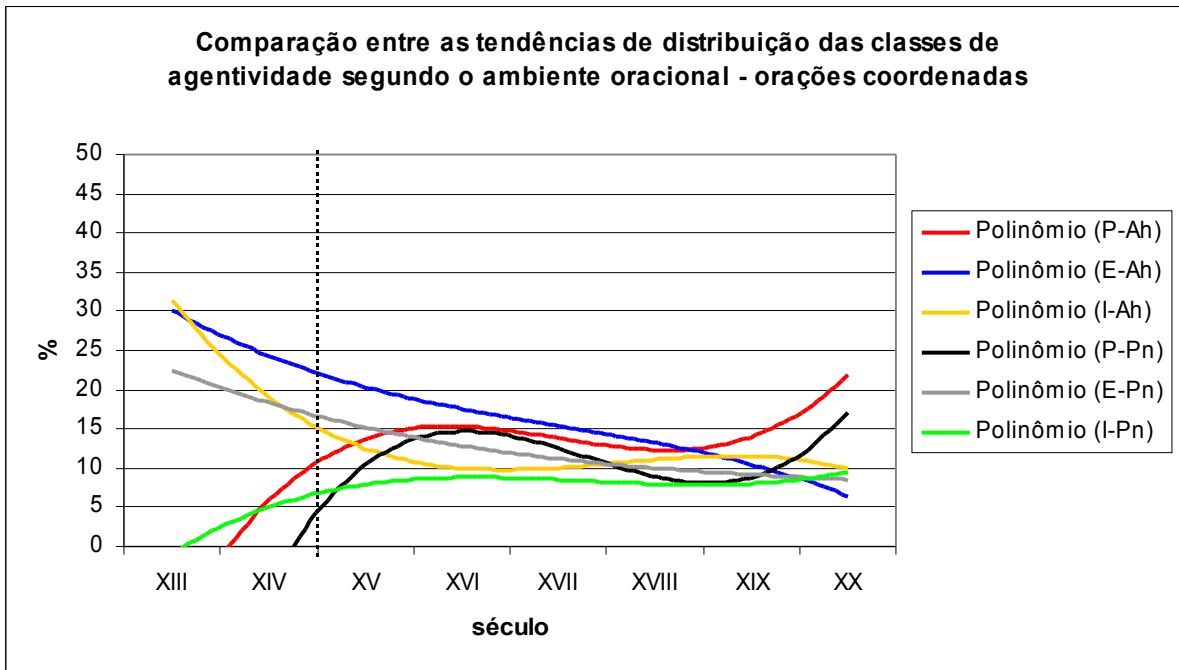


Gráfico C21 – sobreposição das curvas de frequência absoluta de ocorrência das classes mais abundantes de agentividade em orações coordenadas.

Uma dispersão muito mais sensível é observada em C22, que resume os comportamentos dos *corpora* em relação à distribuição das classes Ah e Pn em orações subordinadas. Isso mostra que as características individuais dos sistemas aqui representados transparecem mais claramente nas orações subordinadas do que em orações principais (cf. Gráf. C20) ou coordenadas (cf. Gráf. C21), nas quais se observam tendências semelhantes entre as curvas / sistemas. Mais uma vez, não parece haver relação de implicatura direta entre as curvas, que ocupam uma região superior do gráfico, correspondendo a freqüências absolutas mais elevadas, como já foi tratado anteriormente; as de italiano e português, ademais, representam segmentos cuja configuração pode mudar se se ampliar o intervalo de amostragem.

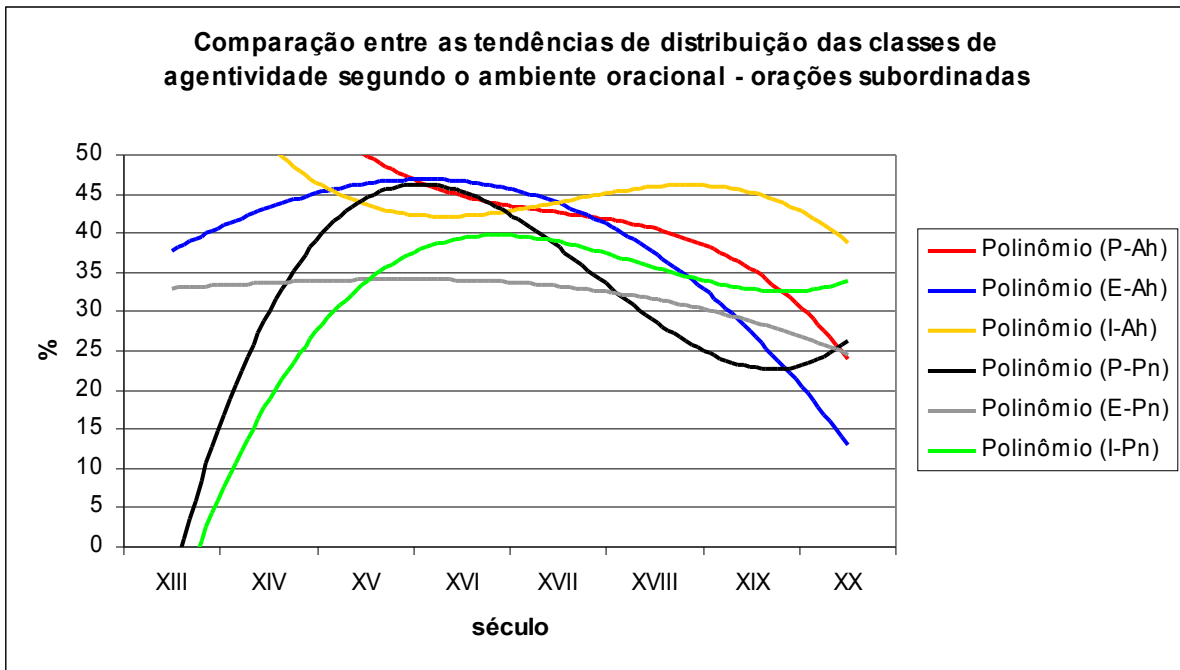


Gráfico C22 – sobreposição das curvas de freqüência absoluta de ocorrência das classes mais abundantes de agentividade em orações subordinadas.

O fator *peso de S e de C*, associado ao ambiente oracional, aponta outras particularidades dos conjuntos de dados. Observe-se, em primeiro lugar, que a distribuição das freqüências das classes mais comuns, **k** (sujeito pleno) e **n** (complemento verbal pleno), leva à construção de curvas que indicam, para as orações principais, uma dispersão das tendências a partir do século XVII – apesar de, no todo, as iRRLs dessas classes serem positivas, chegando a avançar acima dos 20%.

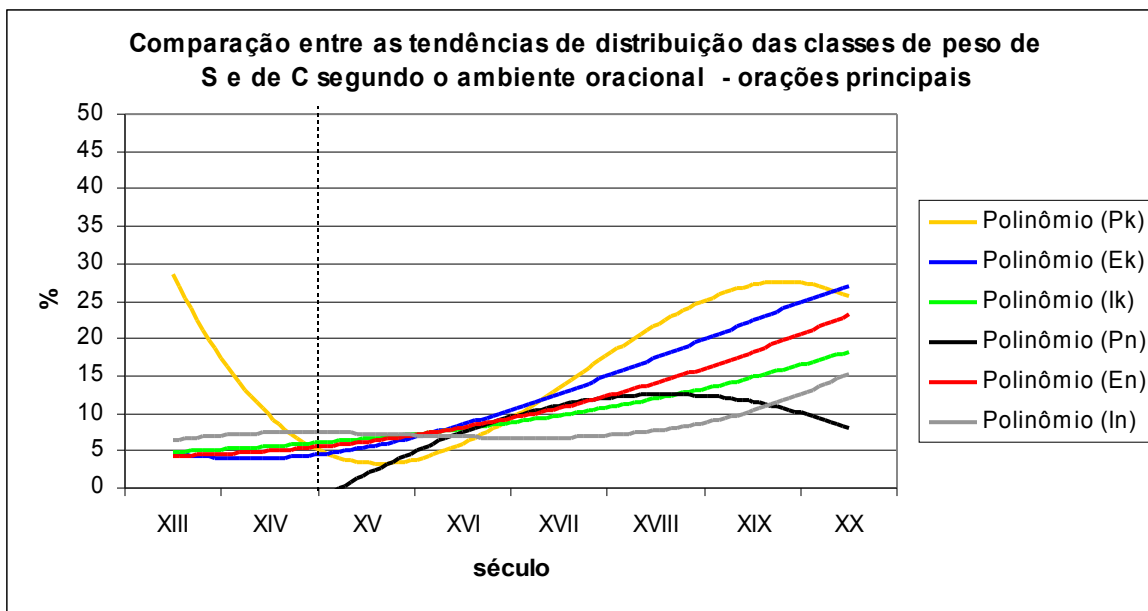


Gráfico C23 – sobreposição das curvas de freqüência absoluta de ocorrência das classes de *peso de S e de C* em orações principais.

Já para o ambiente oracional 2 (coordenadas), o feixe de curvas é mais compacto, indicando que os sistemas apresentam o mesmo comportamento na maior parte do intervalo de amostragem; nos dois últimos séculos, entretanto, haveria também uma tendência à diferenciação dos comportamentos. Esse fato sugere, no mínimo, que a dispersão das curvas em orações principais (cf. Gráf. C23 acima) antecede o mesmo movimento em orações coordenadas; conclusões mais seguras nesse sentido demandam, entretanto, estudo com maior volume de dados e extensão dos limites de amostragem. Para o conjunto, identifica-se, no Gráfico C24, a mesma tendência subhorizontal com leve declínio que está associada às orações coordenadas (q.v. também o Gráf. C27).

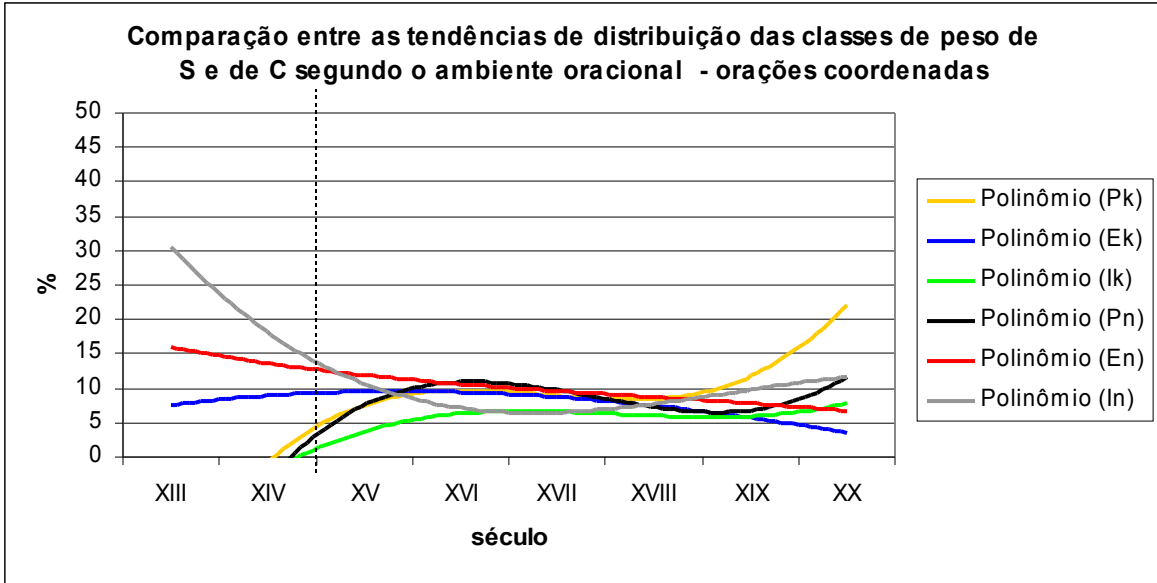


Gráfico C24 – sobreposição das curvas de frequência absoluta de ocorrência das classes de *peso de S e de C* em orações coordenadas.

Em orações subordinadas, as classes **k** (sujeito pleno) e **n** (complemento verbal pleno) mostram uma faixa de dispersão mais larga do que nos outros ambientes oracionais – i.e., as frequências de orações subordinadas distinguem melhor os sistemas do que as de principais e de coordenadas (cf., respectivamente, os Gráfs. C23 e C24 acima). Também neste caso as curvas ocupam uma região superior do gráfico e tendem ao decréscimo previsto para esse ambiente oracional em C28.

