

A diachronic constructional analysis of English registers: grammar and style hand in hand

Javier Pérez-Guerra (jperez@uvigo.es)
Ana Elina Martínez-Insua (minsua@uvigo.es)

Language Variation and Textual Categorisation
Research Group
University of Vigo

Conference 'Grammar of genres or styles:
which approaches to prefer?'
16 January 2015 · University Paris 3

Outline

- Introduction
- Case study
- Data
- Analysis of the data
- Concluding remarks
- Further research
- References

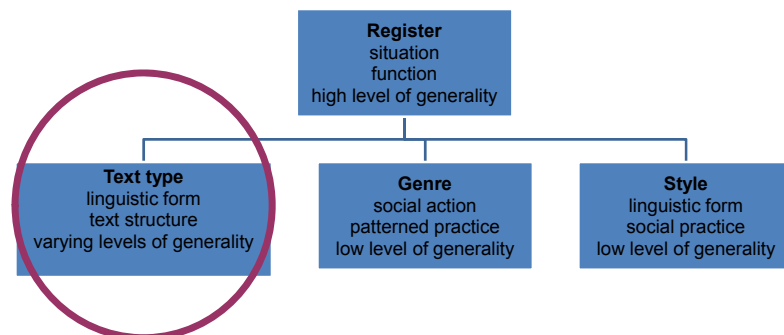
Introduction

- controversy as regards the linguistic analysis of genres and/or registers: intangible status of the concepts 'genre' and 'register'
- multi-faceted studies couched in different theoretical frameworks
 - Swales (1990: 46): “[t]he principal criterion that turns a collection of communicative events into a genre is some shared set of communicative purposes”
 - Halliday’s (1978: 122) Systemic Functional Grammar: genres analysed in terms of three variables: content (or ‘field’), participants (‘tenor’) and channel of communication (‘mode’), that is, focus on the communicative elements and purposes
 - more ‘linguistic’ approaches such as Biber’s multidimensional analysis

3

Introduction

- terminology:
text type / genre / style // register:



(Dorgeloh and Wanner 2010: 10)

4

Introduction

- This paper: two-fold view: linguistic and situational
 - (i) focus on **registers** (in fact ~ text-types) ...
 - as “grouping of texts that are similar in their linguistic form” (Biber 1988: 170)
 - as codifications of linguistic features (Taavitsainen 2001: 141)
 - “clearly relate to the form that [discourse functions] will take through aggregates of linguistic exponents of the particular text strategies that are associated with them” (Virtanen 2010: 57)

5

Introduction

- (ii) focus on **register variation**: relevance of **syntax** to register variation:
 - “Genre variation (...) is language variation beyond the limits of semantic equivalence, which is why syntax (...) provides a promising area of study.” (Dorgeloh and Wanner 2010: 8)
 - “It is form, and here morphosyntactic form in particular, that constitutes ‘a prior condition for reasoning about genre’” (Dorgeloh and Wanner 2010: 9)

6

Introduction

(iii) with a flavour of Biber's (1988, 1995, etc.) multidimensional analysis=> empirical description of registers:

- spread of a number of **objectively depicted linguistic variables** in a selection of **texts**
- **functional or situational interpretation** of the results offered by **factor analysis**

➡ **variation across registers**

Most of the variables are lexical/phrasal (frequency of lexical and functional categories) or complexity-focused (ratios incorporating length, type/token).

Only a few are clausal (coordination and subordination strategies, pied-piping versus stranded relativisation, passivisation).

This paper: **supra-phrasal features**, specifically **word-order** phenomena.

7

Case study

- syntactic supra-phrasal variables => social, situational or functional interpretations
- linearisation in sentence-initial and sentence-final clausal position:
 - topicalisation (TOP)
 - left dislocation (LFD)
 - subject-last (SBJ-LAST)
- quantitative perspective *versus e.g.*
 - Virtanen (2010): qualitative scrutiny of sentence openers in narratives texts
 - Kreyer (2010): on sentence-initial locatives in inversion constructions and qualitative analysis of the co-called immediate-observer effect function

