Table of Contents

Themes and theme progression in Swedish advanced learners' writing in English.......... 1
Jennifer Herriman

Existential *there* and catenative concord. Evidence from the British National Corpus.......... 29
Arne Olofsson

Recurrent word-combinations in English student essays.......... 49
Signe Oksefjell Ebeling

Spanish EFL undergraduate students' perceptions of learning styles.......... 77
Joana Salazar Noguera and Justin Wageman

Variation and language engineering in Yoruba-English code-switching 103
Emmanuel Adedayo Adedun and Mojisola Shodipe

Rosewater, *wheel of fortune*: Compounding and lexicalisation in seventeenth-century scientific texts.......... 135
Begoña Crespo

Student perceptions of English studies in Bulgaria.......... 155
Milena Katsarska and Donka Keskinova

Critical Discourse Analysis, An overview 183
Encarnacion Hidalgo Tenorio

Reviews.......... 211

About the Contributors.......... 223

Themes and theme progression in Swedish advanced learners' writing in English

Jennifer Herriman, University of Gothenburg

Abstract
This study investigates how the Themes and Theme progressions in a sample of Swedish advanced learners' writing contribute to the method of development in their texts. A comparison is made to a sample of similar writing by British university students. It was found that the advanced learners' sample contains more interactional Themes, which create a dialogic method of development similar to that found in conversational language. Themes and Theme progressions which are typical of expository writing, on the other hand, are used less frequently. These include discourse label Themes, summative progressions, which manage the build-up of information as it accumulates in the text, and split progressions, which signal hierarchical relations between parts of the informational content.

Key Words: Theme, method of development, thematic progression, information structure, learner corpora

1. Introduction

Advanced learners have reached a level where they have mastered most of the formal features of the language and make few grammatical mistakes. Nevertheless, it has been found that there are certain characteristic features in their writing that tend to distinguish it from writing by native-speaker students and give it a "non-native" sound, for instance an overuse of high frequency vocabulary, an overly spoken style and a high degree of involvement (Granger 2004:135). Another area which appears to be difficult for advanced learners is the organisation of information in the discourse (Mauranen 1993:141, Bülow-Møller 1996, Lorenz 1998, Caroll *et al* 2000:443, Callies 2009:2). This concerns, among other things, selecting what information to place in Themes (Boström Aronsson 2005). The Theme is the starting point of an utterance, "the element which serves the point of departure of the message; it is the element which locates and orients the clause within its context." (Halliday & Matthiessen 2004:64). This is followed by the Rheme, which is "part of the assembly of the new information that the text offers" (Cummings 2003:133). The choice of what to place in the

Themes of a text is not, as Halliday & Matthiessen (2004:99) point out, a "haphazard matter". It reflects the "way in which a text develops its ideas", thereby constituting what has been called the "method of development" of the text.' (Fries 1995:323). Furthermore, as the text progresses, the Themes connect to the Themes and Rhemes of preceding clauses in various ways, picking up or repeating their meanings and developing them further. These connections form patterns of thematic progression (Daneš 1974). To create a well-formed text, then, learners must select Themes which contribute successfully to its development and connect to other Themes and Rhemes in a way that controls the flow of information and creates coherence in the text.

Earlier studies of Swedish advanced learners' writing have found that it displays some of the typical characteristics of advanced learners' interlanguage. Among other things, there is a tendency to use a rather emphatic style with features typical of spoken language, in particular expressions of subjective involvement (Boström Aronsson 2005:197). This is evident, for instance, in a frequent usage of first person references, *it*-clefts (e.g. *It was Pete that told her*), which typically have a contrastive, emphasizing function, and Themes containing interpersonal items, such as *perhaps, of course* (Boström Aronsson 2005). This study investigates the Themes in Swedish advanced learners' writing with regard to how they contribute to the method of development in their texts and what patterns of thematic progression they form.

2. Material and Method

Investigating how Themes contribute to the thematic progression of a text requires in-depth manual analysis, which restricts the size of the material that can be investigated. The material examined here is a small sample of 16 essays (comprising a total of approx. 8000 words) from the Swedish component of the *International Corpus of Learner English* (ICLE) (Granger 2002), which I will refer to as the SWICLE sample. The essays have been written by Swedish students in their second year of university studies of English and they are all written on the same topic: *Man and Nature*. *How do we develop a borrower's rather than consumer's mindset*? These essays will be compared with a sample of similar size (16 essays, approx. 8000 words) from the *Louvain Corpus of Native English Essays*, which were written by British university students.

These essays were also all written on an environmental topic: *Road and Rail Transport Problems in the UK*. Although the essays in both samples are problem-solution texts, the problems addressed in these texts are slightly different. The SWICLE problem (the need to change people's attitude to their environment) is somewhat more abstract than the LOCNESS problem (the shortcomings of British road and rail transport), which is of a more practical and concrete nature. The assignments are also worded slightly differently, which may also affect how the writers construct their texts. In the SWICLE sample, the essay topic is formulated as a question appealing to a shared responsibility, represented by *we*, whereas in the LOCNESS sample the essay topic is an impersonal, informational noun phrase.

Each essay was divided into "T-units", i.e. units "slightly larger than the clause, but smaller than the sentence" consisting of "an independent clause together with all hypotactically related clauses and words that are dependent on that independent clause" (Fries 1995: 49).[1] Theme and Rheme structure can, of course be found in all types of clauses that express mood and transitivity (Halliday & Matthiessen 2004:98), and there has, therefore, been some variation in what has been treated as the unit of analysis in earlier studies of Themes and Rhemes. Some have analysed the Themes and Rhemes of all types of clauses (Francis 1989, 1990). Others have used the orthographic sentence as their unit of analysis (e.g. Hawes & Thomas 1997) or the macrosyntagm, which is a coordinate, complete main clause capable of standing as a well-formed sentence, including sentence fragments (e.g. Enkvist 1974:131, Crompton 2004:227). What is used as the unit of analysis will, of course, affect the empirical results of the investigation and the comparisons that can be made between them. One advantage of using the T-unit as the unit for the analysis is that it allows analysts to take into account whether a dependent adverbial clause has been placed in initial position in the Theme or later in the Rheme. This would be missed if analysis was carried out on the clause level only. Dependent clauses in initial position provide meaningful frames within which the rest of the sentence develops (Downing & Locke 2006:236), and if they are taken as the point of departure for the whole clause complex, the way in which

[1] Coordinated clauses with an ellipted subject in the second clause have been counted as two T-units.

Themes work to signal the 'method of development' of the text emerges more clearly (Thompson 2004:155). For this reason, those studying the thematic development of texts usually use the T-unit as their unit of analysis (Thompson & Thompson 2009:46).

In each of the T-units, the Themes were identified as the items placed in initial position. As thematic and focal prominence forms a cline which has a wave-like effect (Halliday 1994:337), the boundary between where the Theme ends and the Rheme begins is inherently indeterminate (Fries 2009:13). In this study, I will follow Halliday & Matthiessen (2004:81) and draw the boundary after the first experiential constituent, i.e. the constituent which represents a participant, circumstance or process, which Halliday & Matthiessen label the topical Theme, e.g. the adverbial, *On Friday,* in *On Friday the storm will hit Sweden.* This analysis equates the Theme with a spatial metaphor i.e. the Theme as the point of departure in the message. In some studies, (e.g. Dubois 1987, Hawes & Thomas 1997, Mauranen 1993, 1996, Martin & Rose 2007) the Theme is equated with a matter metaphor i.e. what the clause is 'about' (Gómez-González 2001:115) Here, the boundary between the Theme and the Rheme would be drawn after both the adverbial and the subject, i.e. *On Friday the storm,* as this includes what the clause is about, *the storm*.[2]

The topical theme may be preceded by items which have textual meaning, i.e. they are concerned with the text as a message, such as conjunctions and conjuncts which link the clause to the surrounding text, e.g. *but, nevertheless, therefore,* and/or interpersonal meanings, which provide an attitudinal orientation to the message, such as modal adjuncts e.g. *surely, indeed, etc.* As these items are inherently or, in the case of interpersonal items, characteristically thematic, they do not exhaust the full thematic potential of the clause, and it is only after the topical Theme

[2] Thompson & Thompson (2009:58) argue, that "if Subject is taken as Theme because it is Subject, this means that the distinction between the metafunctions in terms of their different contributions to meaning is blurred. Whereas Theme is the 'starting point of the message', Subject is the "'resting point' of the argument'; Subject 'specifies the entity in respect of which the assertion is claimed to have validity' (Halliday & Matthiessen 2004:117,118). Therefore "it would seem theoretically more secure to consider the possibility that the two threads of meaning, while both contributing to texture, operate independently, and that, in order to gain a full picture of the logogenetic growth of the text, both need to be traced separately."

appears that the thematic grounding of the message is completed (Halliday & Matthiessen 2004:85). This study focuses on the topical Theme, as it is this part of the Theme that contributes most to the method of development of a text. Interpersonal and textual components in the Swedish advanced learners' Themes will not be considered here.[3] I will first investigate the topical Themes in the two samples and then examine what kinds of Theme progressions they form.

3. Topical themes

3.1 Identifying types

A text may be seen as consisting of different levels of meaning, i.e. a propositional information content level, and a writer-reader level, where the writers interact with their readers (vande Kopple 1985, 1988, Crismore 1989 and Crismore *et al.* 1993). The Themes of a text may therefore be selected from these two levels. On the propositional information content level, the Themes refer to actions, events, states of affairs or objects in the real world outside the text.[4] On the writer-reader level, the Themes refer to the world of the text itself, by, for instance, signalling the presence of the writer and/or reader, commenting on the writing process itself or features of its structure and organisation.