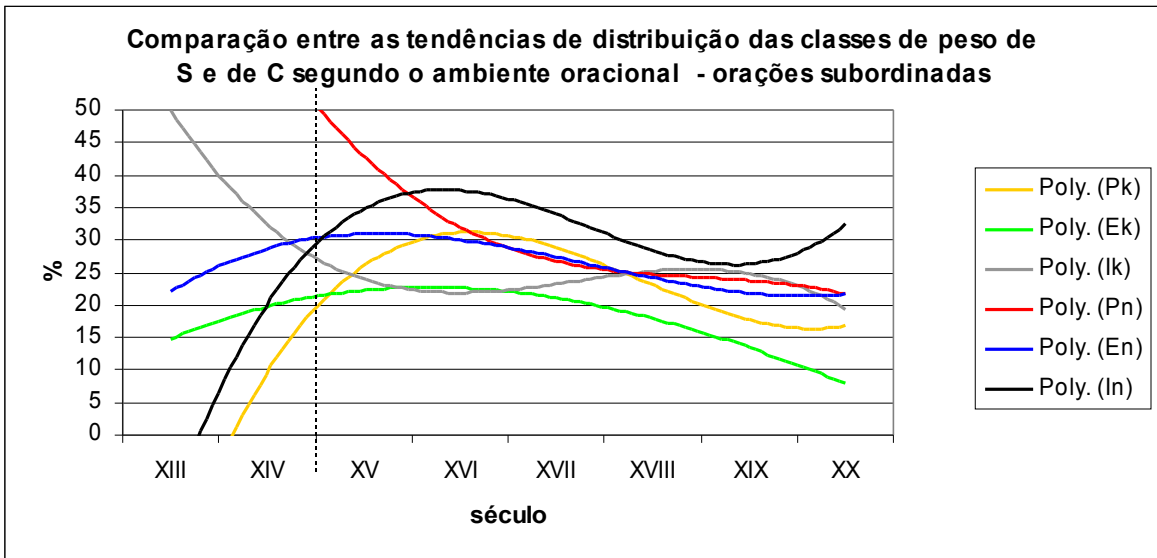


Gráfico C25 – sobreposição das curvas de frequência absoluta de ocorrência das classes de *peso de S e de C* em orações subordinadas.

Contemplando-se unicamente o ambiente oracional, observa-se que a sobreposição das curvas de cada sistema mostra que os comportamentos são parcialmente semelhantes: o aumento das frequências desse ambiente oracional é característico dos três *corpora*, mas cada base de dados tem, naturalmente, suas particularidades. Assim é que a curva para o português (P1) é a única que apresenta um traçado senoidal no intervalo de amostragem, enquanto a do espanhol (E1), cujo perfil é semelhante ao da curva de italiano (I1), afasta-se das outras duas e alcança os valores mais elevados do gráfico.

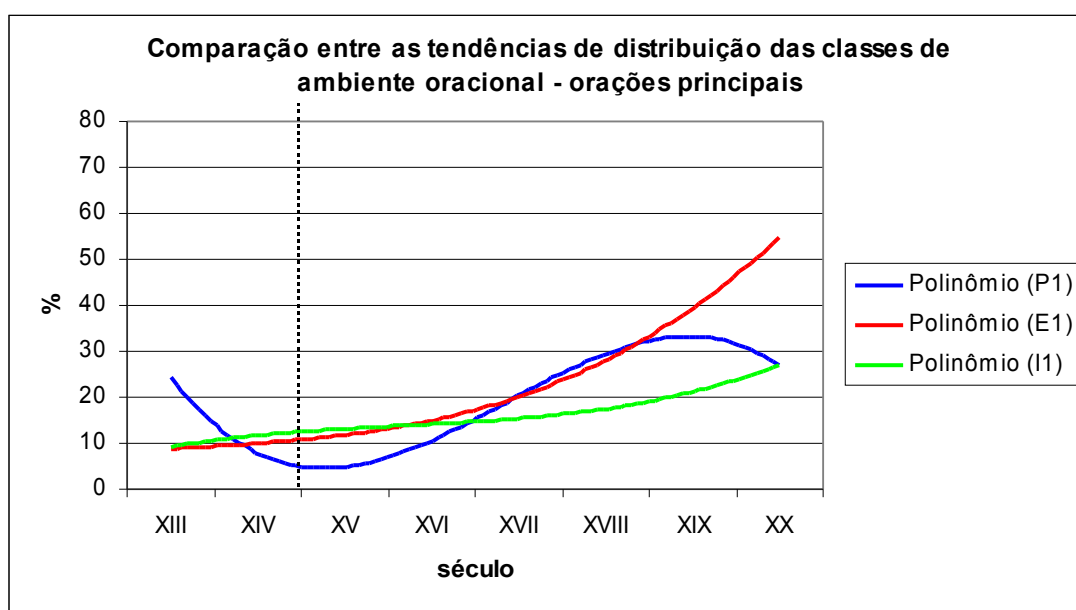


Gráfico C26 – sobreposição das curvas de frequência absoluta de ocorrência das classes de *ambiente oracional* em orações principais.

As orações coordenadas também mostram perfis distintos para cada sistema ao final do período compreendido pela amostra: em português (P2), a tendência é de crescimento discreto, como já apresentado no Gráfico C24 para a classe Pk; no caso do espanhol (E2), verifica-se uma diminuição regular em todo o intervalo de amostragem; por fim, a frequência de ocorrência de orações coordenadas em italiano (I2) se mantém praticamente inalterada no *corpus*.

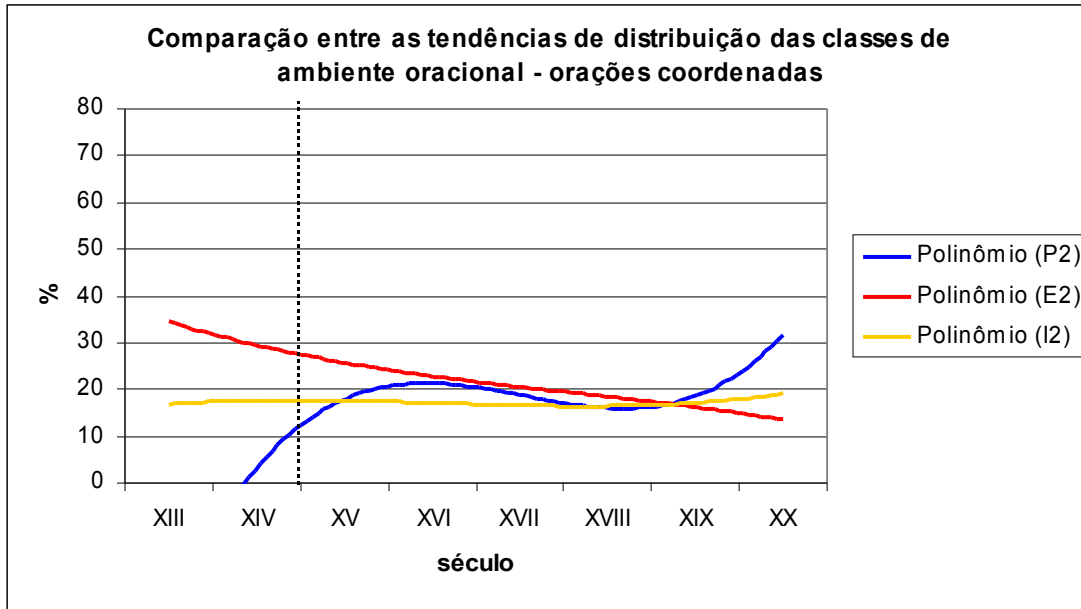


Gráfico C27 – sobreposição das curvas de frequência absoluta de ocorrência das classes de *ambiente oracional* em orações coordenadas.

Em C28, temos os perfis de frequência de orações subordinadas nos *corpora*, que repetem os padrões gerais encontrados em C25: os valores de frequência são decrescentes e ocupam a região superior do gráfico, além de manterem uma dispersão mais elevada que nos outros ambientes oracionais.

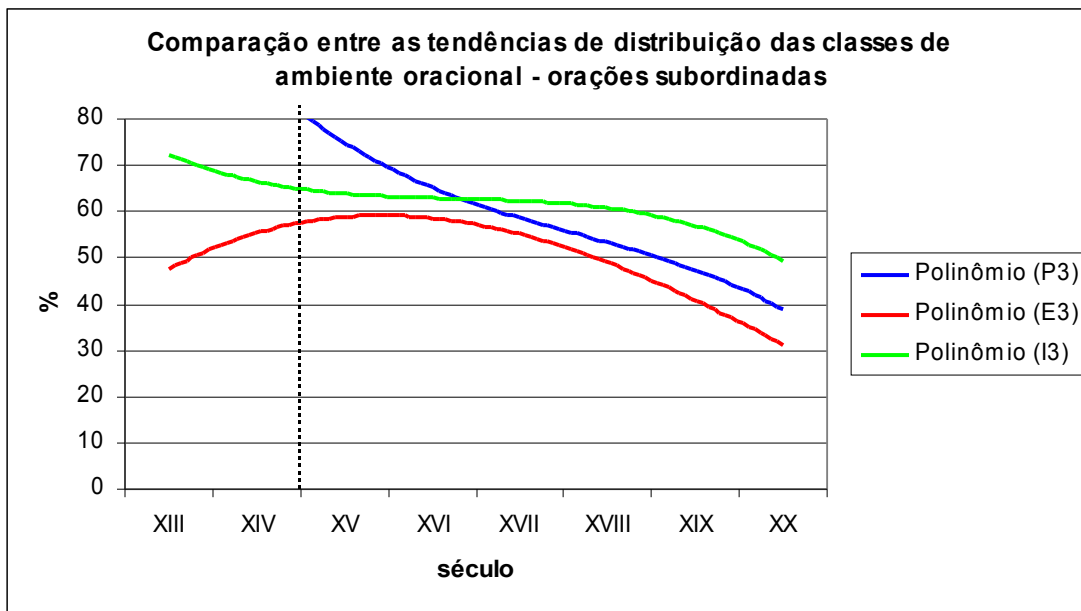


Gráfico C28 – sobreposição das curvas de frequência absoluta de ocorrência das classes de *ambiente oracional* em orações subordinadas.

Com respeito às classes de conjugação verbal, observam-se, nos Gráficos C29 a C31, distinções entre os sistemas. Em C29, nota-se uma leve tendência à diminuição das frequências absolutas de ocorrência da classe *_a_* em italiano, enquanto as curvas dos subconjuntos ibéricos sugerem aumento dessas mesmas frequências; em C30, invertem-se esses comportamentos nas curvas de frequência de verbos de segunda conjugação. As faixas de dispersão máximas das curvas, para os dois gráficos, são próximas dos 15% - inclusive ao final do intervalo de amostragem, indicando uma ótima correlação entre os dados e sugerindo que os pontos finais do intervalo podem, neste caso, representar com segurança as projeções dos comportamentos dos *corpora* até o século XIX.

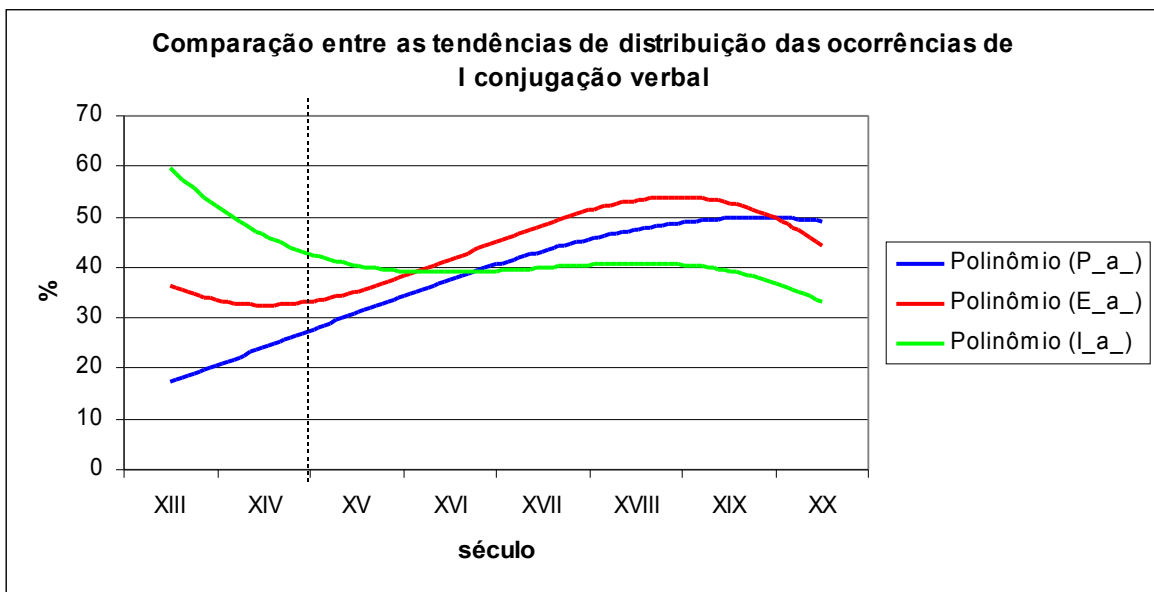


Gráfico C29 – sobreposição das curvas de frequência absoluta de ocorrência de verbos de vogal temática *_a_*.