8

Case study

- justification:
 - marked clausal (non-subject-initial) syntactic designs => hypothesis: marked social, situational or functional role
 - “the sentence-initial slot itself constitutes a rich source of discourse meanings precisely because of its cognitive relevance for our processing capacities and memory constraints” (Virtanen 2004: 12)
 - TOP: marked (complement) constituent in sentence-initial position
 - LFD: marked (non-argument) constituent in sentence-initial position
 - SBJ_LAST: marked (non-subject) constituent in sentence-initial pre-verbal position

9

Case study

- TOP:

That I had received such from Edward also I need not mention; (AUSTEN-180X,187.621) [TOP]

“Starting points are assumed to be light, small in size, and consist of **given** information. The reader’s main inferencing effort is expected to take place later in the sentence (...). Secondly, elements placed at the outset of a sentence also help readers **anticipate** what is to come as they pinpoint what the sentence is about and how it relates to the discourse topic (...). Furthermore, it is occasionally profitable to start with what is regarded as ‘**crucial information**’ (...) Sentence-initial adverbials (...) tend to form chains of text-strategic markers which have two basic functions in the discourse. They help create coherence and at the same time they signal text segmentation” (Virtanen 2004: 80-82) [our boldface]

10

Case study

- **LFD:**

[He that beleeueth on me, as the Scripture hath saide]_i, out of his_i belly shall flow riuers of liuing water. (AUTHNEW-E2-H,VII,20J.945)

- simplifying:

- “simplify discourse processing by removing a **Discourse-new** entity from a position in the clause which favors Discourse-old entities, replacing it with a Discourse-old entity (i.e. a pronoun)” (Prince 1997: 138-139; our boldface)
- making/introducing a new topic into discourse (Gundel 1985, Gelyukens 1992)
- marking a new information-unit (Halliday 1967)
- marking contrast (Chafe 1976, Gelyukens 1992).

- poset:

- “trigger an inference that the entity represented by the initial NP stands in a **salient** partially-ordered set relation to some entities **already in the discourse-model**” (Prince 1997: 138-139; our boldface)

11

Case study

- **SBJ-LAST (I):**

and very great was [my pleasure in going over the house and grounds]_{Subject}. (AUSTEN-180X,168.182)

- subject inversion:

- Green (1980: 583), Birner (1994: 241): given-new is not at work
- Dorgeloh (1997: 46), Green (1980): **anti-prominence** of the subject
- Takahashi (1992: 138): Subtopically-Presentational-Focus-emphasizing function, i.e. **presentative characterisation of the discourse topic**

12

Case study

- **SBJ-LAST (II):**

yt was enacted ordeigned and graunted by auctorite of the same p~liament, [that for x. yeres then next folowyng seväll Comyssions of Sewers shuld be made to dyv~s p~sones]_{Subject} (Statutes(II):524)

– subject extraposition:

- McCawley (1988): **newness** device (e.g. subject extraposition as a syntactic way of accommodating informatively new subjects following ‘old before new’)
- Bolinger (1992: 294): **focusing** or **presentative** effect of inversion, which is almost physical or ‘on-stage’

13

Data

- periods: (ME,) EModE and LModE:

- for Middle English (ME; 1150-1500, the second edition of the **Penn-Helsinki Parsed Corpus of Middle English** or PPCME2 – 1,155,965 words from the Helsinki Corpus of English Texts, with certain additions and deletions (see Kroch & Taylor 2000)
- for Early Modern English (EModE; 1500-1710), the **Penn-Helsinki Parsed Corpus of Early Modern English** or PPCEME – 1,737,853 words from the Helsinki directories of the Penn-Helsinki Parsed Corpus of Early Modern English, plus two supplements (see Kroch et al. 2004)
- for (Late) Modern English (LModE; 1700-1914), the **Penn Parsed Corpus of Modern British English** or PPCMBE – 948,895 words (see Kroch et al. 2010)

14

Data

- parsed corpora, with (almost) identical parsing conventions
- parsed files (.psd), using P&P-based part-of-speech and syntactic tags
- retrieval by means of CorpusSearch
- (extensive) manual revision

15

Analysis of the data

- Database:

	lfd	top	sbj-last	clauses
ppcme2	1638	1878	2989	74092
ppceme	575	359	611	34896
ppcmbe	369	352	677	60100
Total	2582	2589	4277	169088