In this study I have investigated to what extent the topical Themes in the two samples develop their texts on these different levels. Those which develop the propositional informational content level have been labelled informational Themes. These represent, for instance, animate or inanimate participants, as in (1) and (2), respectively, or circumstances such as spatial or temporal location, as in (3), or contingency relations (concession, reason, condition, purpose etc), as in (4).

---

[3] See Boström Aronsson (2005) for a comparison of the interpersonal and textual components of Themes in samples of the SWICLE and LOCNESS corpora.

[4] In dictionaries and encyclopedias, for instance, the Themes are usually centred round the object which is being defined or described (Matthiessen 1995:578); in guide books they are often centred round spatial locations (Matthiessen 1995:570, Fries 1995:325); in obituaries, they are typically centred on the main person and times in his or her life (Fries 1995:320), and in academic research papers, they are often abstractions, in particular nominalisations (Corbett 2008).

(1) Children should also be taught that there will be people coming after them, who will have to take care of everything we leave behind. (SWICLE)

(2) Public transport is the obvious solution to the increase in population (LOCNESS)

(3)
a. In the west we are obsessed with things. (SWICLE)
b. Today environmental issues are considered to be much more important than a couple of decades ago. (SWICLE)

(4)
a. Although catalytic converters are now compulsory on new cars, fumes in cities have become so bad that scientists have linked them to a marked increase in the number of asthma sufferers. (LOCNESS)
b. Because of dangerous UV-B rays from the sun we may have to stay in-doors. (SWICLE)
c. and if Nature is not going to fight back, we will have to change. (SWICLE)
d. To get from A to B, they stop at, and sometimes divert several miles to reach C, D and E. (LOCNESS)

On the writer-reader level, topical Themes are fusions of experiential meaning with interpersonal or textual meanings, respectively. These have therefore been divided into two categories, which I have labelled interactional and discoursal Themes, respectively.

The label interactional Theme is from Berry (1995:64), who makes a distinction between informational Themes which foreground aspects of the topic and interactional Themes which foreground the interactiveness of the discourse by referring to the writer or reader. Interactional Themes occur thus when the writers use first and second person pronouns, as in (5a) and (5b) and when the writer addresses the reader directly in e.g. rhetorical *wh*-questions, as in (5c), in imperatives, as in (5d), or in conditional clauses with *you*, as in (5e). Interactional Themes create a dialogic method of development in a way similar to the collaborative

development across turns found in interactional conversational speech (Matthiessen 1995:575).

(5)

a. I believed that I was the only one thinking about the environment and that my actions taken to improve the world were small and insignificant. (SWICLE)
b. We want to give our children the same experiences of nature as we have had ourselves; as a place for recreation and inspiration. (SWICLE)
c. What can we do with rapid economic growth if we do not have fresh air to breathe? (SWICLE)
d. Just look at the case of the planned Newbury bypass. (LOCNESS)
e. If you look around the world you will notice that countries with a clean "image", meant to attract tourists for example have a more environmental-oriented view in common. (SWICLE)

The label discoursal Theme is from Gibson (1993:324), who proposed this category for Themes which combine textual meaning with experiential meaning. Discoursal Themes are concerned with the current discourse as a text. They include instances of reflexive metadiscourse, i.e. they are used by writers to "comment on language itself, the communicative situation, or their own roles in it." (Ädel 2009:70). In (6a & b), for instance, the Themes comment explicitly on the writing process itself, and in (6c & d), they clarify the writers' purpose.[5]

(6)

a. However, writing this essay will allow me to indulge in some wishful thinking (SWICLE)
b. And as I said initially, I think we are on the right track. (SWICLE)

[5] There may, of course, be a certain degree of overlap between interactional and discoursal Themes, as in (6b) and (6d), for instance, where the writers refer to themselves in the writing process.

c. In concluding I wish to propose clean, efficient comfortable and cheap public transport for the near future. ((LOCNESS)
d. To illistrate my point, if every time you took a train, it stops for 2 hours on the track, everyone would stop taking it.[6] (LOCNESS)

Some discoursal Themes are discourse labels, or what Matthiessen (1995:569) refers to as "lexicalizations of aspects of the discourse organisation". These are abstractions such as *cause, reason*, *question, example, way, factor,* etc. (also referred to as "shell" nouns by Schmid, 2000:37*),* which encapsulate a chunk of information by conceptualising and characterising its function in the discourse (Schmid 2000:14). Some of these labels are genre-sensitive, for instance *problem,* and *solution,* which refer to parts of problem-solution texts. Mauranen (1993:79) notes that discourse labels such as these "both change perspective and provide information. The referent now consists not only of the content of the text stretch picked out by its reference, but also its status in the text. Such references make the reader aware of the text as a text (as a textual phenomenon which has parts with functions, like conclusions, examples, questions, etc.) along with the meaning that is referred to." This is illustrated by (7), where the Theme *This question* labels its immediately preceding content.

(7) What can we do with rapid economic growth if we do not have fresh air to breathe?
This question has made many people start thinking about what they can do to be more friendly to the environment. (SWICLE)

Thematic discourse labels may in some cases, as Matthiessen (1995:569) points out, be regarded as metaphorical interpretations of conjunction, i.e. they are nominalisations of conjunctive relations between stretches of text. In (8a) the Theme, *the problem,* expresses a relation of contrast similar to that expressed by the conjunctions, *but*, and in (8b) the Theme

[6] Spelling mistakes have not been corrected.

the result expresses a relation similar to that expressed by the conjunct *therefore.*

(8)

a. I do not think that people really want their children to grow up in a world where you can not drink the water from a stream in the mountains, or breath the air in the cities without a gas-mask. The problem is that we are in a period where people have not yet experienced what pollution can do to the world. (SWICLE)

b. But he sells it to smart businessmen at losing prizes and the result is that he remains the exploitation victim, unable to get out of the circle. (SWICLE)

As Schmid (2000:37) points out, discourse labels are "remarkable from the point of view of information distribution". When they occur in identifying clauses they have a strong focalising affect similar to that created by *wh*-clefts. In (9), for instance, which are both topic sentences, they are a powerful text-structuring device which marks a transition from one phase of the text to another by linking to the preceding text and predicting how the text will unfold. In (9a) the Theme *A way around these problems* signals a move in the text from the description of the problem to a discussion of a possible solution, and in (9b), the Theme *The reasons for this* links the description of the problem to a discussion of its causes.

(9)

a. A way around these problems is, however, in sight. Road transport will … (LOCNESS)

b. The reasons for this are many and varied. People of today want … (LOCNESS)

In the category of discourse Themes, I have also included existential *there*[7] and pronouns which have extended text reference, such as the demonstratives *this* and *that* (Halliday & Hasan 1976:52) as these items

[7] Berry (1995:66) regards existential *there* as a "pass option, an option not to make use of the thematic slot to foreground any particular type of meaning."

are also used to signal the organisation of information.[8] Existential *there*, as in (10), performs a presentative function for newsworthy information. When writers select *there* as the Theme of their utterances, they indicate to their readers that new information is about to be presented in the Rheme. Thus, by flagging the existence the new information at the end of the clause, *there* plays an important role in the development of a text (Martin 1995: 232).

(10) There must also be a cooperation on international level where governments draw up regulations together. (SWICLE)

Demonstrative pronouns with extended text reference "point to" a stretch of text in the surrounding discourse, as in (11), and "[i]f such an item serves as Theme, it "distils" meaning referentially and contributes a "summary" of a segment of text meaning" (Matthiessen 1995:568).

(11)

a. It is a great danger to ourselves if we continue to consider indigenous peoples and their environment as a museum. We must stop being sentimental when it comes to ecology. This is the major problem in the question of how to save our world. (SWICLE)

b. Our need and dependance on road and rail transport has created supply. The car industry is now one of the world's largest. In the last 30 years we have seen the construction of a vast motorway network throughout the U.K. All of this has come at a price though. (LOCNESS)

The distribution of these three types of topical Themes will be discussed in the next section.

3.2 Topical Themes in the SWICLE and LOCNESS samples

Table 1 compares the proportions of Informational, Interactional and Discoursal Themes in the two samples.

[8] Three instances of the pronoun *it* in *it*-clefts in the SWICLE sample have been included in this category.

Table 1. Informational, Interactional and Discoursal Themes in the SWICLE and LOCNESS samples

	SWICLE		LOCNESS	
Informational	358	65%	366	69%
Interactional	96	17%	31	6%
Discoursal	99	18%	134	25%
Total T-units	553		531	

The proportion of informational Themes in the SWICLE sample is slightly smaller than in the LOCNESS essays (65% vs 69%). There is also a difference in the type of content these informational Themes represent. In the SWICLE sample, there are more animate informational Themes (107 vs. 41) (See Table 2). There are also more Themes representing temporal (44 vs. 17) and spatial location (17 vs. 7). The method of development on the informational level of the SWICLE sample tends thus to be more often centred on humans in a temporal or spatial context in contrast to the LOCNESS sample which is predominantly centred on inanimate concepts. To a certain extent this reflects the different essay topics. The SWICLE essays are concerned with how a change is to be brought about in people's attitudes, and a change implies a temporal perspective. The LOCNESS essays deal with concrete problems concerning infrastructure.