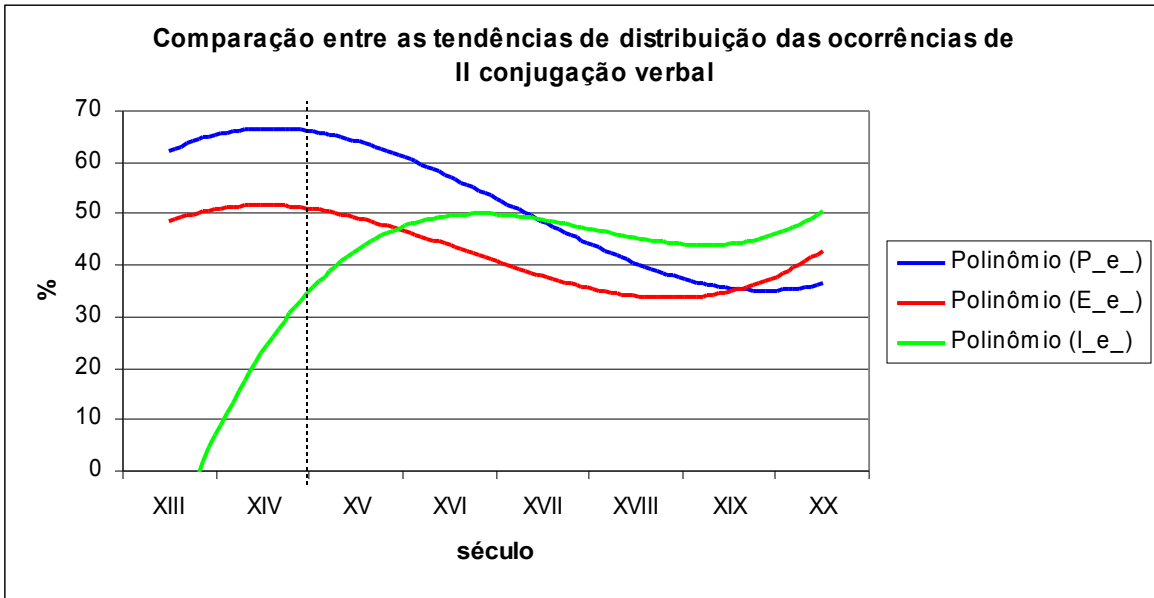


Gráfico C30 – sobreposição das curvas de frequência absoluta de ocorrência de verbos de vogal temática *_e_*.

Para a III conjugação, nota-se que as curvas ocupam a região inferior de C31, e que há um sensível estreitamento da faixa de variação das curvas – o que sugere que a frequência absoluta de ocorrência de verbos dessa classe não distingue os sistemas.

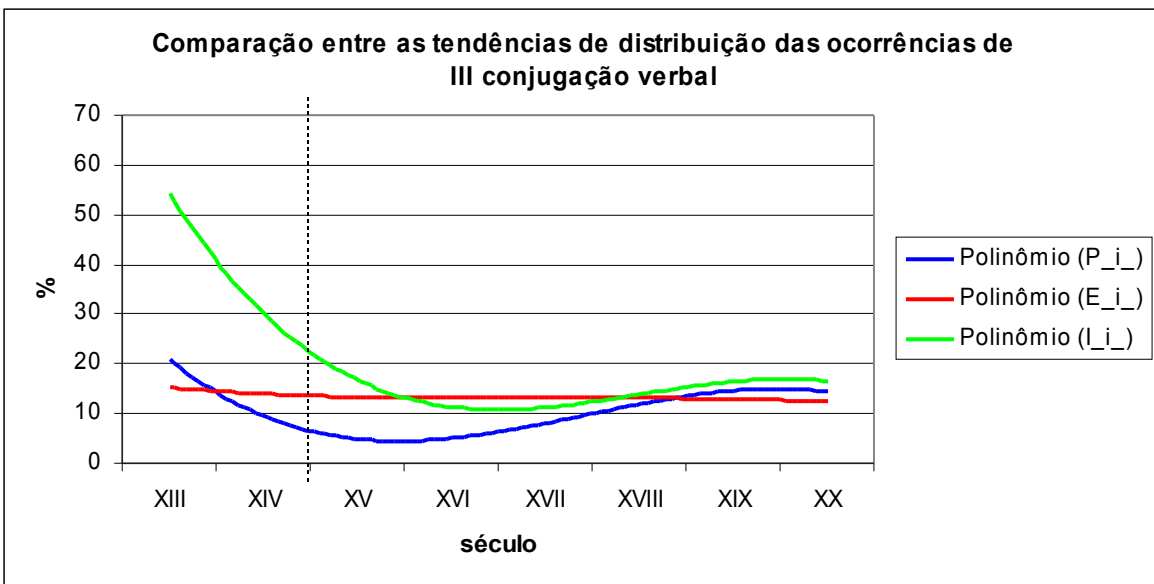


Gráfico C31 – sobreposição das curvas de frequência absoluta de ocorrência de verbos de vogal temática *_i_*.

As curvas relativas à distribuição do item lexical *ser* (P) / *ser* (E) / *essere* (I) segundo o ambiente oracional também acompanham o traçado geral definido nos Gráficos C26 a C28 para os diferentes tipos oracionais. O perfil ascendente característico da distribuição de frequências de orações principais aparece também no Gráfico C32, que mostra as proporções relativas de ocorrência do verbo *ser* nos três *corpora*: a frequência do item lexical é, portanto, função do *ambiente oracional*. O italiano se distingue dos sistemas ibéricos por apresentar a iRRL mais baixa do conjunto: 2.79, contra 7.64 (P) e 8.04 (E). Esse fato sugere a possibilidade de os sistemas ibéricos responderem mais rapidamente à mudança do que o italiano, refletindo a observação de Vidos 1968 sobre a estabilidade do italiano em relação às outras línguas neolatinas.

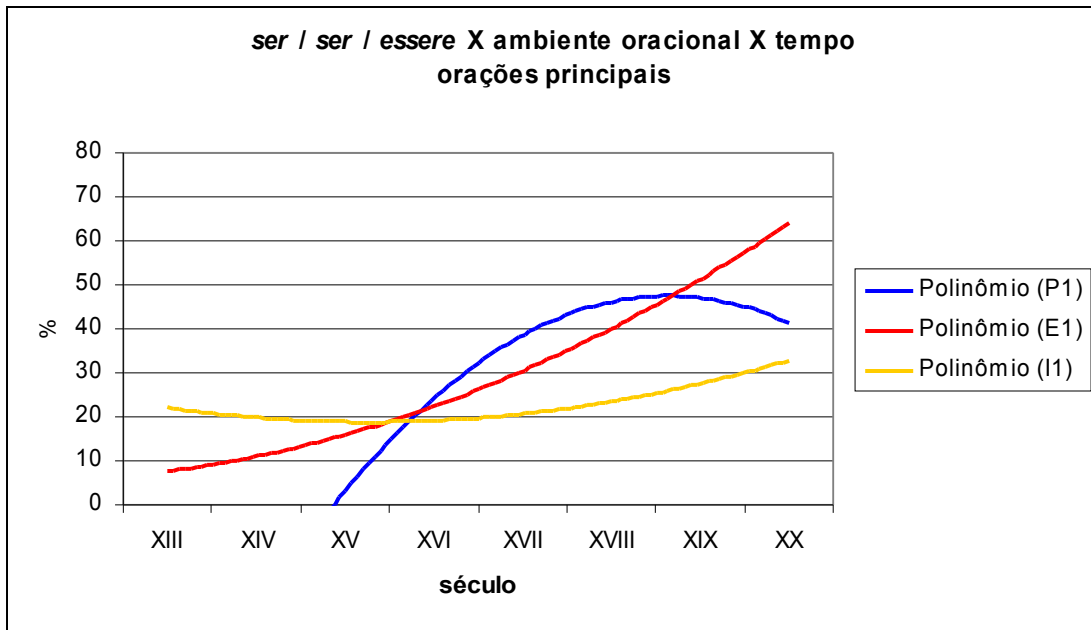


Gráfico C32a – sobreposição das curvas de frequência do *item lexical ser/ser/essere* em orações principais.

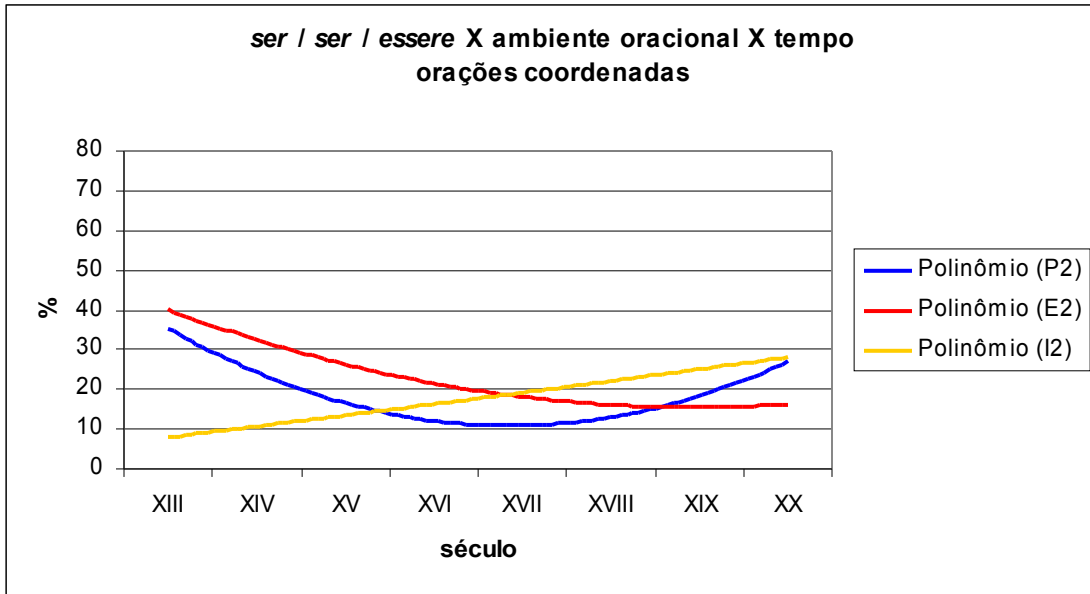


Gráfico C32b – sobreposição das curvas de frequência do *item lexical ser/ser/essere* em orações coordenadas.