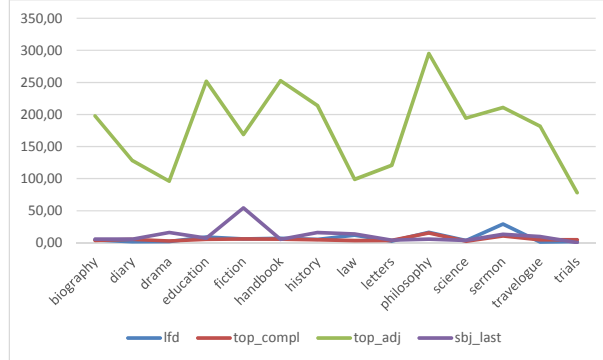
- Registers (text types): [Bible (and Fiction in ME) excluded]

biography	diary	drama
education	fiction	handbook
history	law	letters
philosophy	science	sermon
religious treat.	travelogue	trials
romance		

16

Analysis of the data

- Preliminary research on TOP_{compl}/TOP_{adj}:



- [After that a childe is come to seuen yerres of age,]Adjunct I holde it expedient that he be taken from the company of women: (ELYOT-E1-H,23.27)

17

Analysis of the data

- ME:

	freq /1,000 cls		
	lfd	top_compl	sbj_last
biography	15,11	34,62	41,93
handbook	16,65	11,10	19,26
history	9,25	13,26	36,30
law	30,70	38,28	17,54
philosophy	44,18	16,83	21,71
religious treat	29,32	31,95	35,12
romance	5,76	17,10	109,40
sermons	29,05	31,57	23,38
travelogue	14,79	15,09	72,72
mean	21,65	23,31	41,93

18

Analysis of the data

- EModE:

	freq /1,000 cls		
	lfd	top_compl	sbj_last
biography	30,33	13,31	19,22
diary	4,07	5,63	25,42
drama	4,18	8,12	19,77
education	24,29	10,05	9,12
fiction	8,77	10,96	59,90
handbook	33,44	7,17	5,30
history	19,55	17,46	21,03
law	91,65	8,15	3,64
letters	14,65	8,99	4,50
philosophy	26,20	16,16	3,50
science	41,50	9,94	16,89
sermon	32,95	6,71	18,17
travelogue	6,47	5,55	29,29
trials	6,28	4,63	12,42
mean	24,59	9,49	17,73

19

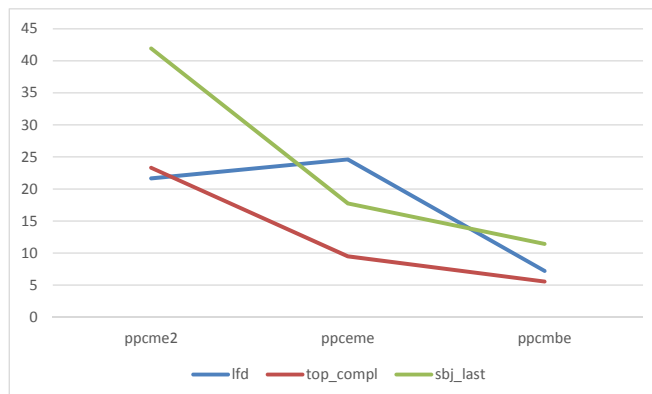
Analysis of the data

- LModE:

	freq /1,000 cls		
	lfd	top_compl	sbj_last
biography	4,34	3,67	5,67
diary	1,61	5,37	5,55
drama	1,84	2,61	15,96
education	8,97	5,12	7,04
fiction	5,99	5,99	54,17
handbook	6,87	5,62	5,31
history	5,29	4,70	15,88
law	11,94	3,47	13,86
letters	2,42	3,70	4,12
philosophy	16,27	15,18	5,42
science	3,49	2,33	3,72
sermon	29,07	10,65	13,10
travelogue	1,13	4,75	9,73
trials	1,85	4,51	0,21
mean	7,22	5,55	11,41