Table 2. Informational Themes in the SWICLE and LOCNESS samples (raw frequencies)

	SWICLE	LOCNESS
Animate	107	41
Temporal Location	44	17
Spatial Location	17	7
Contingency	51	56
Other Inanimates	139	245

The proportion of interactional Themes is more than three times as high in the SWICLE sample as in the LOCNESS sample (17% vs. 6%). This

is mostly due to more *wh*-interrogatives (23 vs. 6) and 1st and 2nd personal pronouns (63 vs 16), especially inclusive *we* with generic reference (See Table 3)[9]. Using questions is a strategy often used by learners (Virtanen 1998:97), in particular to highlight a piece of information for emphasis (Callies 2009:209). The frequent usage of inclusive *we* as a personal generic is characteristic both of the high degree of involvement found in advanced learners' writing in general (Petch-Tyson 1998: 117) and of their tendency to use features of spoken lang. It also reflects a tendency for learners to write on a somewhat vague and general level, often making sweeping statements.

Table 3. Interactional Themes in the SWICLE and LOCNESS samples (raw frequencies)

	SWICLE	LOCNESS
1st and 2nd person pronouns	63	16
Imperatives	10	9
Wh-interrogatives	23	6

The proportion of discoursal Themes is lower in the SWICLE sample than in the LOCNESS sample (18% vs. 25%). This is mainly due to fewer discourse labels (54 vs. 70) and summative demonstrative pronouns (17 vs. 35), whereas there is no great difference in the numbers of the other types (i.e. writers' comments on the writing process and existential *there*, See Table 4).

[9] In a comparison of the usage of *we* in the SWICLE and LOCNESS corpora (Herriman 2009), it was found that generic *we* was used four times more frequently by the Swedish students.

Table 4. Discoursal Themes in the SWICLE and LOCNESS samples (raw frequencies)

	SWICLE	LOCNESS
Discourse Labels	54	70
Summative demonstratives	17	35
Existential *there*	18	19
Writing process	7	10
It in *it*-clefts	3	

In sum, the Swedish advanced learners tend to use more interactional and fewer informational Themes than the native writers. This reflects the tendency for advanced learners to use a subjective style of writing. It may also be partly due to the wording of the essay topics. In the LOCNESS sample, the essay topic clearly elicits a problem and solution text, whereas in the SWICLE sample, it is formulated as a question and therefore invites a dialogic perspective with interactional Themes. The advanced learners also use fewer discoursal Themes of the type which condense and label chunks of information in the text. In contrast to the SWICLE essays, then, the method of development in the LOCNESS essays tends to be more informational and discoursal.

These findings suggest, then, that, in contrast to the native speakers, the Swedish advanced learners are less aware of how to use Themes which create a monologic method of development of the type normally found in expository writing. I will now go on to examine how the Themes in the SWICLE and LOCNESS samples form Theme progressions.

4 Thematic progression in the SWICLE and LOCNESS samples

4.1 Background

Thematic progression is concerned with connectivity between information in Themes and other parts of the text. Three main types of thematic progressions may be formed, as the following examples from

Daneš (1974:119) illustrate,[10] a linear progression where the Theme relates to the Rheme of the preceding utterance, as in (12), a constant progression where the same Theme appears in a series of utterances, as in (13), and a derived pattern, where successive Themes are derived from a common topic, a hypertheme (of a paragraph, or other text section), as in (14), where the hypertheme is *New Jersey*.

(12) Linear
The first of the antibiotics was discovered by Sir Alexander Fleming in 1928. He was busy at the time investigating a certain species of germ which was responsible for boils and other troubles.

(13) Constant
The Rousseauist especially feels an inner kinship with Prometheus and other Titans. He is fascinated by any form of insurgency…

(14) Derived from hypertheme
New Jersey is flat along the coast and southern portion; the northwestern region is mountainous. The coastal climate is mild, but …..

According to Enkvist (1974), linear progressions create a "dynamic" style of writing and constant progressions a "static" style of writing. Certain correlations have also been found between these progressions and text type. Linear progressions occur frequently in expository and argumentative texts, for example in editorials (Francis 1989, Hawes & Thomas 1996) and popular medical texts (Nwogu & Bloor 1991). Constant progressions occur frequently in narratives (Fries 1995, Wang 2007) and in news stories (Francis 1989 & 1990, Gómez (1994). The third type of progression, the derived from hypertheme progression, deals with sets of interrelated ideas and therefore relies largely on shared knowledge. This kind of progression has been found to be usual in legal

[10] Daneš, as Crompton (2004:200) points out, had a combining approach to Theme, i.e. Theme was identified contextually as given information, not syntactically by its position.

texts (Kurzon 1984) and professional medical texts (Nwogu & Bloor 1989). Identifying derived from hypertheme progressions is, however, somewhat problematic as there are no criteria for an unambiguous distinction between this and constant or linear progressions which are indirectly connected to each other by relations of collocation or other associative relations. Leong Ping (2005:713) points out, for example, that the Themes, *the corner kick* and *the curling ball,* in (15) are, on the one hand, connected to a common topic, in this case the football game and may therefore be regarded as a derived from hypertheme progression. On the other hand, both of the Themes are also collocationally associated to each and may therefore also be regarded as a constant progression.

(15) The corner kick was taken skilfully, and the curling ball soon found the back of the net.

As both views are equally possible, it is not clear how these two patterns are to be distinguished from each other. Some studies (e.g. Dubois 1987, Mauranen 1996, McCabe 2004, Leong Ping 2005) have therefore subsumed Daneš' derived from hypertheme progression under linear and constant progressions. In this study, I will do the same.

Some Themes form progressions which connect in hierarchical relations of superordination or subordination to other Themes and Rhemes. McCabe (1999:181) observed a summative Theme progression which summarises a stretch of the preceding text, as in (16), where *the social effects* is a summary of the preceding content.

(16) Even more significant than the social effects were the psychological consequences.

Daneš (1974:121) and McCabe (2004) identified a split Rheme progression where a Rheme is expanded by a series of subordinate Themes, as in (17).

(17) All substances can be divided into two classes: elementary substances and compounds. An elementary substance is a substance which consists of atoms of only one kind … A compound is a substance which consists of atoms of two or more different kinds … (Daneš 1974:121).

Similarly McCabe (1999:175) observed a split Theme pattern where a Theme is expanded by a series of subordinate Themes, as in (18).

(18) The upward movement of wages and the downward price of cereals led….
Better wages in both town and countryside enabled the population to …
While the price of wheat fell, wine, beer, oil, butter, cheese, meat, fruit, …

Split Theme and Rheme progressions such as these are expository in character as they list related points of information which illustrate a main point.

Some studies have included gapped progressions. They differ in how far these gaps extend. McCabe (1999) for instance, delimits progressions to connections with a maximum gap of three clauses, whereas Dubois (1987:94) counts all connections as progressions, even those 80 clauses apart. In this study I will, however, only consider connections between contiguous T-units as linear or constant progressions. The Themes which connect to earlier parts of the text will be labelled "Back". Themes, with the exception of Themes in split and summative progressions, as these form hierarchical relations of subordination and superordination over longer stretches of text.

Themes which do not form progressions at all have been regarded as "unmotivated Themes" and consequently signs of bad writing (Mauranen 1993). It has been pointed out by McCabe (1999:180-189), however, that even though these Themes do not form progressions, some of them do, nevertheless, contribute to the development of the text in various ways. These fall into four main categories: pragmatic Themes, which are key concepts frequently referred to in the text, grammatical Themes, i.e. dummy subjects *it* and *there*, extralinguistic Themes, which are situationally evoked, e.g. *we* referring to the writer and reader, or imperatives addressing the reader, and metatextual Themes, which are Themes which refer to text itself, e.g. *chapter, figure,* etc. In this study, I will divide the Themes which do not connect to the preceding text into two categories depending on whether they represent new information or may be derived from the context (labelled New and Contextual Themes, respectively).

The distribution of these different types of theme progressions in the SWICLE and LOCNESS samples will be discussed in the following section.

4.2 Thematic progressions in the SWICLE and LOCNESS samples

Table 5 compares the Themes which form progressions in the two samples.

Table 5. Theme Progressions in the SWICLE and LOCNESS samples

	SWICLE		LOCNESS	
Constant	109	20%	90	17%
Linear	180	32%	156	29%
Summative	14	3%	38	7%
Split Theme	0	-	24	4%
Split Rheme	15	3%	38	7%
Total	318	58%	346	65%

In both samples, nearly two thirds of the Themes (58% in SWICLE and 65% in LOCNESS) form progressions. Linear progressions are most usual. These are formed by about a third of the total number of Themes (with slightly more, 32%, in SWICLE than in LOCNESS, 29%). Linear progressions tend to predominate in the passages of the essays which describe causes of problems and their effects, as in (19).

(19) The basic dilemma facing the UK's rail and road transport system is the general rise in population. This (LINEAR) leads to an increase in the number of commuters and transport users every year, consequently putting pressure on the UKs transports network. The biggest worry to the system (LINEAR) is the rapid rise of car users outside the major cities. Most large cities (LINEAR) have managed to incourage commuters to use public transport thus decreasing major conjestion in Rush hour periods. (LOCNESS)

Constant progressions were formed by about a fifth of the Themes in both samples (with slightly more in SWICLE, 20%, than in LOCNESS, 17%). These progressions develop the text from the same angle by adding new information in each sentence. Others form pairs of contrastive Themes, juxtaposing two members of a set of possible candidates, as in (20).

(20) People living in big cities do not see any change in their lives, whether the nature around the city is damaged or not, because they will got their everyday goods transported to them anyway. Those on the other hand who live in close contact with Nature and depend on it for their survival (CONSTANT) have another approach to the environment. (SWICLE)

Series of successive constant progressions were found in the SWICLE sample only. These formed narrative or descriptive passages, or lists of proposed solutions, such as (21).