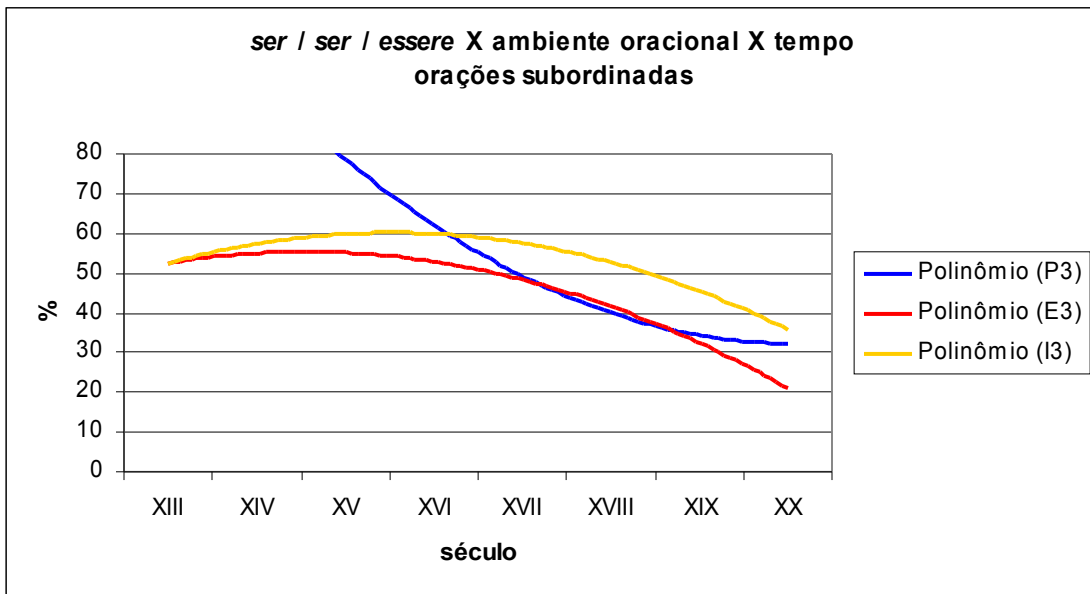


Gráfico C32c – sobreposição das curvas de frequência do *item lexical ser/ser/essere* em orações subordinadas.

Voltamo-nos, por fim, à justaposição das curvas de ordem **a** (SVC), **i** (VC) e [VC] em função do ambiente oracional. Paralelamente aos movimentos apontados nos gráficos precedentes, observam-se a) a localização dos perfis referentes às orações principais (cf. Gráf. C33a) e coordenadas (cf. Gráf. C33b) em regiões de baixa frequência absoluta, além b) da setorização das curvas em faixas diferentes do gráfico referente às subordinadas (cf. Gráf. C33c). Neste último caso, há uma distribuição semelhante à exibida no Gráfico C14, mas as iRRLs dos conjuntos apresentados em C33c são decrescentes. As frequências de ocorrência são crescentes nos três sistemas para as principais, e nota-se a formação de um feixe de curvas mais compacto e de inclinação levemente decrescente para as coordenadas em C33b.

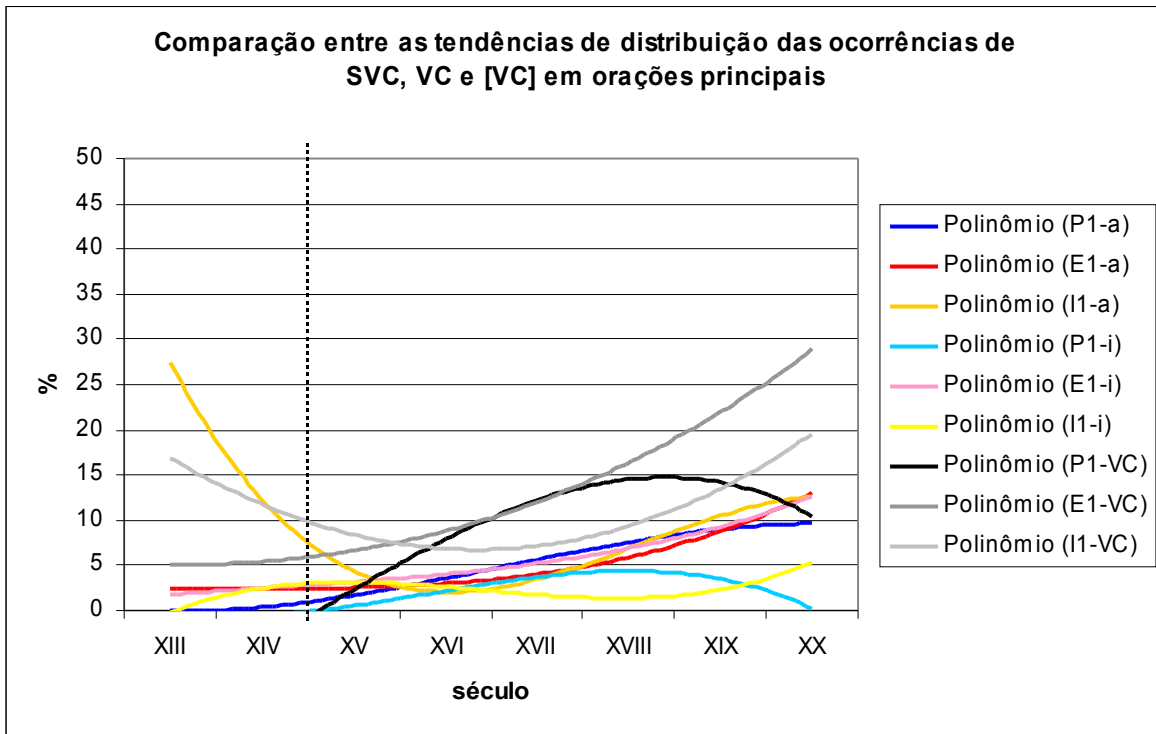


Gráfico C33a – sobreposição das curvas de frequência absoluta de ocorrência das ordens **a**, **i** e VC em orações principais.

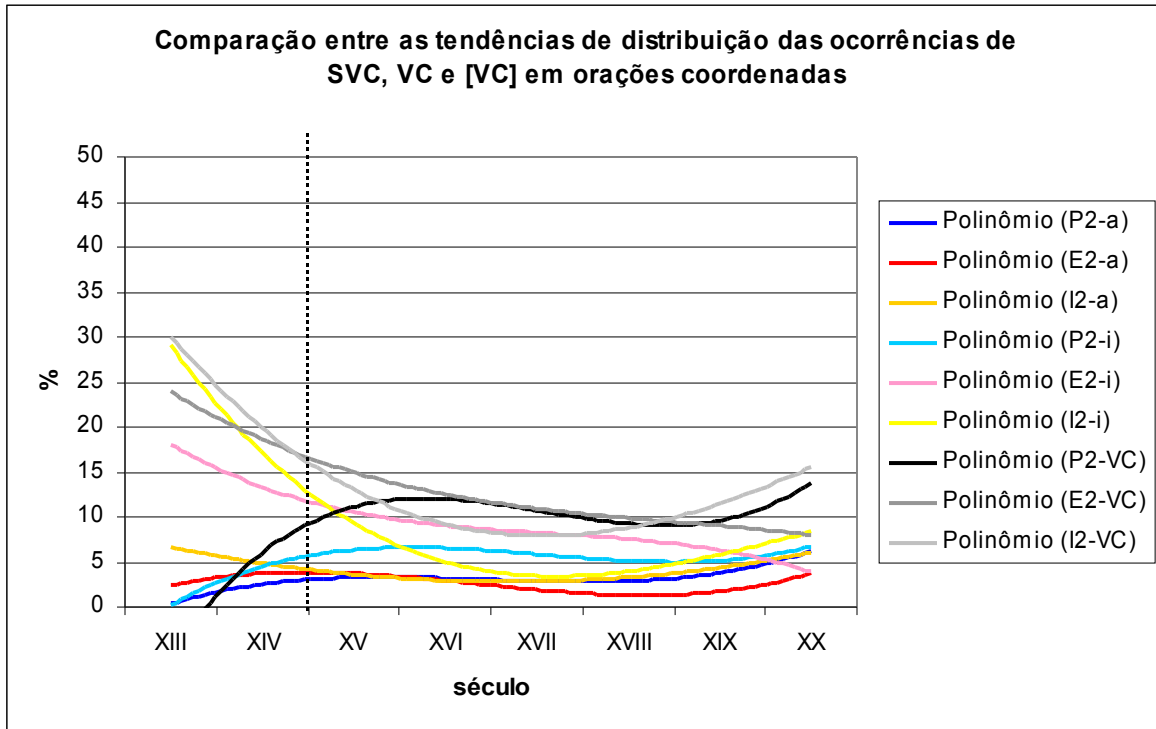


Gráfico C33b – sobreposição das curvas de frequência absoluta de ocorrência das ordens a, i e VC em orações coordenadas.

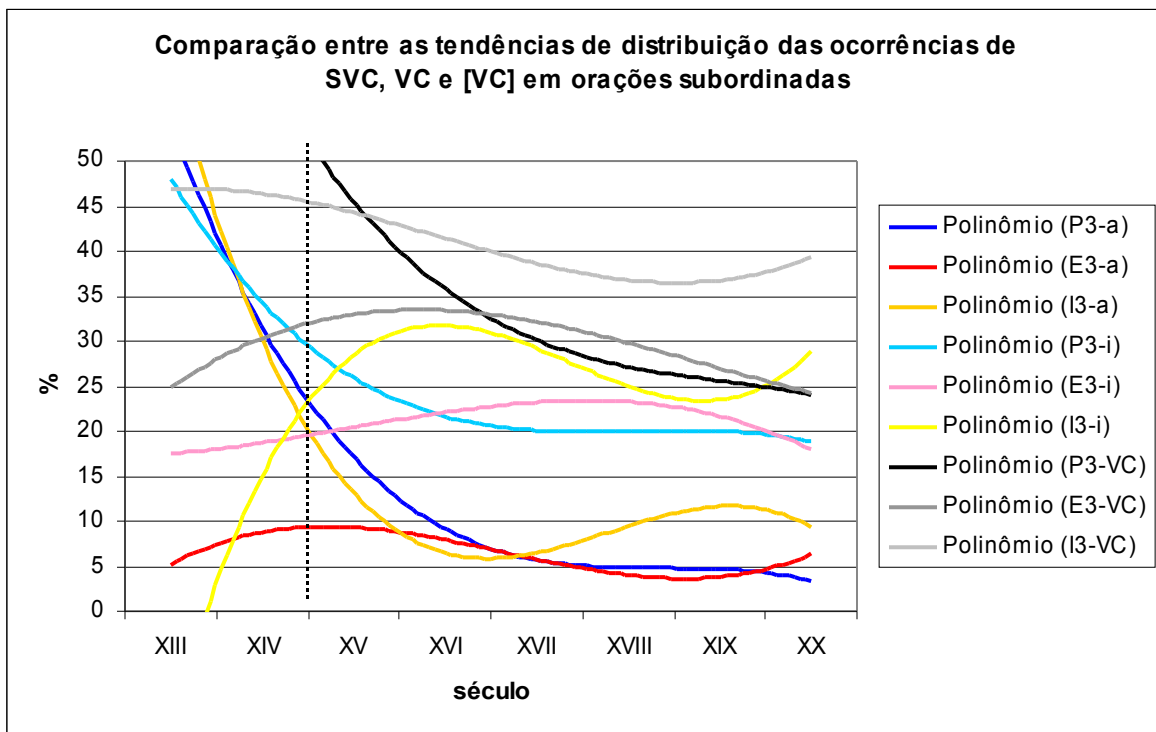


Gráfico C33c – sobreposição das curvas de frequência absoluta de ocorrência das ordens a, i e VC em orações subordinadas.

Uma comparação entre as frequências absolutas de ocorrência das ordens que têm os três constituintes S-V-C explícitos mostra que, de fato, o português é predominantemente SVC – apesar de a frequência dessa ordem sofrer uma diminuição (iRRL = -1,91) diante de outras em que um dos argumentos verbais não é expresso. A curva Σ , que reflete a somatória de todas as frequências, tem iRRL = -3,97, indicando que a ocorrência simultânea de S e de C é, em linhas gerais, cada vez menos comum em português.