20

Analysis of the data



21

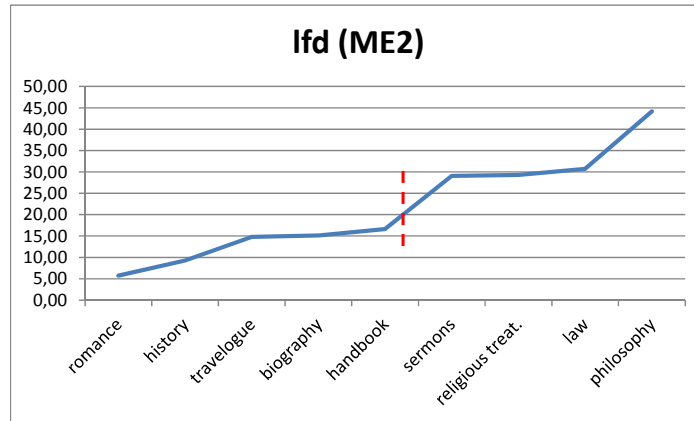
Analysis of the data

- (statistically significant) decrease of all these marked cxns ME>LModE, in line with the syntacticisation of subject-verb (-complement) word order in English
- statistical marked condition of the 3 cxns:
 - LFD: approx. 20-25/1,000 cls. in ME/EModE > 7/1,000 cls. in LModE
 - TOP: approx. 23 cls/1,000 cls. in ME > 9-5/1,000 cls. in LModE => rare construction (less frequent than even LFD in Modern English)
 - SBJ-last: approx. 42 cls. in ME > 17-11/1,000 cls. in LModE => also rare (more frequent than LFD and TOP in Modern English)

22

Analysis of the data

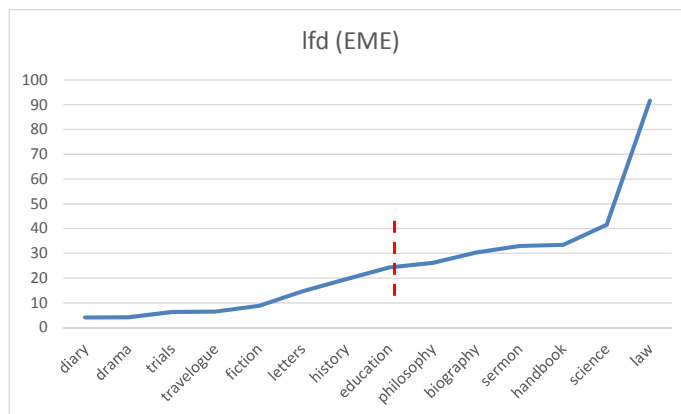
- LFD: ME:



23

Analysis of the data

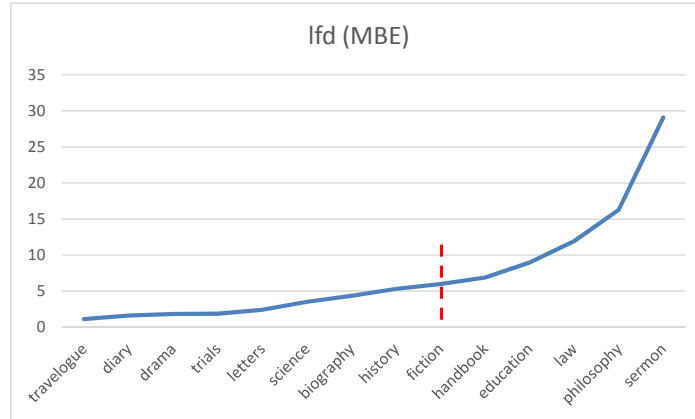
- LFD: EModE:



24

Analysis of the data

- LFD: LModE:



25

Analysis of the data

	LOWER PROPORTIONS	HIGHER PROPORTIONS
ME	Romance	Religious treatises Law Philosophy
EModE	Diary Drama Trials Letters	Science Law
LModE	Travelogue Diary Trials Letters	Philosophy Law

LFD and registers across time

26

Analysis of the data

- LFD: ME:
 - The registers which are stylistically less literate (biography, romance, travelogue) contain fewer examples of LFD.
 - The registers which are stylistically more literate (law, philosophy, religious treatises) contain more examples of LFD.
 - The sermons are grouped with the more literate registers, although they illustrate speech-based strategies (*Boethius* is included).
- LFD: EModE and LModE:
 - The relative proportions of registers in EModE and LModE are quite similar => stylistic literacy seems to be the factor also in these periods.