(21) So what we have to do ti change our mindset to a “borrower’s”, is first of all to put pressure on all polluting companys and make them either re-cycle or find non-polluting alternatives. Secondly we (CONSTANT) have to fight poverty in order to save the land from deforestation and desertification. We (CONSTANT) need to educate the poor people how to use their land without destroying it. And then we (CONSTANT) have the most difficult task: the one of changing peoples view of money. We (CONSTANT) have to make everyone see that you can not eat, breath or drink money. (SWICLE)

Summative progressions were formed by 3% and 7% of the Themes in SWICLE and LOCNESS samples, respectively. This is illustrated by (22) where the Theme, *the way of consuming and hurrying through our lives,* encapsulates the content of the preceding chunk of text

(22) In the west we are obsessed with things and the ironic saying: ‘The one who has got the most things when he dies, wins’ tells us a lot about what we believe in. It is also an alarm bell trying to make us realize how crazy we act. Our way of consuming and

hurrying through our lives (SUMMATIVE) must come to an end and we have to understand the consequences of our actions. (SWICLE)

Summative progressions allow a cumulative "compacting" of the text by nominalising chunks of information. The lower proportion of summative progressions in the advanced learners' sample may be due to the fact that this kind of progression often involves nominalisation of the propositional information content. As nominalisation increases the level of abstraction in the text (Martin 1992:407), it tends to be acquired later in language development and therefore it is perhaps not surprising that it is used less by learners writing in a second language. A similar tendency has been observed in the English used by Danish speakers. In a study of the language used in negotiations, (Bülow-Møller 1996) found that the Danish speakers lacked a ready ability to condense information in abstract superordinate terms, and therefore had difficulties "in creating both local and global anaphoric and cataphoric connections and in contrasting old and new information by reduction or focus to forge links in their discourse." (1996:38). A further reason why the Swedish learners use fewer nominalisations may, be transfer from Swedish. Nordrum (2007) found that nominalisations are frequently translated into clauses with the generic subjects (*man* "one" or *vi* "we") in Swedish. If nominalisations are less usual in Swedish, this would also explain why the Swedish learners prefer to select personal generics as their Themes rather than nominal Themes which encapsulate information and label its discourse function.

Split Rheme progressions were formed by 3% and 7% of the Themes in the SWICLE and LOCNESS samples, respectively. Split Theme progressions, on the other hand, were found in the LOCNESS sample only (4%). Split progressions organise information into subordinate parts. In (23), for instance, the discoursal Themes, *one* and *another reason,* form a split Rheme progression which expands on the Rheme *many reasons for this,* and in (24) the informational Themes *the carrot* and *the stick* form a split Theme progression which expand on the Theme, *A 'carrot and stick' approach*.

(23) The forests, especially in eastern Europe, are dying. There are many reasons for this: one (SPLIT RHEME) is the extensive use

of the car, another reason (SPLIT RHEME) is that many of the power-plants are run on coal or oil. (SWICLE)

(24) A 'carrot and stick' approach should probably be used. The carrot (SPLIT THEME) would be cheaper fares on buses, trams and trains and improved services, together with an advertising campaign to try to persuade people of the benefits of public transport. The stick (SPLIT THEME) would be making cars more expensive to run, either by increasing road tax or increasing the duty on petrol. Neither measure would be popular, but in the end they would help to create a better transport system in Britain. (LOCNESS)

In contrast to the advanced learners, the LOCNESS writers tend thus to use summative and split progressions more frequently as a means of organising the content of their essays. They are often used to scaffold the content of longer stretches of text, as in (24) where a split Rheme progression, and a summative progression organise the discussion of a number of proposed solutions presented in the Rheme of the topic sentence at the beginning of this stretch of text (*three main ways of doing this: sharing cars, using public transport and walking or cycling*).

(25) So, the only way around the problem is to have less cars. There are three main ways of doing this: sharing cars, using public transport and walking or cycling.
Sharing cars (SPLIT RHEME) seems feasible, but is ….
Buses (SPLIT RHEME) are almost without exception, old. This makes them … Trains have similar problems. They don't ...
People walking or using a bike (SPLIT RHEME) are not protected from the elements and, …
These methods of transport (SUMMATIVE) are slow, and allow you to carry much less. (LOCNESS)

Similarly, in (25) the themes, *Solutions for these problems in Britain* and *The problem with these schemes*, are summative progressions in topic sentences at the beginning of each paragraph. These label the content of the preceding paragraphs (*problems* and *schemes*) and predict the content of the paragraph which follows (*solutions* and *problem*), thereby

providing a transition between the different parts of these problem-solution texts.

(26)
a. Solutions for these problems in Britain (SUMMATIVE) have been fairly unimaginative and unsuccessful. Park and ride schemes are … (LOCNESS)
b. The problem with these schemes (SUMMATIVE) is that they are unpopular …(LOCNESS)

The remaining Themes (42% in SWICLE and 35% in LOCNESS) are "Back", Contextual and New (see Table 6).

Table 6. Back, Contextual and New Themes in the SWICLE and LOCNESS samples

	SWICLE		LOCNESS	
Back	74	13%	53	10%
Contextual	103	19%	82	15%
New	58	10%	50	9%
Total	235	42%	185	35%

"Back" Themes do not connect to their immediately preceding T-unit but reintroduce meanings mentioned earlier in the essays. Back Themes are found, for instance, at the beginning of the concluding paragraphs to remind the reader of the essay topic, as in (26) and after intervening explanations or evaluations, as in (27).

(27) In order to develop people's minds towards "borrowers" (BACK) society must be permeated with ecological awareness. (SWICLE)

(28) These are the common opinions and arguments when green consumers describe their difficult situation and they explain why well-meaning people sometimes fail to fulfil their obligations as green consumers. Unfortunately this is the stage where most of us stagnate in our efforts to become environmentally conscious

and responsible people. Being a green consumer (BACK) is a step in the right direction, but it is not enough to solve the problem of our planet's fate. (SWICLE)

Back Themes are somewhat more usual in the SWICLE sample than the LOCNESS sample (13% vs. 10%). This is partly due to a tendency for the SWICLE writers to intersperse their argumentation with interactional rhetorical questions, generalisations with *we, you* and evaluative comments, as in (27) above.

Contextual Themes represent information which is situationally evoked (Prince 1981:236) and therefore does not interrupt the flow of information. These include deictic references to time and place, e.g. *now, here, today,* references to the writer, reader(s) or to the writing process itself. In this category I have also included the nonrepresentational subjects *there* and *it* and personal generics, e.g. *people*, *we, man, nobody.* There is a slightly greater proportion of contextual Themes in the SWICLE sample (18% vs. 15%). This is mainly due to the frequent usage of the personal generic *we.*

New Themes, finally, are either brand-new and have to be created by the reader, as in (29) or "Unused", i.e. assumed to be part of the readers background knowledge but introduced into the discourse for the first time, as in (30) (Prince 1981:235). New Themes may also be unknown information represented by the *wh*-words in *wh*-interrogatives, as in (5c) above.

(29) A recent survey in the Observer calculated that these measures, and many more, could be paid for out of the government's road budget, and enough would be left to maintain all existing roads. (LOCNESS)

(30) Bangkok, the capital of Thailand is today the most polluted capital in the world. (SWICLE)

When new information is placed in the Theme, it is backgrounded, and, as a result, more difficult to challenge. In this way, the writer is able to "smuggle" more information into the message and to present it as shared knowledge, as in (31)

(31) But with petrol companies unwilling to investigate water and alcohol powered cars for fear of losing petrol sales these still seem a long way off as our everyday form of transport. (LOCNESS)

New Themes are more or less equally usual in both samples (10% vs. 9%). There is however a qualitative difference. In the SWICLE sample, a large number of these (22 instances) are *wh*-words in rhetorical questions, whereas in the LOCNESS samples most new or "Unused" Themes are informational, as in (29) and (31).

In sum, linear progressions predominate in both samples, which is to be expected in argumentative texts such as these (Crompton 2004:237). Both linear and constant progressions are slightly more usual in the SWICLE sample as well as Themes which do not form progressions (Back, Contextual Themes). They also contain fewer summative and split Theme/Rheme progressions. The advanced learners' differ thus from the LOCNESS writers in that tend to use fewer progressions of the kind typically found in expository texts, i.e. progressions which form hierarchical relations between parts of its informational content.

5. Conclusion

As Thompson points out (2004:165), in order to build up the framework of a text, Themes need to 1) signal what the speaker thinks is a suitable starting point; 2) signal the maintenance or progression of 'what the text is about'; 3) specify or change the framework for the interpretation of the following clause (or clauses) and 4) signal boundaries between sections of the text. It has been found here that, when it comes to signalling a suitable starting point and 'what the text is about', the Swedish advanced learners, in contrast to the native speakers, tend to frequently select an interactional starting point, thereby creating a dialogic method of development similar to that found in conversational language. As a dialogic method of development foregrounds interpersonal meaning (Martin 1992:437 & 448), it follows, then, that their essays contain fewer informational Themes which signal the content of the text. It has also been found that the Swedish advanced learners select fewer Themes and Theme progressions which specify the framework of their texts and signal boundaries between its sections, i.e. they use fewer discoursal

Themes, and fewer summative and split Theme progressions. It has been pointed out earlier that one of the reasons for the non-native soundedness of Swedish advanced learners' writing is their lack of awareness of register differences, which results in a tendency to use an overly informal and colloquial style of writing (Altenberg and Tapper 1998:92, Boström Aronsson 2005: 188-190). One remedy for this is, as Altenberg and Tapper have pointed out, more extensive training in writing expository texts. The results of this study have shown that this should aim at increasing the advanced learners' understanding of how Themes and Theme progressions contribute the method of development in texts. They would, in particular, benefit from an increased awareness of how Themes and Theme progressions may be used to manage the logogenetic build-up of information as it accumulates in their texts.