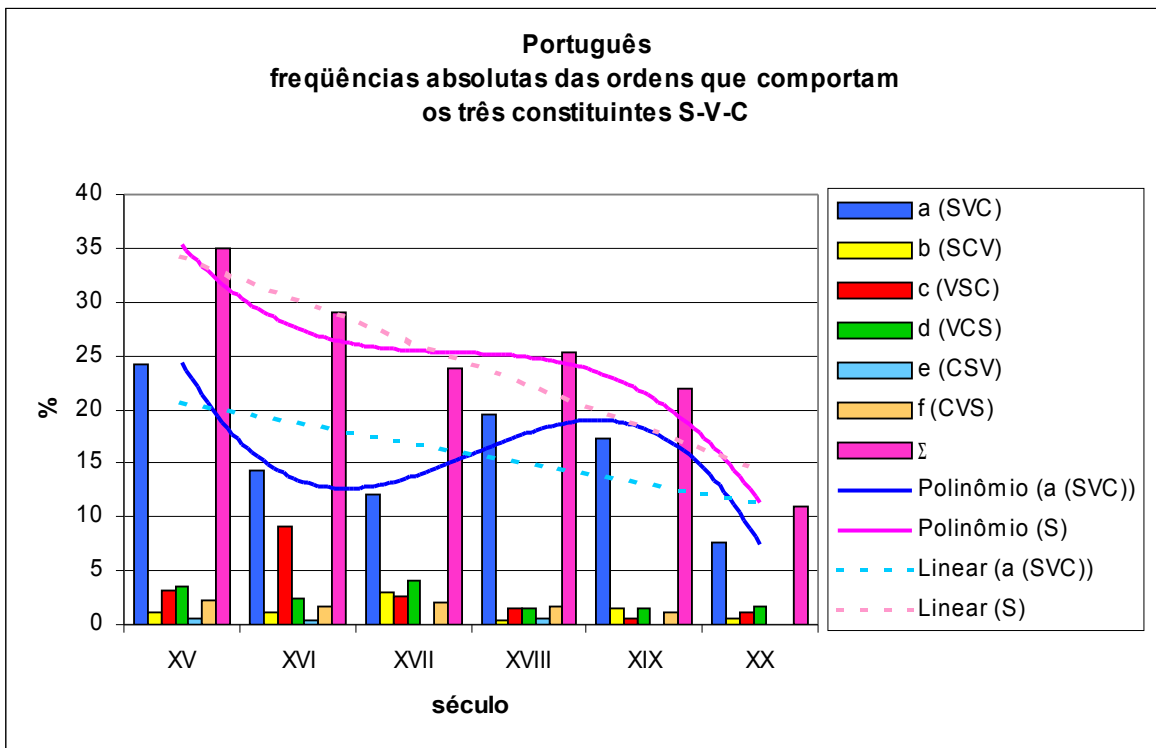


Gráfico C34 – sobreposição das curvas de frequência absoluta de ocorrência das ordens que têm os três constituintes S-V-C expressos no *corpus* português.

No *corpus* espanhol, as freqüências absolutas de ocorrência das ordens que têm os três constituintes S-V-C explícitos são dominadas pela curva **a** (SVC), que assinala a orientação principal do sistema; a despeito da oscilação das proporções associadas a essa classe e contrariamente ao que se observou para o português, o subconjunto espanhol mostra iRRLs positivas tanto para essa ordem principal (.68) quanto para a somatória Σ (.73). A curva Σ indica, neste caso, que a ocorrência simultânea de S e de C é, em linhas gerais, cada vez mais freqüente em espanhol, ao contrário do português.

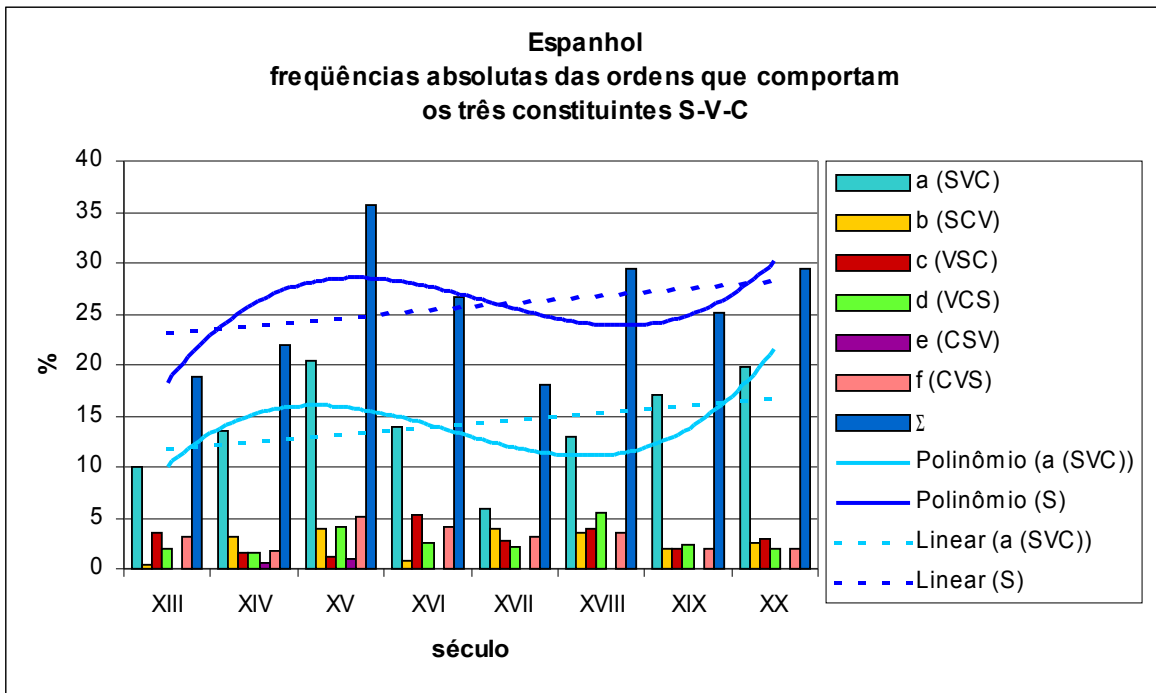


Gráfico C35 – sobreposição das curvas de freqüência absoluta de ocorrência das ordens que têm os três constituintes S-V-C expressos no *corpus* espanhol.

Para o *corpus* italiano, divisa-se a mesma prevalência da ordem **a** (SVC) já apontada nos casos português e espanhol. Essa classe, entretanto, mostra a iRRL mais elevada (2.04) dentre os três sistemas, e, em decorrência disso, a somatória Σ é influenciada de forma similar (iRRL(S) = 1.02). Isso significa que os três *corpora* mostram uma orientação definitivamente SVC, dada pelas freqüências absolutas das classes que contêm os três constituintes; contudo, essa tendência se afirma, através das iRRLs, cada vez mais fracamente a partir do italiano para o português, passando pelos valores intermediários do espanhol – i.e., num sentido de leste para oeste. Em outras palavras, os sistemas polares na România Ocidental de Vidos 1968, português e italiano, apresentam propensões sintáticas claramente díspares em diversos aspectos.

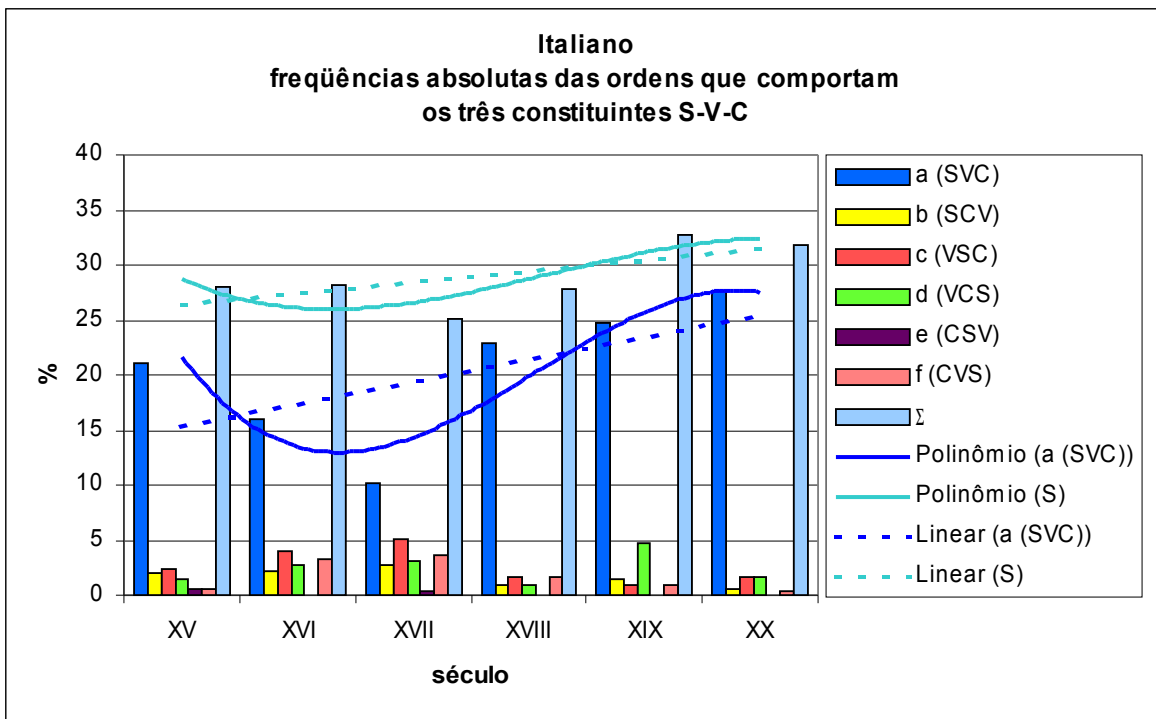


Gráfico C36 – sobreposição das curvas de freqüência absoluta de ocorrência das ordens que apresentam os três constituintes S-V-C expressos no *corpus* italiano.

Considerações finais

Os gráficos apresentados na seção anterior compreendem, como já destacado, dois tipos de comparação entre os *corpora*: os Gráficos C1 a C13 comparam as tendências dos sistemas, enquanto os Gráficos C14 a C34 permitem comparar os perfis de mudança (cf. metodologia proposta por Brill et al. 1990, Elman 1998 e McMahon & Smith 2001). Minhas conclusões estão baseadas principalmente na análise desses Gráficos, e pretendem chegar a generalizações que recuperem os pontos fundamentais desta proposta de trabalho.

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; assim é que o *ambiente oracional* pode ser desdobrado, por exemplo, segundo a natureza das orações subordinadas, e que os fatores *item lexical*, *regência* e *tempo do verbo* podem ser repensados de maneira a se precisarem melhor as características e o papel dos elementos envolvidos na mudança.

limitações da proposta

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 do que o atribuído a *ordem dos constituintes* e *item lexical*. Assim, as classes mais freqüentes de *agentividade* (Gráfs. C20 – C22) e de *peso de S e de C* (Gráfs. C23 – C25) mostram perfis que reproduzem as tendências do ambiente oracional a que o grupo de curvas se relaciona em cada gráfico; dito de outra forma, esses fatores são dependentes da variável *ambiente oracional* de forma semelhante, ou seja, os perfis dos gráficos que relacionam essas variáveis são similares. Mesmo o fator *item lexical* apresenta índices de correlação com o *ambiente oracional* tão elevados quanto *agentividade* ou *conjugação verbal*, o que leva à consideração de que a difusão entre ambientes sintáticos não envolve apenas os itens lexicais mais ou menos freqüentes, mas sim, níveis estruturais diversos (cf. também proposto por Hawkins 1979, 1983, 1990). Apesar de minha proposta se basear, enfim, numa tentativa de identificação de fatores associados a uma difusão de mudança entre ambientes sintáticos, são poucos os casos em que se observam prováveis seqüências entre os pontos críticos das curvas de tendências;

esse é o caso do português nos Gráficos C23 (orações principais) e C25 (orações subordinadas): o mínimo absoluto da curva de Pk(princ), no século XV, antecede o máximo absoluto de Pk(subord), localizado no século XVI. É 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 & Lemle 1976 e Davies 2000. Esses autores se baseiam em critérios de saliência fônica e de identificação de estruturas marcadas para demonstrar a difusão associada a itens lexicais específicos; os dados aqui apresentados indicam, entretanto, 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.

Vejam, então, alguns aspectos particulares da aplicação dos diferentes critérios de análise. Inicialmente, a partir do exame dos Gráficos C14 e C15, nota-se que, apesar de a ordem VC ser mais freqüente que SV ou SVC para os sistemas estudados, esta última é, de fato, relativamente mais comum que qualquer outro ordenamento dos três constituintes. Portanto, a orientação dos *corpora* no sentido de uma mudança SCV – SVC de Canale 1992 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 (von Wartburg 1966), como proposto por Canale *op. cit.* 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 dos ajustes apontados neste estudo, como mostram os Gráficos C26 e C28, para o ambiente oracional, e C29 e C30, para a conjugação verbal.