27

Analysis of the data

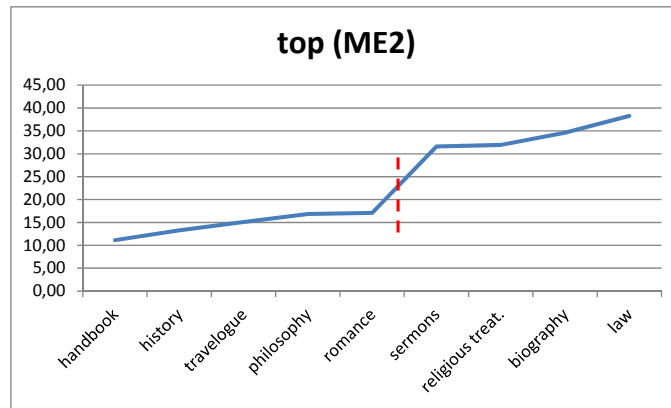
- LFD:
 - highly marked syntactic device:
 - leads to **non**-sentence-initial designs
 - left-dislocated constituents do not fulfil a syntactic function within the clause
 - hypothesis: linguistic markedness is connected with functional specificity in register analysis:

LFD is a formal indicator of stylistic literacy, at least in the recent history of English => LFD can be taken as a linguistic feature which positively contributes to the minus-plus dimension 'less literate versus more literate'

28

Analysis of the data

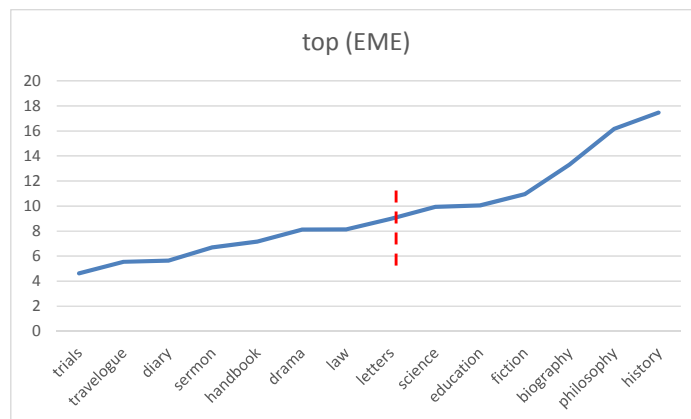
- TOP: ME:



29

Analysis of the data

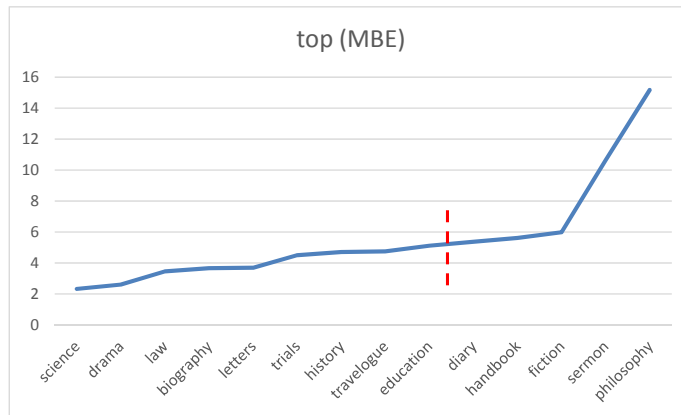
- TOP: EModE:



30

Analysis of the data

- TOP: LModE:



31

Analysis of the data

	LOWER PROPORTIONS	HIGHER PROPORTIONS
ME	Handbook History Travelogue	Religious treatises Biography Law
EModE	Trials Travelogue Diary Sermon	Biography Philosophy History Fiction
LModE	Science Drama Law Biography	Philosophy Sermon Diary Fiction

TOP and registers across time

32

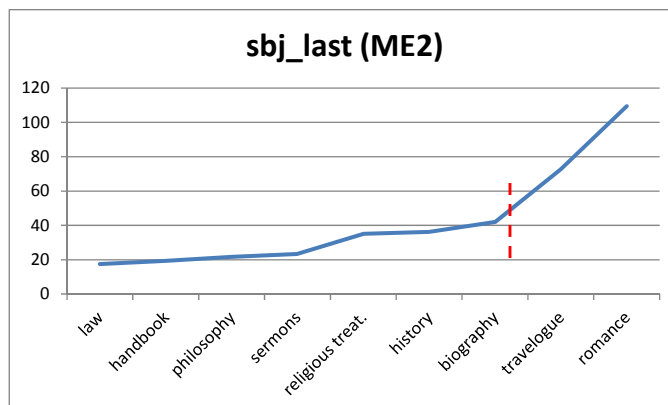
Analysis of the data

- TOP: more complicated distribution:
 - ME: less formal registers (handbook, travelogue and (speech-purposed) philosophy) in the group of registers with fewer examples of TOP, and more formal registers (religious treatises and law) with many more instances of TOP
 - But... EModE and LModE: TOP is no longer a textual marker (frequent in formal registers – law, history – and also in indisputably less formal registers – fiction, diary) => **textually unmarked status of TOP as a functional or situational marker**