References

Ädel, A. 2009. "Just to give you a kind of a map of where we are going: A taxonomy of metadiscourse in spoken and written academic English." *Nordic Journal of English Studies* 9:2, 41-68.

Altenberg, B. & M. Tapper. 1998. "The use of adverbial connectors in advanced Swedish learners' written English." In S. Granger (ed.) *Learner English on Computer*. London/New York: Longman. 80-93.

Berry, M. 1995. "Thematic options and success in writing." In Ghadessy, M. (ed.) *Thematic Development in English Texts*. London & New York: Pinter. 55-84.

Boström Aronsson, M. 2005. *Themes in Swedish Advanced Learners' Written English*. Unpublished PhD thesis: University of Gothenburg.

Bülow-Møller, A. M. 1996. "Control from the background: a study of information structure in native and non-native discourse." *International Journal of Applied linguistics* 6: 1, 21-42.

Callies, M. 2009. *Information Highlighting in Advanced Learner English*. Amsterdam/ Philadelphia: John Benjamins.

Carroll, M. & J. Murcia-Serra, M. Watorek, A. Bendiscioli. 2000. "The relevance of information organization to second language acquisition studies." *SSLA* 22. Cambridge University Press. 441-466.

Corbett, J. 2009. "Theme, field and genre: thematic realisations in academic articles and their popularisations." In Forey, G. & G. Thompson (eds.) *Text Type and Texture*. London: Equinox. 70-93.

Crismore, A. 1989. *Talking with Readers: Metadiscourse as Rhetorical Act*. New York: Peter Lang.

Crismore, A, R. Markkanen & M. Steffensen. 1993. "Metadiscourse in persuasive writing: A study of texts written by American and Finnish university students." *Written Communication* 10: 39-71.

Crompton, P. 2004. "Theme in discourse. 'Thematic progression' and 'method of development' re-evaluated." *Functions of language* 11:2, 213-249.

Cummings, M. 2003. "The role of Theme and Rheme in contrasting methods of organization of texts." In Butler, C. (ed.) *Dynamics of Language Use: Functional and Contrastive Perspectives.* Amsterdam/Philadelphia: John Benjamins. 129-154.

Daneš, F. 1974. "Functional sentence perspective and the organization of the text." In Daneš, F. (ed.) *Papers on Functional Sentence Perspective*. Mouton The Hague, Paris. 106-127.

Downing, A. & P. Locke. 2006. *English Grammar. A University Course*. 2nd edn. London & New York: Routledge.

Dubois, B.L. 1987. "A reformulation of thematic progression typology." *Text* 7.2, 89-116.

Enkvist, N.E. 1974. "Theme dynamics and style: an experiment." *Studia Anglica Posnaniensia* 5: 127-135.

Francis, G. 1989. "Thematic selection and distribution in written discourse." *Word* 40, 201-221.

Francis, G. 1990. "Theme in the daily Press." *Occasional Papers in Systemic Linguistics* 4, 51-87.

Fries, P. H. 1995 "Themes, methods of development and Texts." In Hasan, R. & P. Fries (eds.) *On Subject and Theme: From the Perspective of Functions in Discourse*. Philadelphia: John Benjamins. 317-359.

Fries, P. H. 2009. "The textual metafunction as a site for a discussion of the goals of linguistics and techniques of linguistic analysis." In Forey, G. & G. Thompson *Text Type and Texture*. London: Equinox. 8-44.

Gibson, T.R. 1993. *Towards a Discourse Theory of Abstracts and Abstracting* (monographs in Systemic Linguistics 5) Nottingham. Dept of English Studies, University of Nottingham.

Gómez, A. 1994. "The relevance of Theme in the textual organization of BBC news reports." *Word* 45.3, 293-305.

Gómez-González, M. Á. 2001. *The Theme-Topic Interface. Evidence from English.* Amsterdam/ Philadelphia: John Benjamins.

Granger, S. (ed.) 1998. *Learner English on Computer*. London/New York: Longman.

Granger, S. 2002. *The International Corpus of Learner English*: Handbook and CD-ROM. Louvain-la-Neuve: Presses Universitaires de Louvain. Available from http://www.i6doc.com

Granger, S. 2004. "Computer learner corpus research: current status and future prospects." In Connor, U. & T. Upton (eds.) *Applied Corpus Linguistics: A Multidimensional Perspective.* Amsterdam/New York: Rodopi. 123-145.

Halliday, M.A.K. & R. Hasan. 1976. *Cohesion in English*. London: Longman.

Halliday, M.A.K. 1994. *An Introduction to Functional Grammar*. 2nd edn. London: Arnold.

Halliday M.A.K. & C.M.I.M. Matthiessen. 2004. *An Introduction to Functional Grammar* 3rd edn. London: Arnold.

Hawes, T. & S. Thomas. 1997. "Problems of thematisation in student writing." *RELC Journal* 28:35, 35-54.

Herriman, J. 2009. "The use of *we* in argumentative essays by Swedish advanced learners of English." In Bowen, R., M. Mobärg & S. Ohlander (eds.) *Corpora and Discourse—and Stuff. Papers in Honour of Karin Aijmer*. Gothenburg Studies in English 96. Acta Universitatis Gothoburgensis. 105-120.

Kurzon, D. 1984. "Themes, hyperthemes and the discourse structutre of British legal texts." *Text* 4:1-3, 31-55.

Leong Ping A. 2005. "Talking themes: the thematic structure of talk." *Discourse Studies* 7: 6, 701-732.

Lorenz, G. 1998. "Overstatement in advanced learners' writing: stylistic aspects of adjective intensification." In S. Granger (ed.) *Learner English on computer*. London: Longman 53-66

Martin, J. 1992. *English Text*. Amsterdam: John Benjamins.

Martin, J. 1995. "More than what the message is about: English Theme." In Ghadessy, M (ed.) *Thematic Development in English Texts*. London & New York: Pinter. 223-258.

Matthiessen, C.M.I.M. 1995. *Lexicogrammatical Cartography: English Systems*. International Language Sciences Publishers: Tokyo.

Mauranen, A. 1993. *Cultural Differences in Academic Rhetoric. A Textlinguistic Study*. Frankfurt am Main: Peter Lang.

Mauranen, A. 1996. "Discourse competence—Evidence from thematic development in Native and Non-native texts." In Ventola, E. & A. Mauranen (eds.) *Academic Writing. Intercultural and Textual Issues*. Amsterdam/Philadelphia: John Benjamins. 195-239.

McCabe, A. 1999. *Theme and Thematic Patterns in Spanish and English History Texts*. Unpublished PhD Dissertation, Aston University.

Nordrum, L. 2007. *English Lexical Nominalisations in a Norwegian-Swedish Contrastive Perspective.* Unpublished doctoral dissertation, University of Gothenburg.

Nwogu K. & T. Bloor. 1989. "Thematic progression in professional and popular medical texts." In Ventola, E. (ed.) *Functional and Systemic Linguistics. Approaches and Uses*. Berlin & New York: Mouton de Gruyter. 367-384.

Petch-Tyson, S. 1998. "Writer/reader visibility in EFL written discourse." In Granger, S. (ed.) *Learner English on Computer*. London: London. 107-118.

Prince, E.F. 1981. "Toward a taxonomy of given-new information." *Radical Pragmatics*, (ed) Cole, P. New York: Academic Press. 223-255.

Schmid, H. 2000. *English Abstract Nouns as Conceptual Shells*. Berlin & New York Mouton de Gruyter.

Thompson, G. 2004. *Introducing Functional Grammar*. 2nd ed. London: Arnold.

Thompson, G. & S. Thompson. 2009. "Theme, subject and the unfolding text." In Forey, G. & G. Thompson *Text Type and Texture*. London: Equinox. 45-69.

vande Kopple, W. J. 1985. "Some exploratory discourse on metadiscourse." *College Composition and Communication*, 36: 82-93.

vande Kopple, W. J. 1988. "Metadiscourse and the recall of modality markers." *Visible Language* XXII: 233-272.

Wang, L. 2007. "Theme and Rheme in the thematic organization of texts: implications for teaching academic writing." *Asian EFL Journal* 9:1. Article 9. http:/www.asian-efl-journal.com/March_07_Iw.php

Virtanen, T. 1998. “Direct questions in argumentative student writing.” In S. Granger (ed.) *Learner English on computer*. London: Longman. 94-106.

Existential *there* and catenative concord. Evidence from the British National Corpus

Arne Olofsson, University of Gothenburg

1. Introductory[1]

1.1 Existential sentences

Like many other languages, English has a presentative construction that mainly serves to create end-focus. A subject that contains new information (normally expressed by an indefinite noun phrase) is given a position after the verb and the typical subject slot is filled by *there*.[2] The result is, for instance, *There is a teacher in the classroom* in preference to *A teacher is in the classroom*. Syntactically, *there* behaves like a subject, for instance occurring in postverbal position in a direct question (*Is there ...?*).