Outros movimentos, contudo, acompanham as oscilações de freqüência das ordens que envolvem os constituintes S-V-C no intervalo de amostragem e é possível, portanto, situá-los em relação à mudança dos constituintes A-N. O fator *ambiente oracional*, por exemplo, compreende um avanço das freqüências absolutas de orações principais (especialmente em espanhol, cujo perfil ultrapassa os 50% no século XX – cf. Gráfico

C26), e um recuo nas frequências de orações subordinadas nos três sistemas para baixo dos 50%, segundo a seqüência espanhol (séc. XVIII) – português (séc. XVIII-XIX) – italiano (séc. XX) (cf. Gráfico C28). 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 (cf. Cohen 1986/1990 e Totaro 1998); entretanto, como o italiano passa a NA entre os séculos XV e XVI, não há, até onde pude compreender, relação com o ambiente oracional semelhante à observada para o português e para o espanhol. Apesar de isso indicar que o controle das mudanças aqui apontadas não é sintagmático, mas, de fato, sintático, esse ponto também demanda estudos adicionais a fim de determinar as causas do comportamento divergente do sistema italiano nesse aspecto.

Por sua vez, fatores de natureza distinta, como *agentividade* e *conjugação verbal*, mostram a) perfis de flutuação segundo curvas mais excêntricas (nos casos de *agentividade* associados a ordens cujo S é expresso, cf. Gráficos C16, C18 e C19), ou b) perfis mais amplos em que não há oscilação, mas uma mudança segundo uma tendência geral de aumento ou de declínio (para o fator *conjugação verbal*, cf. Gráficos C29 e C30). Para os dois casos, não parece haver associação clara entre as mudanças S-V-C ou A-N e os movimentos definidos nos gráficos mencionados, como proposto por Canale 1992; a análise dos *corpora* indica, portanto, que os ajustes sintáticos a) não guardam, necessariamente, relação de contemporaneidade ou de causalidade diretas com fatores como *agentividade* e *conjugação verbal*, e b) operam em seqüências nem sempre paralelas, mas são necessários à produção de configurações harmônicas similares entre os sistemas estudados. Além disso, é digno de nota o fato de que, muitas vezes (a exemplo do que se passa com as ordens que envolvem os constituintes S-V-C), a característica dominante (SVC, no mesmo caso) corresponde a uma fração do universo de possibilidades combinatórias dos traços envolvidos (cf. Gráfs. P1, E1 e I2).

Uma comparação entre as frequências absolutas de ocorrência das ordens que têm os três constituintes S-V-C explícitos mostra que, de fato, o português é predominantemente SVC – apesar de a frequência dessa ordem sofrer uma diminuição (iRRL = -1,91) diante de outras em que um dos argumentos verbais não é expresso. A curva Σ , que reflete a somatória de todas as frequências, tem iRRL = -3,97, indicando que a ocorrência simultânea de S e de C é, em linhas gerais, cada vez menos comum em português.

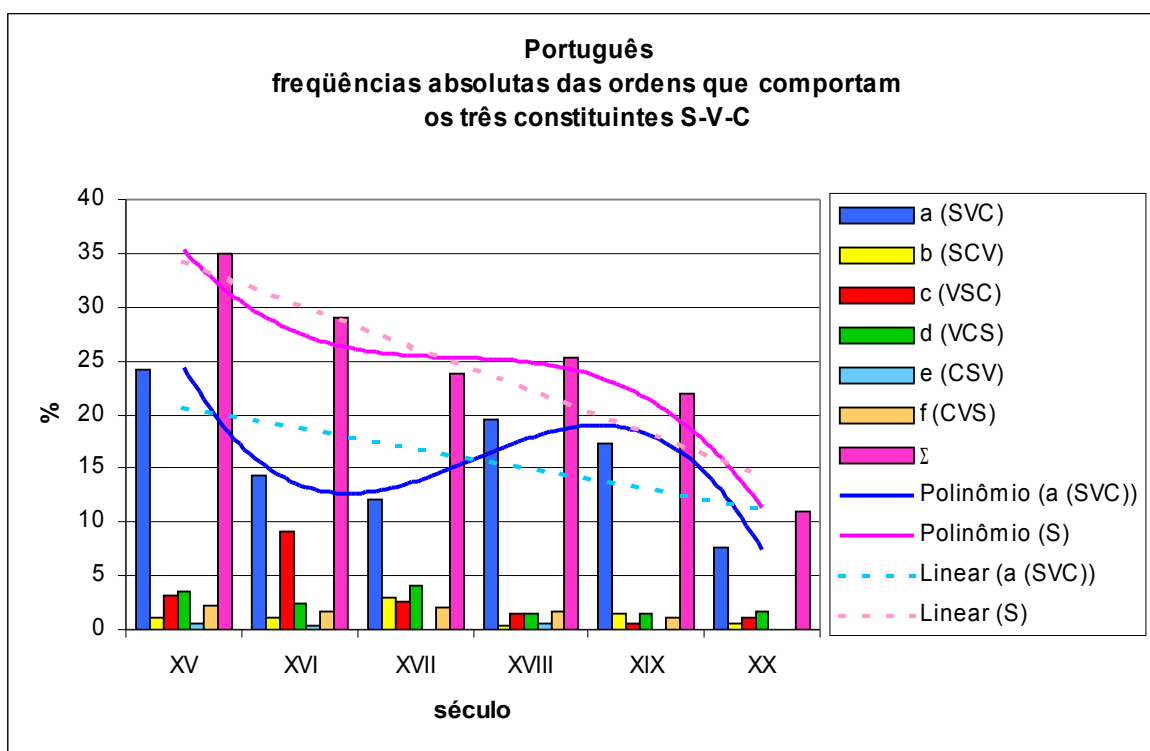


Gráfico CF1 – distribuição geral das ordens que compreendem os constituintes S-V-C no *corpus* português.

No *corpus* espanhol, as freqüências 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 freqüência de ocorrência dessa ordem é crescente (iRRL = 1,37). A curva Σ , que reflete a somatória de todas as freqüências, tem iRRL também positiva e de mesmo módulo (1,42), o que aponta a probabilidade cada vez maior de se encontrarem ordens que envolvam S e C simultaneamente em textos espanhóis.

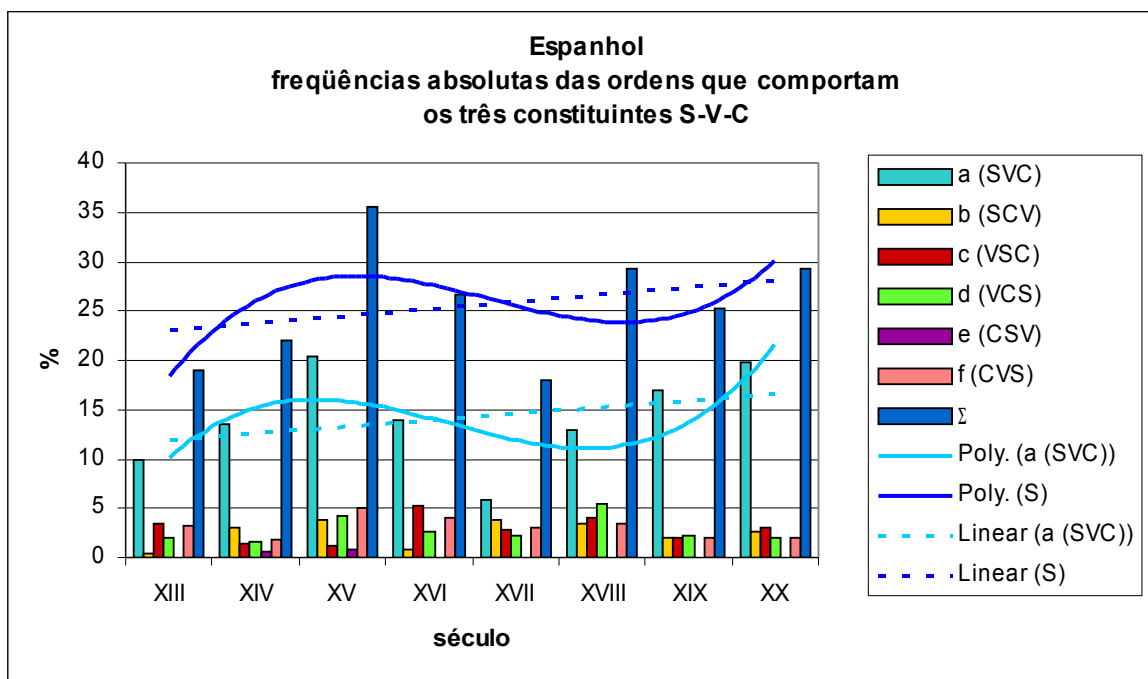


Gráfico CF2 – distribuição geral das ordens que compreendem os constituintes S-V-C no *corpus* espanhol.

Os dados dos textos italianos apontam para um comportamento semelhante ao espanhol: as proporções de ocorrências SVC, dominantes em todo o intervalo de amostragem, tendem a aumentar em termos relativos, apresentando $iRRL = 1,93$; em decorrência da diminuição das freqüências 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, mais uma vez, uma individualização do português diante dos outros dois sistemas, uma vez que se mostra menos sensível às pressões que levam a um aumento de SVC em espanhol e em italiano.

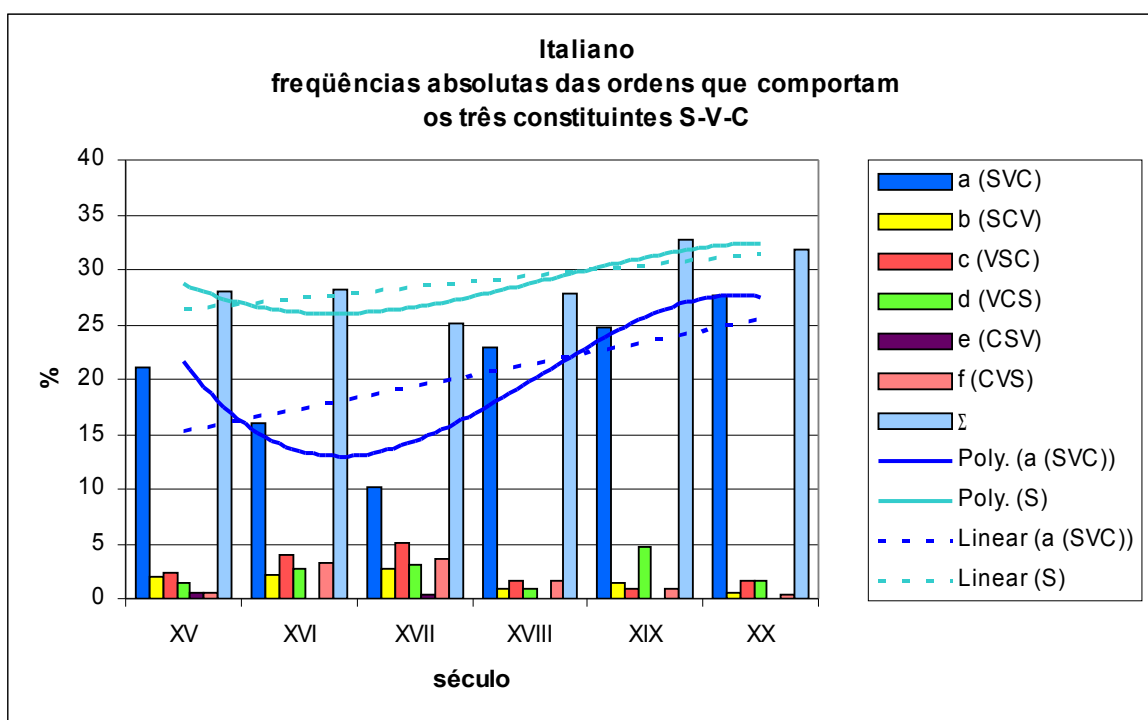


Gráfico CF3 – distribuição geral das ordens que compreendem os constituintes S-V-C no *corpus* italiano.