33

Analysis of the data

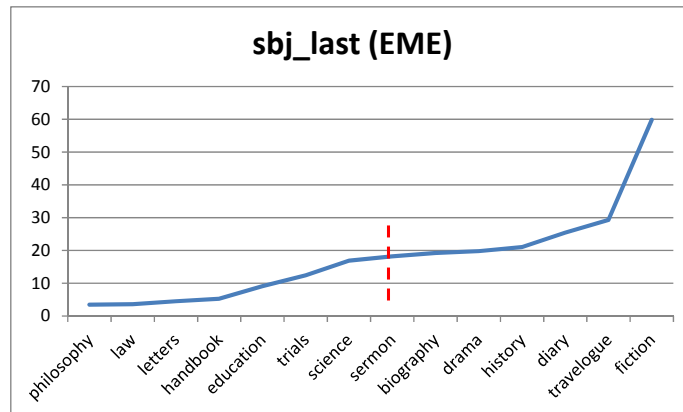
- SBJ-LAST: ME:



34

Analysis of the data

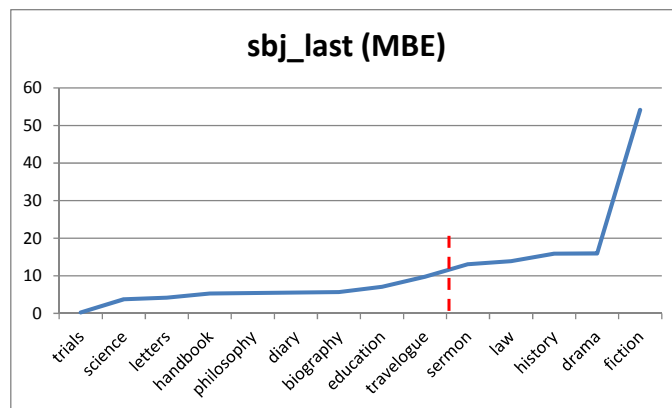
- SBJ-LAST: EModE:



35

Analysis of the data

- SBJ-LAST: LModE:



36

Analysis of the data

	LOWER PROPORTIONS	HIGHER PROPORTIONS
ME	Handbook Law Philosophy	Travelogue Romance
EModE	Philosophy Law	Fiction Diary Travelogue Drama
LModE	Science Trials Handbooks	Drama History Fiction

SBJ-LAST and registers across time

37

Analysis of the data

- SBJ-LAST:
 - **connected to subject-involvement:**
 - fewer examples in registers which prototypically avoid speaker/writer- or hearer/reader-oriented linguistic features (law, science, handbooks)
 - many examples in subject-oriented registers (travelogue, romance, fiction, diary, drama)
 - So... in favour of the connection ‘highly marked syntax of a construction’ and ‘substantive functional defining role as far as register analysis is concerned’

38

Concluding remarks

- registers as (basically) linguistic units associated with functional/textual/stylistic interpretation:
“[t]he register perspective combines an analysis of linguistic characteristics that are common in a text variety with analysis of the situation of use of the variety” (Biber and Conrad 2009: 1)

39

Concluding remarks

- **LFD** as a linguistic strategy associated with **literate registers** from ME to LModE
What about PDE?: assumed conversational character of LFD
[“Prefaces [LFD] (...) are almost exclusively conversational features (...) Prefaces are a sign of the evolving nature of conversation.” (Biber et al. 1999: 957-958)]
- **TOP** as a **literacy** strategy in ME and progressively **more textually unmarked** in ModE
- **SBJ-LAST** as a feature of subject-hearer **involvement**
- Word-order strategies can be added to the list of linguistic features on which register analysis relies.