1.2 Subject–verb concord

There is general agreement that in the existential construction, the predicate verb (prototypically a form of *be*) shows concord with the notional (postponed) subject, at least in formal Standard English:

(1) *There* ***is*** *a book on the table*
(2) *There* ***are*** *some books on the table*

---

[1] A similar analysis based on CobuildDirect and limited to *seem* is to be found in Olofsson (2007). For an analysis of the concord patterns of the basic *there* + *be* construction in contemporary spoken British English, see Martinez Insua & Palacios Martinez (2003) and Breivik & Martinez-Insua (2008). Spoken Canadian English has been similarly analysed by Meechan & Foley (1994).

[2] Often in the past, this *there* has been regarded as an empty place-holder. However, as noted by e.g. Breivik & Martinez-Insua (2008:356) it can also be described as a presentative signal "to tell the addressees that they must be prepared to direct their attention towards an item of new information."

In less formal English, many speakers use and accept the singular verb, but only if it is contracted with *there*: *There's some books on the table*. This contracted invariable form has often been referred to as "frozen *there's*". From a communicative point of view it has also been called a "fixed pragmatic formula", for instance by Breivik & Martinez-Insua (2008:361).

Apart from the formal/informal dimension, the situation with *be* is fairly clear-cut in the sense that it is possible to claim that the postponed element is "really" the subject of the clause in which it is positioned.[3] Concord in this type of construction has been extensively studied and reported on, for instance by Meechan & Foley (1994), Martinez Insua & Palacios Martinez (2003) and Breivik & Martinez-Insua (2008).

1.3 Catenative constructions

The prototypical form of the existential sentence has factual meaning. Modality, for instance possibility or probability, can be expressed by means of modal auxiliaries (*can, may*), but from the point of view of concord, such constructions are of no interest. Alternative modes of expression involve catenative verbs (e.g. *appear* or *seem*), which are full verbs in the sense that they show concord distinctions in the present tense.[4]

---

[3] Meechan & Foley (1994:82) claim that "orthodox" concord based on the number of the notional subject in this construction is just an artefact: "We assert that these prescriptive rules of grammar taught in high school ... serve to obscure the fact that nonconcord is the norm."

[4] There are different definitions of the class of catenatives. Two major grammars, Quirk et al. (1985) and Huddleston & Pullum (2002), represent two extremes. According to Huddleston & Pullum (2002:1177), any verb that takes a non-finite complement, be it gerund-participle or infinitive, with or without *to*, is a catenative verb. Quirk et al. (1985:236) declare: "We may use the expression 'marginal auxiliaries' to apply to all four categories (marginal modals, modal idioms, semi-auxiliaries, and catenative verbs) discussed in 3.40–51." In the passage referred to (1985:146), the catenative category is narrowed down even further: "The term CATENATIVE will in practice be used to denote verbs in such constructions as *appear to, come to, fail to, get to, happen to, manage to, seem to, tend to,* and *turn out to* followed by the infinitive." In Biber et al. (1999), the term is not used at all.

One important function of *appear* and *seem* is that of hedging, as mentioned briefly by Biber et al. (1999:944). Sometimes this kind of hedging is a sign of excessive politeness and writers can use it for humorous effect, as in the following lines, spoken by a butler to his employer:

(3) "I'm very sorry, sir, but there appears to be a tiger in the dining room. Perhaps you will permit the twelve-bores to be used?" (BNC AR3 384)[5]

When catenatives are used, the syntax, including concord, is more complicated.[6] As noted by Huddleston & Pullum (2002:242), the noun phrase in question is not located inside the clause with the finite verb, so it is difficult to argue that it governs the choice of form (singular/plural) there. Instead, in their view, *there* can be seen as "inheriting" the number of the noun phrase, and this inheritance is decisive for the choice of verb form after the introductory subject.

Contrary to the situation with "frozen *there's*", which speakers can get away with, there is no accepted frozen **thereseems* or **thereappears*, at least not yet, but in spite of that, unorthodox choice of singular verb forms with plural notional subjects is well attested in both spoken and written native production. According to Quirk et al. (1985:1406), the catenative verb in existential sentences "often agrees with the notional subject in number", but the authors go on to say that "often informally" the singular of the verb is used also with a plural subject.[7] The question asked in this article is the same as in Olofsson (2007), though the answer is now sought in a different, larger material, viz. "How often?" (which implies "How acceptable is the construction?" and "Should learners be penalized for using it?").

---

[5] The source is *The Remains of the Day* by Kazuo Ishiguro.

[6] Breivik (1983), a standard work for existential *there*, presents (1983:158) the examples *there appears to be no one else* and *there appear to be one million policemen outside* (from the Survey of English Usage) without commenting on the S–V concord complication.

[7] Thagg Fisher (1985:54) categorically condemns the construction: "… there can be no doubt about the error status of the following types of singular verb concord: … there seems [seem] to be animals crawling under it and everywhere". (Square brackets are her way of introducing the correct form.)

2. Material and method

The answers to the questions asked above will be based on the entire British National Corpus (BNC), which consists of about 100 million words of British English, about 90 per cent written and 10 per cent spoken.[8] The material was searched for strings containing *there* in conjunction with present-tense forms of *appear* and *seem* followed by *to be* or *to have been*.[9] The overwhelming majority of the instances contain *be* in the sense of "exist", but there are occasional findings of auxiliary functions (bold added):

(4) ... there seems to **be** some kind of disturbance **going** on ... (EFP 843)

(5) Within the context of the Mysteries, there seems to have **been enacted** a Sacred Marriage ... (G3C 132)

From a methodological point of view it seems worth mentioning that the statistics automatically generated in a search are not quite the whole truth. For absolute precision, ocular inspection of the matches returned is necessary. Consider the following examples of spurious instances of *there seem to* (double slashes and bold added):

(6) Eliot hailed from St Louis, but the Eliots **there // seem to** have regarded themselves as Bostonian ... (A05 237)

---

[8] The entire corpus was needed to provide a sufficient number of instances. In contrast, for Martinez Insua & Palacios Martinez (2003) and Breivik & Martinez Insua (2008) subcorpora were sufficient for their investigations of the much more frequent constructions with *be*.

[9] *Seem* is the best attested catenative verb in the corpus (about 30,000 of the 60,735 instances of the lemma *seem*). There are 30,513 instances of the lemma *appear*, but only slightly more than a third of them represent the catenative function. The actual strings of *there* + present tense *seem* + present or perfect infinitive of *be* number 65 in the spoken part of the BNC, i.e. 6.5 per million words; in the written part the corresponding figure is 583, i.e. practically the same pmw. The corresponding strings for catenative *appear* are less than half as frequent.

(7) ... to note how indiscriminately trigger-happy so many people over **there // seem to be** ... (CJP 215)

The obvious span of the search is just two positions, with *there* and *seem(s)* or *appear(s)* next to each other, but, for good measure, in all the queries on the catenatives, the window was expanded to include the positions –2 and –3, thus catching cases such as (7)–(9).

(8) Nevertheless, **there** still **appear** to be notable differences … (J7D 153)

(9) **There** thus **appear** to be two relationships … (HPM 14)

(10) **There** also sometimes **seems** to be a feeling that … (EX6 1321)

In contrast to Olofsson (2007), instances with present-tense *do*-constructions (emphatic and question-forming) were not included. Not that there are many of them, but their exclusion makes for more clear-cut statistics, where present-tense *appear* and *appears* (or *seem* and *seems*) are the only contrasting elements.

The focus of the investigation is on constructions with plural notional subjects.[10] The percentages of non-concord indicated in the results below stand for the proportions within this subset, not for the relation to the total number of *there*-constructions with the catenative in question.[11]

[10] As noted by e.g. Meechan & Foley (1994:72, 75–76) and Olofsson (2007), the phenomenon under scrutiny is unidirectional in the sense that the verb often varies with plural notional subjects but hardly ever with singular subjects.

[11] In this respect, my investigation resembles that by Meechan & Foley (1994) rather than those by Martinez Insua & Palacios Martinez (2003) and Breivik & Martinez Insua (2008), who have chosen the equivalent of the latter principle. This means that our results cannot be compared, apart from very general tendencies.

3. Results

3.1 Introductory

In order to set the stage, the following table presents a few pairs of the corpus cases with *seem*, chosen for being as parallel as possible in structure, the difference being the choice between an unexpected singular and an orthodox plural form of the catenative verb. As in the fuller presentation that will follow, the parallelism as often as possible includes determiners of the notional subjects.[12]

Table 1. Pairs with different concord

Singular verb (unexpected)	***Plural verb (orthodox)***
… there seems to be fewer episodes of rebleeding … (HU2 6914)	… there seem to be few reliable cues … (HX9 1187)
… there seems to be many instances … (B7H 1376)	… there seem to be many cultures … (CLP 1144)
There seems to be two possible ways … (EW1 831)	… there seem to be two basic ways … (J2K 706)

Judging from the examples presented, there seems to be a fair degree of free variation. However, the corpus represents a spectrum of genres, registers and styles, and, given the well-known differences between planned (and perhaps even edited) and spontaneous "online" production, it can be assumed that the constructions are not evenly distributed. Therefore, in what follows, examples and statistics for each of the two catenative verbs will initially be presented in two main sections, spoken and written. The material does not seem large enough for further systematic subgrouping to be meaningful.

[12] The distinction sometimes made between "weak" and "strong" determiners (for a listing, see e.g. Meechan & Foley 1994:67) does not seem to have played a part in the choice of concord in the present material. Meechan & Foley (1994:79) found that, according to the variable rule program they used, the only truly relevant distinction was between *no* and all the other determiners.