Em relação aos resultados fornecidos pelo tratamento das ocorrências segundo os critérios de análise selecionados, verifica-se que as alterações nas freqüências das classes de dados favorecem os traços prototípicos através das estruturas não-marcadas (corroborando, portanto, as previsões de Naro & Lemle 1981 nesse sentido): sobem as proporções de orações principais, paralelamente ao avanço generalizado das ordens [VC],

das classes **k** (sujeito pleno) e **n** (complemento verbal pleno), do subgrupo **u** (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 lingüística, a um alinhamento de vários traços estruturais cuja dependência mais estreita deve ser melhor estudada, ou a oscilações cuja amplitude não implica mudança tipológica; em termos da dinâmica de sistemas complexos (Niyogi & Berwick 1994, 1995), essas variações de frequência correspondem a órbitas em torno de um conjunto de atratores SVC / principais / Ah etc., como ilustra a Fig. F1 abaixo.

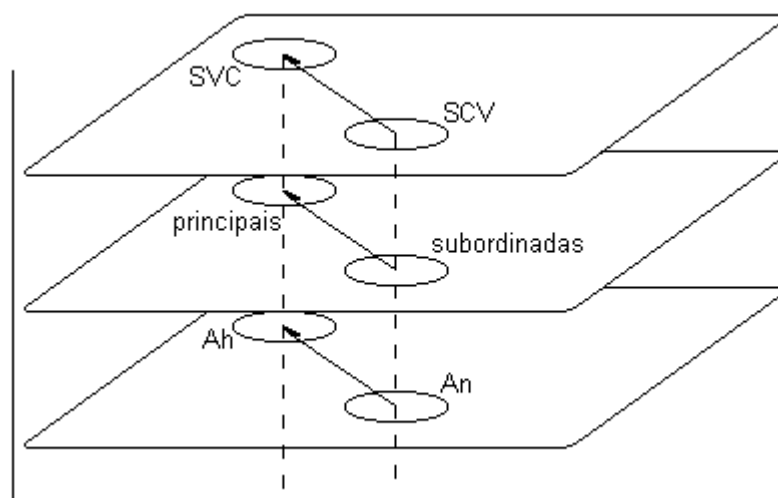


Fig. CF1 – representação da mudança tipológica em vários níveis estruturais.

Esses traços *default* comuns aos três sistemas estudados podem ser tratados sob uma única perspectiva, assumindo-se a ausência de transformações para a produção deles em relação aos traços não-prototípicos. Em outras palavras, 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; dá-se o mesmo em outras categorias, como a) argumentos verbais plenos (prototípicos) diante de seus correspondentes pronominais e oracionais (ambos transformados) e b) agente [+humano] e paciente [-humano] (*default*) em relação a agente [-humano] e paciente [+humano], respectivamente

(ambos submetidos a operações semânticas tratadas como 'figuras de linguagem'). 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 1979, 1983), 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 Fodor 1965 e retomado por Hawkins 1983, Lightfoot 1991, Dryer 1992 e Kaufmann 1993, dentre outros. Os fatores que contribuem para essa estabilidade de combinações não-*default* ainda são uma questão em aberto.

À luz de uma avaliação positiva dos resultados oferecidos pela orientação teórica e pela metodologia adotadas, atribuo maior peso à necessidade de se ampliarem os estudos a fim de abarcar *corpora* mais extensos e oferecer uma análise difusionista similar para as mudanças processadas nas demais línguas românicas, de maneira a permitir contemplar os aspectos diacrônicos que caracterizam o grupo e individualizam cada sistema. 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.

Em linhas mais gerais, nota-se que a individualização das línguas ibéricas diante do italiano não é clara a partir dos critérios selecionados, e é mais razoável afirmar que a metodologia aqui utilizada permite identificar, nos três sistemas, similaridades e particularidades estruturais. No primeiro caso, considerem-se a distribuição das ocorrências das ordens mais comuns segundo faixas de frequência (cf. Gráfs. C14 e C15) e as tendências à elevação das proporções de orações principais e à diminuição de subordinadas (cf. Gráfs. C3, C26 e C28); no segundo, tomem-se as iRRLs claramente decrescentes de Ah e de Pn no português em relação aos valores crescentes para o italiano (cf. Gráf. C4) e a dispersão das curvas relativas ao comportamento das ocorrências em orações subordinadas, segundo vários critérios de análise (cf. Gráfs. C22 e C25, por exemplo). Esses resultados

Semelhanças e diferenças entre os *corpora*

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 α para β , 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*.

Para Heine & Kuteva 2006:68ss, a transmissão de caracteres estruturais entre sistemas lingüísticos em contato – denominada, por eles, ‘replicação gramatical’ – obedece a restrições severas; apesar de poder atingir qualquer domínio da estrutura lingüística, esse processo não envolve preferencialmente estruturas morfossintáticas. Segundo essa proposta, portanto, a seqüência de mudanças sintáticas internas a cada sistema em apreço neste trabalho (italiano (>francês) > espanhol > português) 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 lingüísticos apenas reforçaria ou aceleraria uma mudança estrutural em andamento, e não criaria irregularidades nas línguas¹. Essa última afirmação está em 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.

Naturalmente, o fato de o italiano ter passado, por exemplo, de uma orientação predominantemente AN para outra NA antes das demais línguas estudadas neste trabalho ainda leva a questionar a causa dessa mudança no próprio italiano. Ao discutirem uma série de alterações ocorridas nas línguas européias ocidentais, Heine & Kuteva 2006:94 sugerem a) que essas mudanças primárias sejam o resultado de uma combinação de fatores sistêmicos e externos – ou seja, verifica-se também aqui a causação múltipla –, b) que elas são limitadas à disponibilidade de estruturas semelhantes nas línguas envolvidas, e c) que a difusão entre sistemas seja conseqüência do prestígio cultural atribuído à comunidade-fonte das mudanças. Esse modelo, em consonância com a proposta de Harris & Campbell 1995

¹ Como exemplificado por Herbstzeit > Zeit des Herbstes, na Bélgica Oriental, por influência do francês temps d’automne (Riehl 2001 in Heine & Kuteva 2006:55).

(cap. 6, principalmente) e com as observações de Nielsen 1989 sobre o grupo germânico, prevê, portanto, a operação de mecanismos semelhantes na difusão de uma inovação através de uma comunidade que fala uma única língua (como previsto pelo modelo laboviano e seus desdobramentos – q.v. Milroy 1987, Milroy 1992, Santa Ana & Parodi 1998, Labov 2006 a fim de apreciar algumas perspectivas sobre a difusão lingüística na comunidade de fala) e em situações de contato entre comunidades vizinhas que falam línguas distintas, geneticamente aparentadas ou não (q.v. Hock & Joseph 1996, Dixon 2002, Harris & Campbell 1995, Heine & Kuteva 2006a, b).

Uma das hipóteses de trabalho que lancei na fase inicial deste estudo envolvia o prestígio da língua italiana como determinante na transmissão de caracteres estruturais entre as línguas EIP, mas essa hipótese teve de ser revista porque os perfis obtidos para os *corpora* indicam que o italiano não lidera a mudança sintática em todos os ambientes e que os comportamentos dos sistemas são diferentes. Se o contato e o prestígio, portanto, facilitam a transmissão de um padrão através de domínios lingüísticos distintos, deve haver outros fatores que limitam sua relevância ou competem com eles, o que torna o modelo mais complexo do que o sugerido por esses autores; a meu ver, as melhores perspectivas são oferecidas pela caracterização das línguas como sistemas dinâmicos não-lineares, como proposto por Briscoe 2000, o que, aliado a um tratamento detalhado das demais línguas românicas e dos outros grupos lingüísticos europeus, pode levar a uma compreensão mais acurada da natureza dos mecanismos responsáveis pela mudança lingüística e das relações entre eles.

Sugere-se, por fim, a ampliação das bases de dados de maneira a contemplar outras categorias sintáticas, outros níveis estruturais e outros sistemas lingüísticos (a começar pelo francês, no grupo românico, e pelas línguas germânicas, conforme proposto por Nielsen 1989). A condução dessa proposta de trabalho 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 lingüísticos e de outros fatores externos envolvidos nessas mesmas mudanças.

Referências bibliográficas

ABEILLÉ, A. & GODARD, D. - French word order and lexical weight - SOAS (mimeo) 1997.

ADRADOS, F. R. - Hacia una tipología de las combinaciones de rasgos lingüísticos - in JUSTUS, C. F. and POLOMÉ, E. C. - **Language change and typological variation**: in honor of Winfred P. Lehmann on the occasion of his 83rd anniversary - vol. II: Grammatical universals and typology - Journal of Indo-European Studies Monograph 31 - Washington 1999.

ALDRETE, B. - Del origen, y principio de la lengua castellana ò romãce que oi se usa en España - in MANCINI, G. - **Textos y estudios clásicos de las literaturas hispánicas** - Hildesheim: Georg Olms Verlag 1970 - documento n° 26785 da *Bibliothèque Gallica* (BNF): gallica.bnf.fr/ - 15.10.2001.

ALIGHIERI, D. – **De vulgari eloquentia** – in GIULIANI, G. – **Le opere latine di Dante Alighieri** – Firenze: Successori Le Monnier, 1878 – gallica.bnf.fr, doc. 84326.

ALONSO, M. - **Evolución sintáctica del español** - Madrid : Aguilar, 1972.

BASSETTO, B. F. – **Elementos de filología românica** – SP: EdUSP, 2001.

BATALI, J. - Computational simulations of the emergency of language - in HURFORD, J. R et al. (eds.) - **Approaches to the evolution of language** - Cambridge: Cambridge Univ. Press 1998.

BAUER, B. L. M. – **The emergence and development of SVO patterning in Latin and French**: diachronic and psycholinguistic perspectives – Oxford: OUP, 1995.

de BEAUGRANDE, R. - **New foundations for a theory of text and discourse** - cópia em CD - 1997.

BENCINI, A. & MANETTI, B. - **Le parole dell'Italia che cambia** – Roma: Edmond Le Monnier, 2005.

BIBER, D., CONRAD, S. & REPPEN, R. – **Corpus linguistics: investigating language structure and use** – Cambridge: CUP, 1998.

BITTENCOURT, V. O. - Da expressão da causatividade no português do Brasil: uma viagem no túnel do tempo - tese de doutoramento - São Paulo: PUC 1995.

BOLINGER, D. - **Linear modification** - *Publications of the Modern Language Association* 67, 1952.

BOURCIEZ, E. – **Éléments de linguistique romane** – Paris: Klincksieck, 1956

BRILL, E., MAGERMAN, D., MARCUS, M. & SANTORINI, B. - Deducing linguistic structure from the statistics of large corpora, *in* Proceedings of the DARPA Speech and Natural Language Workshop, 1990.

BRISCOE, T. – Evolutionary perspectives on Diachronic Syntax. In: PINTZUK, S., TSOULAS, G. & WARNER, A. (eds.) – **Diachronic syntax: models and mechanisms** – Oxford: OUP, 2000.

BRUNOT, F. – **Histoire de la langue française des origines a 1900** – Paris: Librairie Armand Colin 1947.

BRUNOT, F. & BRUNEAU, C. - **Précis de grammaire historique de la langue française** - Paris: Masson 1949 (3^e. éd.)

BYBEE, J. – **Phonology and language use** – Cambridge Studies in Linguistics 94 – Cambridge: CUP, 2001.

-, R. PERKINS, and W. PAGLIUCA - **The evolution of grammar: tense, aspect and mood in the languages of the world.** Chicago: University of Chicago Press, 1994.

BYNON, T. - **Historical linguistics** - Cambridge: Cambridge University Press 1977.

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.

CAPELLI, A. – **Dizionario di abbreviature latine ed italiane** – Milano: Ed. Hoepli, 2001.

COHEN, M. A. A. M. - Syntactic change in Portuguese: relative clauses and the position of the adjective in the noun phrase - tese de Doutorado, inédita - Campinas: Unicamp 1986/1990.

- “O posicionamento do adjetivo no sintagma nominal português - um estudo diacrônico” - Boletim do Centro de Estudos Portugueses - Belo Horizonte: UFMG 1988.

- “A possibilidade do adjetivo no SN português sob uma perspectiva diacrônica” - II Encontro de Pesquisa da FaLe, Boletim de Resumos - Belo Horizonte: UFMG 1995.

COMPANY, C. C. - **La frase sustantiva en el español medieval: cuatro cambios sintácticos** - México, D. F.: Universidad Nacional Autónoma de México 1992.

COMRIE, B. - **Language universals and linguistic typology** - Chicago: The University of Chicago Press 1981.