40

Concluding remarks

- So... tentative conclusion:
connection **'highly marked syntax of a construction'**
and
'substantive functional defining role as far as register analysis is concerned'

41

Further research

- Present-Day English data
- hybrid registers, so text-by-text analysis:
 - “for some genre categories, greater linguistic differences exist among texts within the categories than across them.” (Biber and Finegan 1988: 3)
 - “texts are seldom *unitype*; text types usually appear in embedded hybridized forms, resulting in *multiple* texts” (Virtanen 2010: 58)
 - “One and the same text type can be put to use in very different genres, and one and the same genre easily manifests texts that can be related to very different types” (Virtanen 2010: 76)
- joint treatment of some registers (factor analysis / logistic regression)

42

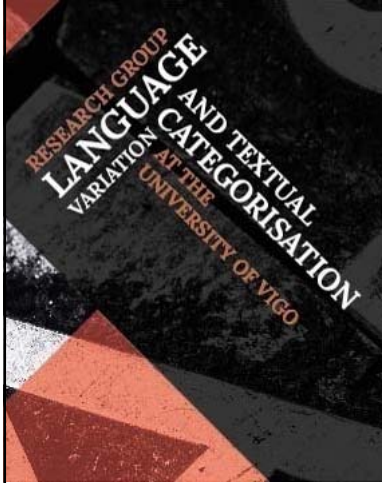
References

- Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, Douglas. 1995. *Dimensions of register variation. A cross-linguistic comparison*. Cambridge: Cambridge University Press.
- Biber, Douglas and Susan Conrad. 2009. *Register, genre, and style*. Cambridge: Cambridge University Press.
- Biber, Douglas and Edward Finegan. Adverbial stance styles in English. *Discourse Processes* 11: 1-34.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *Longman grammar of spoken and written English*. Essex: Longman.
- Birner, Betty J. 1994. Information status and word order: an analysis of English inversion. *Language* 70/2: 233-59.
- Bolinger, Dwight. 1992. The role of accent in extraposition and focus. *Studies in Language* 16/2: 265-324.
- Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles N. Li ed. *Subject and topic*. New York: Academic Press, 25-55.

- Dorgeloh, Heidrun. 1997. *Inversion in modern English. Form and function*. Amsterdam: John Benjamins.
- Dorgeloh, Heidrun and Anja Wanner. 2010. Introduction. In Dorgeloh and Wanner eds., 1-26.
- Dorgeloh, Heidrun and Anja Wanner eds. 2010. *Syntactic variation and genre*. Berlin: Mouton de Gruyter.
- Geluykens, Ronald. 1992. *From discourse process to grammatical construction: on left-dislocation in English*. Amsterdam: John Benjamins.
- Green, Georgia M. 1980. Some wherefores of English inversion. *Language* 56: 582-601.
- Gundel, Jeanette k. 1985. 'Shared knowledge' and topicality. *Journal of Pragmatics* 9/1: 83-107.
- Halliday, M.A.K. 1967. Notes on transitivity and theme in English. Part 2. *Journal of Linguistics* 3: 199-244.
- Kreyer, Rolf. 2010. Syntactic constructions as a means of spatial representation in fictional prose. In Dorgeloh and Wanner eds., 277-303.

- Kroch, Anthony and Ann Taylor. 2000. Penn-Helsinki Parsed Corpus of Middle English, second edition.
- Kroch, Anthony, Beatrice Santorini and Lauren Delfs. 2004. Penn-Helsinki Parsed Corpus of Early Modern English.
- Kroch, Anthony, Beatrice Santorini and Ariel Diertani. 2010. Penn Parsed Corpus of Modern British English.
- Martínez-Insua, Ana Elina and Javier Pérez-Guerra. 2012. Genre and word order in Modern English. Delivered at 'Grammar and Genre: interfaces and influences'. Turku, Åbo Akademi, 24-26 October.
- McCawley, James D. 1988. *The syntactic phenomena of English. Vols. 1, 2.* Chicago: The University of Chicago Press.
- Prince, Ellen F. 1997. On the functions of left-dislocation in English discourse. In Akio Kamio ed. *Directions in functional linguistics.* Philadelphia: John Benjamins, 117-144.
- Taavitsainen, Irma. 2001. Changing conventions of writing: the dynamics of genres, text types, and text traditions. *The European Journal of English Studies* 5: 139-150.