3.2 SEEM

3.2.1 Table 2. Spoken British English, c. 10 million words. Plural NP with singular verb, 10 instances

Unexpected singular verb	**Close parallels or similar cases with orthodox plural verbs**
There seems to be an awful lot of people … (KCK 1280)	There seem to be an awful lot of local companies … (JA4 487)
There seems to be an awful lot of us going … (KD8 3732)	And there seem to be a lot of those. (KRL 5026)
… there seems to be about four or five about … (KBB 6088)	… there seem to be a lot of Scouts about … (HYJ 998)
… there seems to be more men than women … (KD6 459)	… there seem to have been a lot of problems … (KRW 59)
… there seems to be more flowers going down than up … (KCD 4789)	… there seem to [be] a number of headings here that I couldn't account for … (JA5 95)
… there seems to be many people for whom erm all my effort gets nowhere … (KRL 3122)	There seem to be so many different terms … (KRS 33)
There seems to be quite a number [of churches] … (GYT 81)	How many joint telephones can you have? There seem to be a number in Client Services. (FLS 456)
… there seems to be two stands this morning. (J9B 527)	There seem to be <pause> our two groups … (KB0 1962)
… there seems to be no questions … (J3W 128)	… there seem to be some suggestions … (JNM 25)
… there seems to be very few councillors attending. (J3R 459)	… there seem to be quite a few German [TV stations] … (KDM 5167)

3.2.2 Table 3. Written British English, c. 90 million words. Plural NP with singular verb, 19 instances

Unexpected singular verb	**Close parallels with orthodox plural verbs**
There thus seems to be quantitative differences … (FT0 188)	… there seem to be fundamental differences … (GV0 1541)
… there seems to be two types of studs … (AN2 469)	… there seem to be two anatomically distinct types … (HWS 4384)
… there seems to be two Jamaicas … (HH3 6879)	There seem to be two main differences here … (G2E 1948)
There seems to be three possibilities … (H78 1682)	… there seem to be at least three possible strategies … (GWJ 987)
There seems to be two possible ways … (EW1 831)	… there seem to be two basic ways … (J2K 706)
… there seems to be two centres or loci of power … (EEB 1822)	… there seem to be two bases on which a contractual term can be attacked … (J7B 132)
… there seems to be two concentration optima … (HWS 1948)	There seem to be two distinct periods … (AA8 167)
There seems to be hundreds of supporters' clubs. (K4T 10505)	… there seem to me to be four kinds of responses. (GVX 331)
… there seems to have been hundreds and thousands of years … (J18 1107)	… there seem to have been many years when there was little or no nationally-coordinated resistance. (HXX 318)
… there seems to be a few misconceptions … (C91 1706)	There seem to be quite a few local authorities … (HPP 555)
… there seems to be fewer episodes of rebleeding … (HU2 6914)	… there seem to be few cases of actual censorship. (CDU 1662)
… there seems to be several species … (H10 202)	… there seem to be several quite different types of trial … (EUW 774)
… there seems to be many instances … (B7H 1376)	… there seem to be many cultures … (CLP 1144)

… there seems to be far too many players …HJ3 7647)
There seems to have been too many people writing blockbusters. (GWK 989)
There seems to be rather a lot of the latter [people] about … (B1J 1942)
There seems to be rather a lot of names … (B1J 983)

… there seems to be a number of factors at play … (GWL 647)

... there always seems to be an inadequate number of facilities for women. (EFG 591)

There seem to be too many outlets in Wales … (CRB 2663)
There seem to be too many women with very possessive attitudes towards you! (JY3 3603)
… there always seem to be a lot of people who make the effort … (HDB 1024)
… there seem to have been a lot of gas shells landing … (HRA 3929)
… there seem to be a number of rules to be followed … (BML 1558)
... there seem to be a number of connections ... (HY0 164

3.2.3 General comments on the results

The overall absolute figures for spoken British English are 10 and 14 for singular and plural verb respectively, which looks like this in pie-chart format:[13]

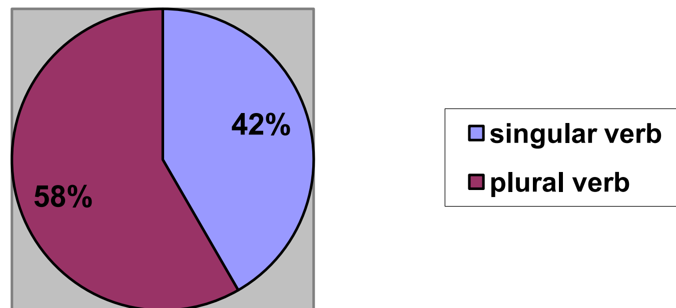


Figure 1. Spoken English. Choice between singular *seems* and plural *seem* with plural notional subjects. N = 24.

It is only to be expected that unorthodox concord in general, not only with existential *there*, has its highest frequency in informal spoken language. However, even within that category, syntactically conditioned differences can be noted. For existential sentences it has been pointed out by Breivik & Martinez-Insua (2008:358) that intervening material between the verb (in their investigation always a finite form of *be*) and the postposed notional subject may make speakers "lose sight" of the number of the notional subject.[14] With catenatives, intervening material is a built-in feature, in the form of *to be* or *to have been*, so it is

[13] It should be kept in mind that with this small total, each instance carries over four percentage points.

[14] The proportion of non-concord cases with and without "extensions" is roughly 2:1 in their material from BNC and COLT.

systematically more difficult for speakers to handle concord in such constructions, as borne out by the proportions in the present material.[15]

The pattern in written British English is illustrated in the following pie-chart, which has a stronger statistical basis than the previous one:

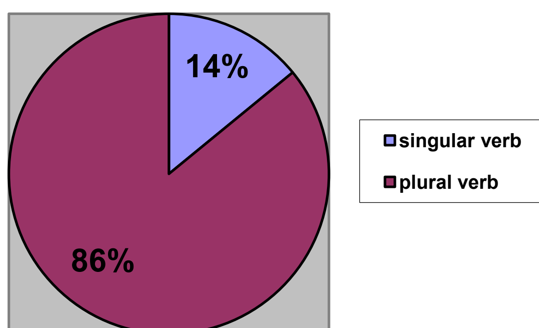


Figure 2. Written British English. Choice between singular *seems* and plural *seem* with plural notional subjects. N = 134.

Written, and perhaps edited, language offers opportunities for revision and amendment which are not available in online spoken production. This is indeed confirmed by the BNC material. Only one out of seven existential sentences with plural notional subjects comes out with unorthodox concord, as against two out of five in the spoken BNC material.[16] Still it is surprising that 13 of the 19 cases originate in the

[15] There are some more striking examples of intervening material, but none of them have contributed to unorthodox concord. Two of them are (bold added): *There seems* ***to me to have been*** *a growing wish to* ... (EA8 282) and *There seems,* ***from past votes on the matter, to be*** *an overwhelming majority* ... (C8R 321).

[16] As long as we regard the BNC as the universe that we are investigating, these figures can be taken at face value. However, owing to the small number of spoken examples (which is a reflection of the relatively small size of the spoken component of the BNC), the difference between spoken and written, though striking enough in pie-chart format, is statistically significant only at the 10 per

edited categories Books and Periodicals, which means that there is no real difference between strictly edited and more informal written material.[17]

3.2.4 Comparisons with CobuildDirect

As shown in Olofsson (2007:5), in CobuildDirect (c. 57 million words) there are 84 instances of plural notional subjects with existential *there* and some present-tense form of *seem*, which means that the overall frequency of this construction in that corpus is c. 1.5 per million words. The pmw figure for the BNC is the same. As for the proportions of singular and plural verbs, the best area for comparison seems to be spoken British English, which is a separate category in both corpora. Unorthodox concord is slightly more frequent in the BNC (42%) than in CobuildDirect (34%), but the totals are far from large enough for general conclusions. More interesting than such a comparison on a weak statistical basis is to conflate the figures for spoken British English from the two corpora, which gives a boosted total of 56 and the proportions illustrated in the following pie-chart:

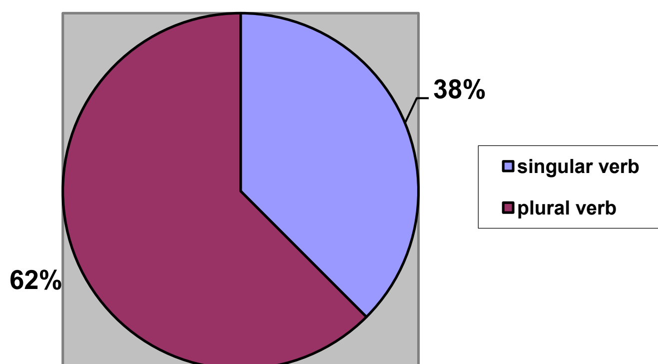


Figure 3. Choice between singular *seems* and plural *seem* with plural notional subjects in the spoken British English of the BNC and of CobuildDirect. N = 56.

---

cent level. Not even with the conflation with CobuildDirect done below (3.2.4 and Figure 3) does the difference quite reach the 5 per cent level.

[17] C. 70 out of the 90 million words in the written component represent Books and Periodicals, so 13 out of 19 is only marginally less than what can be expected.

It seems safe to conclude that in spoken British English with the degree of (in)formality represented in the corpora, roughly two catenative *there*-constructions out of five with a plural notional subject come out as *there seems*.

3.3 APPEAR

3.3.1 Spoken

In the spoken part of the BNC (c. 10 million words) there are two cases of *there + appear* where *appear* is finite and the notional subject is a plural NP. Of *there appears*, there are nine cases in all, none of which has a plural notional subject. There is thus nothing to comment on, except the low frequency of catenative *appear* in spoken English.