COSTA, J. - Word order variation. A constraint-based approach. Leiden, 1998. Tese de doutorado - Universidade de Leiden, 1998.

DANESI 1996 - 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 Univ. Press, 1996.

DAVIES, M. - The diachronic evolution of causative constructions in Spanish and Portuguese - tese de doutoramento - Austin: 1992.

- Syntactic diffusion in Spanish and Portuguese infinitival complements - *in* DWORKIN, S. N. & WANNER, D. (eds.) - **New approaches to old problems**: issues in Romance historical linguistics - Philadelphia: John Benjamins Publ. Co., 2000.

DILLON, M. - Linguistic borrowing and historical evidence – *Language*, vol. 21, n. 1 (Jan. - Mar., 1945), pp. 12-17.

DIXON, R. M. W. – **The rise and fall of languages** – Cambridge: CUP, 2002 (4th ed.).

DRYER, M. S. - The Greenbergian word order correlations - *Language* 68 (1) 1992.

ELMAN, J.L. – Finding structure in time - ucsd.edu/pub/neuralnets/fsit.ps.gz

FODOR, I. – **The rate of linguistic change** – The Hague: Mouton & Co., 1965.

GIVÓN, T. - **Syntax**: a functional-typological introduction - vol. 1 - Amsterdam: John Benjamins Pub. Co. 1984.

GLEICK, J. – **Chaos**: making a new science – NY:Penguin, 1984.

GOERTZEL, B. - Chaotic logic - altair.comoc.com/People/ben/ben/chlog0.html -1995a - 29/07/1997.

- The miraculous mind attractor - altair.comoc.com/People/ben/ben/cm0.html - 1995b - 29/07/1997.

GONZÁLEZ, J. G. - “La colocación del adjetivo atributivo en el español medieval: un problema metodológico e histórico” - XIX Congreso Internacional de Lingüística e

Filoloxía Románicas. Volume de Resumes - Santiago: Universidade de Santiago de Compostela 1989.

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.

GUTIÉRREZ, G. G. - Un fenómeno de convergencia lingüística por contacto en el quechua de Santiago del Estero: el desarrollo del futuro verbal perifrástico - *Estudios Filológicos*, N° 32, 1997, pp. 35-42.

HANSEN, F. - **Gramática histórica de la lengua castellana** - Buenos Aires: Librería y Editorial “El Ateneo” 1945.

HARMON, R. M. – Aspectos lingüísticos dos empréstimos em português - *Hispania*, Vol. 77, No. 3 (Sep., 1994), pp. 463-469.

HARRIS, A. C. – Cross-linguistic perspectives on syntactic change – in JOSEPH, B. D. & JANDA, R. D. – **The handbook of historical linguistics** – Oxford: Blackwell, 2005.

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

HARRIS, M. B. - **Noun phrases and verb phrases in Romance** - *Transactions of the Philological Society* 1980.

HAUGEN, E. - The analysis of linguistic borrowing –*Language*, Vol. 26, No. 2 (Apr. - Jun., 1950), pp. 210-231.

HAWKINS, J. A. – Implicational universals as predictors of word order change – *Language* 55 (3), 618-648 1979.

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

– **Word order universals** – New York: Academic Press 1983.

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

– **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 – NY: Mouton, 1996.

KAUFMANN, S. – **The origins of order**: self-organization and selection in evolution – NY: Oxford UnPress, 1993.

KIRBY, S. - Function, Selection and Innateness: the Emergence of Language Universals. PhD Thesis, University of Edinburgh.
<http://www.ling.ed.ac.uk/%7Esimon/publications.html> - 07.12.2006.

KOKTOVÁ, E. – **Word-order based grammar** – Berlin: Mouton, 1999.

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.

- Syntactic change - ftp://ling.upenn.edu/papers/faculty/tony_kroch/papers/diachronic-syntax-99.pdf - 15.12.2006.

KUNO, S. – The position of relative clauses and conjunctions – *Linguistic Inquiry* 5.117-36, 1974.

KURAMOCHI, M. and KARYPIS, G. - Finding frequent patterns in a large sparse graph. SIAM International Conference on Data Mining (SDM-04), 2004.

<http://citeseer.ist.psu.edu/kuramochi03finding.html>. 14.12.2006.

LABOV, W. – Transmission and diffusion - www.ling.upenn.edu/~wlabov/Papers/TD.pdf - 04.12.06.

LAMBERTI, F. C. C. . Regularidade versus irregularidade na formação de empréstimos lingüísticos no português do Brasil. In: 10o. Simpósio Nacional de Letras e Lingüística, 2004, Uberlândia. Caderno de Resumos. Uberlândia : Universidade Federal de Uberlândia, 2004. p. 129-129.

LAPESA, R. - **Historia de la lengua española** - Madrid: Talleres Graficos Escelicer, S. A. 1962.

- Morfosintaxis histórica del verbo español - *in* AGUILAR, R. C. & ELIZONDO, M. T. E. (eds.) - **Estudios de morfosintaxis histórica del español** (2º vol.) - Madrid: Gredos, 2000.

LEECH, G.; BARNETT, R. & KAHREL P. (1996) – Recommendations for the syntactic annotation of corpora - <ftp.ilc.pi.cnr.it/pub/eagles/corporasagl.ps.gz>

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

LEHMANN, W. P. - On converging theories in Linguistics - *Language* 48, 266-275 1972.
 - A structural principle of language and its implications - *Language* 49, 47-66 1973.
 - The great underlying ground-plans - in LEHMANN, W. P. - **Syntactic typology: studies in the phenomenology of language** - Austin: University of Texas Press, 1978.

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

LIMA, B. F. Z. – Análise da ordem dos constituintes *adjetivo* e *nome* em textos catalães dos séculos XII, XIV, XV e XX – Belo Horizonte: Fac. Letras da UFMG, 2003 (dissertação de mestrado, inédita).

LLORACH, E. A. - **Gramatica de la lengua española** - Real Academia Española - Madrid: Espasa-Calpe, 1995.

McMAHON, J. & SMITH, F.J. – A review of statistical language processing techniques - cs.qub.ac.uk/~J.McMahon/paper3.ps

de MARCO, A. – **Acquisire secondo natura**: sviluppo della morfologia in italiano – Milano: Franco Angelo, 2005.

MARIZ E PÁDUA, M. P. C. - **A ordem das palavras no Português Arcaico** - Coimbra: Fac. de Letras da Univ. de Coimbra / Instituto de Estudos Românicos 1960.

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

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 – SP: Contexto, 1992.

MAURER Jr, T. - **Gramática do latim vulgar** - Rio de Janeiro: Livr. Acadêmica 1959.

MILROY, J. - Social network and prestige arguments in sociolinguistics - *in* BOLTON, K. and KWOK, H. - **Sociolinguistics today: international perspectives** - London: Routledge, 1992.

MILROY, L. **Language and social networks** - Oxford: Blackwell, [1980] 1987, 2nd ed.

MOLL, F. B. - **Gramática histórica catalana** - Madrid: Editorial Gredos 1952.

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* 29 (3) 1976.

NIELSEN, H. F. – The Germanic languages: origins and early dialectal interrelations – Tuscaloosa: Un. Alabama Press, 1989.

NIYOGI, P. & BERWICK, R. C. - A dynamical systems models for language change - MIT: Artificial Intelligence Laboratory Memo 1515 - 1994 - <ftp://publications.ai.mit.edu/ai-publications/1500-1999/AIM-1515.ps.Z> - 29/07/1997.

- The logical problem of language change - MIT: Artificial Intelligence Laboratory Memo 1515 - 1995 - <ftp://publications.ai.mit.edu/ai-publications/1500-1999/AIM-1516.ps.Z> - 29/07/1997.

OLIVEIRA, M. F. – A voz passiva no português do século XX – Inventário: <http://www.inventario.ufba.br/03/d03/03moliveira.htm> - 29.11.2006.

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

PAIVA BOLÉO, M. – O estudo das relações mútuas do português e do espanhol na Europa e na América, e influência destas línguas em territórios da África e da Ásia – *in* PAIVA BOLÉO, M. - **Estudos de lingüística portuguesa e românica** – vol. I, tomo I – Coimbra: Ed. Universitária 1974 – 1975.

PATTERSON, W. T. – Lexical borrowings in Spanish: function, length, genealogy and chronology – Canadian Modern Language Review, v. 42 n.3 p.629-34, 1986.

PENNY, R. - **A history of the Spanish language** - London: Cambridge University Press 1995.

PERINI, M. A. - O sintagma nominal em português: uma análise descritiva - mimeo 1994.
- **Gramática descritiva do português** - São Paulo: Ática 1996 (2ª ed.).

PIDAL, R. M. - **Manual de gramática histórica española** - Madrid: Espasa-Calpe S.A. 1949.

- **Orígenes del español** - Madrid: Espasa-Calpe S.A. 1956.

POSNER, R. – **The Romance languages** – Cambridge: CUP, 1996.

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

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

ROHLFS, G. - **Romanische Sprachgeographie**: Geschichte und Grundlagen, Aspekte und Probleme mit dem Versuch eines Sprachatlas der romanischen Sprachen. München: Beck, 1971.

RUHLEN, M. – **A guide to the world's languages**: Vol. 1: Classification – Stanford: Stanford UnPress, 2000 (3rd ed.).

SALVI, G. P. – Morphosyntactic persistence from Latin into Romance - <http://geocities.com/gpsalvi/cikke/persistence.rtf> - 29.11.2006.

SANTA ANA, O. & PARODI, C. – Modeling the speech community: configuration and variable types in the Mexican Spanish setting – *Language in Society* 27: 23-51, 1998.

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

SCHMIDT, K. H. - On congruence in languages of active typology - *in* JUSTUS, C. F. and POLOMÉ, E. C. - **Language change and typological variation**: in honor of Winfred P. Lehmann on the occasion of his 83rd anniversary - vol. II: Grammatical universals and typology - *Journal of Indo-European Studies Monograph* 31 - Washington 1999.

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

SILVA-CORVALÁN, C. – **Sociolingüística y pragmática del español** – Washington, DC: Georgetown Un. Press, 2001.

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

TEICH, E. – Types of syntagmatic grammatical relations and their representation - acl.ldc.upenn.edu/W/W98/W98-0505.pdf. 1998 / 14.12.2006.

TEŠITELOVÁ, M. D. – **Quantitative linguistics** – NY: John Benjamins, 1992.

THOMASON, S. G. & KAUFMAN, T. - **Language contact, creolization and genetic linguistics** - Berkeley: Un. California Press 1991.

TOTARO, J. H. R. - Recursos computacionais para pesquisa de textos em Lingüística Histórica - V Semana de Iniciação Científica da UFMG - Belo Horizonte: UFMG 1995.

- Mudança de ordem dos constituintes A e N em textos espanhóis dos séculos XIII a XX - dissertação de mestrado, inédita - Belo Horizonte: Fac. Letras da UFMG - 1998.

VENNEMANN, T. - Topics, subjects and word order: from SXV to SVX via TVX - *in* ANDERSON, J. M. & JONES, C. - **Historical linguistics**, vol. I - Amsterdam: North-Holland Linguistics Series 1974.

VIDOS, B. E. - **Manual de lingüística românica** - Madrid: Gredos 1968.

- **Manual de lingüística românica** - Rio de Janeiro: Eduerj 1996.

VIEIRA, S. - Estudo de funções assintótico-sigmóides - tese de livre-docência - Campinas: Unicamp 1975.

VINCENT, N. - Competition and correspondence in syntactic change: null arguments in Latin and Romance. In: PINTZUK, S., TSOULAS, G. & WARNER, A. (eds.) - **Diachronic syntax: models and mechanisms**. Oxford: Oxford University Press, 2000.

WALLACE, W. D. - Object-marking in the history of Nepali: a case of syntactic diffusion - *Studies in the linguistic sciences* 11 (2) 1981.

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.

WHALEY, L. J. – **Introduction to typology** – Thousand Oaks: SAGE Publ., 1997.

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

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

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

ZIPF, G. K. – **Psychobiology of languages** – London: Houghton-Mifflin, 1935. In: CHITASHVILI, R. J. & BAAYEN, R. H. - **Word frequency distributions**. In: ALTMANN, G. & HREBICEK, L. (eds.) - **Quantitative text analysis** – Trier: Wissenschaftlicher Verlag, 1993.