- Takahashi, Kunitoshi. 1992. Constructionally presentational sentences. *Lingua* 86: 119-48.
- Virtanen, Tuija. 2004. Point of departure: cognitive aspects of sentence-initial adverbs. In Tuija Virtanen ed. *Approaches to cognition through texts and discourse.* Berlin: Mouton de Gruyter, 78-97.
- Virtanen, Tuija. 2010. Variation across texts and discourses: theoretical and methodological perspectives on text type and genre. In Dorgeloh and Wanner eds., 53-84.



**A diachronic
constructional analysis of
English registers:
grammar and style hand
in hand**

Javier Pérez-Guerra (jperez@uvigo.es)
Ana Elina Martínez-Insua (minsua@uvigo.es)

Thank you!

Universidad de Vigo

47

The data

- LFD: CorpusSearch query:

```
node: IP*  
nodes_only: t  
ur_text_only: t  
query: (IP* Doms *-LFD)
```

48

48

The data

- TOP: CorpusSearch query (TOP_{complement}):

node: IP-MAT*

nodes_only: t

ur_text_only: t

query: (IP-MAT* iDoms NP-OB* | NP-SPR)
AND (IP-MAT* iDoms NP-SBJ*)
AND (NP-OB* | NP-SPR precedes NP-SBJ*)

- TOP: CorpusSearch query (TOP_{adjunct}):

node: IP-MAT*

nodes_only: t

ur_text_only: t

query: (IP-MAT* iDoms PP* | ADVP*)
AND (IP-MAT* iDoms NP-SBJ*)
AND (PP* | ADVP* precedes NP-SBJ*)

49

49

The data

- SBJ-LAST: CorpusSearch query:

node: IP-MAT*

nodes_only: t

query: (IP-MAT* iDomsLast NP-SBJ)
AND (NP-SBJ iDoms !PRO)

50

50

The data: LFD

- NP:

He that beleeueth on me, as the Scripture hath saide,
out of his belly shall flow riuers of liuing water.
(AUTHNEW-E2-H,VII,20J.945) [Spec resumptive]

Simon, beinge a child of six yers old, his father loved
him above all the reste, (FORMAN-E2-H,3.28) [shared
LFD?]

51

51

The data: LFD

- PP/AdvP:

But of the tree of the knowledge of good and euill, thou
shalt not eate of it: (AUTHOLD-E2-H,II,1G.155)

..

but if it worke vpon it selfe, as the Spider worketh his
webbe, then it is endlesse, (BACON-E2-H,1,20R.49)

and though he suffer'd only the name of a slave, and
had nothing of the toil and labour of one, yet that was
sufficient to render him uneasy; (BEHN-E3-H,193.231)

And as these Languages ought to be well understood,
so they shou'd be learn'd in as short a Time as may be.
(ANON-1711,3.6)

52

52

The data: LFD

- *that*-clause:

That false Locks as they call them of some Hair, being by curling or otherwise brought to a certain degree of driness, or of stiffness, will be attracted by the flesh of some persons, or seem to apply themselves to it, as Hair is wont to do to Amber or Jet excited by rubbing. Of this I had a Proof in such Locks worn by two very Fair Ladies that you know. (BOYLE-E3-H,27E.93)

That this War has been levy'd, and that Acts of Hostilities have been committed by the Prisoner, I \$do \$n't doubt but that it has been sufficiently proved to you, Gentlemen of the Jury; (TOWNLEY-1746,31.236)

53

53

The data: TOP complement

- argument NP: OB1, OB2 or SPR [direct object, indirect object or subject/object predicative]):

This and many good deedes he did to diuers.
(ARMIN-E2-H,43.295) [OB1]

male and female created he them. (ERV-OLD-1885,1,20G.66) [SPR]

54

54

The data: TOP complement

- *that*-clause:

That I had received such from Edward also I need not mention; (AUSTEN-180X,187.621)

They had a hamper and were independent of stoppages for provision, he informed her. (MEREDITH-1895,19,157.529)

55

55

The data: TOP complement

- PP:

- includes (very few) examples of complements:

In the inward Frame the various Passions, Appetites, Affections, stand in different Respects to each other. (BUTLER-1726,235.69)

To them may be applied what St. James says on a like occasion; (BURTON-1762,2,5.116)

- includes (very few) examples of subject dependents:

Of the 'ologies' very few correspond to their derivation. (BAIN-1878,375.248)

56

56