Catenative *appear* is less frequent in general than *seem* and it is particularly infrequent in spoken English as compared with *seem*. A search on the strings *appear to, appears to, appeared to, seem to, seems to* and *seemed to* (with no other collocational restrictions) shows that in the written part of the BNC, the *seem* forms taken collectively are 2.6 times more frequent, and in the spoken part they are over ten times more frequent than the catenative *appear* forms. The *seem* forms are fairly evenly distributed over written and spoken in relation to the amounts of text, whereas the *appear* forms are about four times more frequent in written than in spoken. This medium and register difference comes out very clearly in Biber et al. (1999:945), where a diagram shows that the typical habitat of *there appear(s) to* is academic prose, whereas *seem* is more evenly distributed.

3.3.2 Written

In the written part of the BNC (c. 90 million words) there are 77 cases of *there + appear* where the lexeme *appear* is in the present tense and the notional subject is a plural NP. There are no instances of *there appear* with a notional subject in the singular.[18] There are thirteen cases of *there appears* with a plural notional subject, i.e. 14% of the total situations

---

[18] This statement is not as vacuous as it may seem. Olofsson (2007:6) reports two instances from CobuildDirect.

with a plural notional subject.[19] The picture is identical to that presented in pie-chart format for *seem* (Figure 2 above) but with weaker statistical support.

Table 4. Written British English, c. 90 million words. Plural NP with singular verb, 13 instances

Unexpected singular verb	**Close parallels with orthodox plural verbs**
There appears to be only three front ranking firms not involved … (A55 277)	… there appear to be only three main sources of gesture … (A12 1089)
There appears to be disagreements within Whitehall … (A96 57)	There appear to be mountings for the transformer … (G00 3124)
There appears to be no standard pinouts for p.c.b. (C91 994)	… there appear to be no obvious reasons … (B15 241)
Through the mouth there appears to be the remains of a chain link … (G30 171)	There appear to be those in the Lord Chancellor's Department who perceive its role as being far wider … (FRT 566)
There appears to be just 60% of the residents supporting the scheme. (GXH 220 + HD1 250)	… there appear to be only five possibles … (HTG 3244)
… in some preserved specimens there appears to be interradial rays of thickened skin; (H79 2145)	There appear to be structural differences between the major countries … (H7T 1360)
There also appears to be some quasi-religious motives for the operation … (F9E 1376)	There appear to be some important challenges … (B2A 1465)
With a laminate there appears to be many who will [listen]. (HH0 1577)	There appear to be many small churches … (ADC 1167)
There appears to be substantial numbers of people who join or leave … (HJ2 3140)	Nevertheless, there still appear to be notable differences … (J7D 153)

[19] Strictly speaking, the figure is slightly misleading because one of the cases occurs twice (i.e. verbatim in two different texts). If one of the doubled instances is subtracted, the percentage goes down to 13.

There appears to have been a substantial number of “rank and file” supporters scattered across the whole of both Galilee and Judaea (EDY 955)	Of course, there also appear to have been a good number of people who regarded her as an eccentric … (AN4 339)
… that [misprint for *there*] appears to be quite a significant number of dogs … (K54 3984)	(No more examples with *number* and *appear*) … there seem to be a growing number of collectors. (KAP 67)
There appears to be only a finite number of such theories. (FYX 889)	There seem to be an extraordinary number of people ... (HD4 32)

Somewhat surprisingly in view of the fact that the characteristic environment for *there* + *appear* is written language, no less than seven of the instances of unorthodox concord with *appear* (C91 994, EDY 955, FYX 889, F9E 1376, G30 171, H79 2138 and HH0 1577) hail from the professionally edited Books and Periodicals.

4. Some grey areas

As is well known, it is not always easy to determine whether a noun phrase merits grammatical treatment as singular or plural.

4.1 Coordinated noun phrases

In cases where two noun phrases, with or without different grammatical numbers, are coordinated, there can be doubt as to concord, as discussed by e.g. Quirk et al. (1985:759). I have counted for instance the following sentence as representing orthodox concord.

(11) In the posterior part of the body there appears to be a distinct mid-line streak and diagonal marks running out from it. (B7N 781)

My argument is that the coordinator *and* is very close to the preposition *with* in a sentence of this kind. What is being described by *streak* and *marks* is one pattern rather than two.

4.2 Collective nouns

In British English, a collective noun such as *group* or *team* as subject can take either a singular or a plural verb.[20] The best-known collective nouns are not represented in the present material; the closest we get is an instance of *array* in *There appears to be a bewildering array of choices* (HSW 10), which I have chosen to exclude from my statistics. This noun is not among the examples of (frequent) collective nouns enumerated in Quirk et al. (1985:316), nor is it mentioned in connection with the grammaticalization of noun phrases such as *a number of* in Quirk et al. (1985:264).[21] However, its status as a variable collective noun is illustrated in the *Collins COBUILD Dictionary for Advanced Learners*: (under **array 1**) *A dazzling array of celebrities* ***are*** *expected at the Mayfair gallery to see the pictures* vs. (under **array 2)** *There* ***was*** *an impressive array of pill bottles stacked on top of the fridge* (emphasis mine).[22] Now, using paraphrases, if for the BNC example *A bewildering array of choices* ***was*** *presented* is the "correct" form, the example quoted has normal concord, but if *A bewildering array of choices* ***were*** *presented* is preferable (still in BrE), then the concord of the corpus example is deviant.[23]

4.3 Relative what

Without being an example of the situation in question, the sentence ... *there appears to be a preponderance of narrow gauge or what I would call miniature railways.* (CJ7 78) brings to our attention a potential grammatical ambiguity. Relative *what* is in itself singular, but, as pointed

---

[20] General problems of this kind are discussed extensively in Levin (2001). There is only a brief mention of the existential construction, but his two examples (2001:92) are in line with my discussion.

[21] As can be seen from the arrangement of the examples in the last three rows in Table 4 above, I have regarded *a number of* as a determiner, following Quirk et al. (1985:264), and the grammatical number of the noun that follows (by definition the plural) as decisive for the choice between singular and plural verb.

[22] Somewhat confusingly, the COBUILD "extra column" says "N-COUNT-COLL: usu sing" about the subsense that is illustrated with an example where the verb is plural.

[23] Olofsson (2007:7) uses the CobuildDirect example *There does seem to be a group of women who suffer because of their hormonal cycles ...* and offers two paraphrases to choose from to check the grammatical number: *This group of women suffers* vs. *This group of women suffer*.

out by e.g. Quirk et al. (1985:755), S–V concord is governed by the relevant noun rather than by the formally singular *what*. Although there is no case in the present material, it may be worth while noting the potential problem illustrated by a paraphrase of the above sentence: *There appear(s) to be what are usually called miniature railways*. It would be difficult to determine whether *appear* (based on *railways*) or *appears* (based on *what*) would represent orthodox concord here.

5. Implications for teaching and testing

Taking a Nordic perspective, it can be noted that for most Norwegian and Swedish learners of English, there are two main difficulties involved in the construction under discussion.[24] Firstly, the same introductory subject (*det*) is used for both the existential construction (*There seem/s/ to be ...)* and extraposition (*It seems that there is/are ...*). Thus we find a tendency to produce strings like (**It seems to be something wrong*), certainly helped on by the general frequency of the string *it seems* in English (75 pmw in the written and 91 pmw in the spoken component of the BNC).[25] Secondly, subject–verb concord in Danish, Norwegian and Swedish was gone before the present generation of learners were born, so with any language that has that kind of concord, learners are faced with the need for a new way of thinking.[26] Diagnostic tests and other investigations have shown that texts produced by at least first-term university students of English in Sweden show a considerable amount of uninformed guessing in the choice between singular and plural forms of

---

[24] Danish learners share only one of the difficulties, that of subject–verb concord. For the information on Danish and Norwegian, I have drawn on Breivik (1983:359–370).

[25] There is actually a dialectal direct counterpart in the south of Sweden which uses existential *där* (which is etymologically identical with *there*) much in the same way as Danish (*der*). The same holds for Norwegian dialects along the southern and western coast, which can use *der* (Breivik 1983:364).

[26] It should be noted that when Swedish still had subject–verb concord, it was the notional subject in the dialectal construction mentioned in the previous note that governed the form (singular/plural) of the verb.

predicate verbs, even in straightforward cases without complicating factors like existential *there*.[27]

In the construction with *there* + some form of *appear* or *seem*, the prescriptive rule is quite clear. The notional subject is what is supposed to govern the choice of verb form. However, as demonstrated above, actual usage differs considerably from the rule, particularly with *seem* in spoken English. In a situation that calls for *there* + a present-tense form of *seem* and a plural notional subject in spoken language, two cases out of five produced by native speakers come out containing the "incorrect" alternative. It would seem cruel to penalize learners severely if their pattern emulates that of native speakers, but at the same time it should be pointed out to them that adherence to the general rule for subject–verb concord is what will guarantee universal acceptance of this aspect of their text.

References

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Longman.

Breivik, Leiv Egil. 1983. *Existential* there. *A Synchronic and Diachronic Study*. Studia Anglistica Norvegica 2. Bergen: University of Bergen.

Breivik, Leiv Egil & Ana E. Martinez-Insua. 2008. "Grammaticalization, subjectification and non-concord in English existential sentences." *English Studies* 89/3. 351-362.

Collins COBUILD Dictionary for Advanced Learners. 2001 (3rd edition). Glasgow: HarperCollins Publishers.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: CUP.

[27] Large-scale diagnostic testing done at the University of Gothenburg in the late 1980s and onwards has shown that not even straightforward examples like *John has* are mastered by all prospective first-termers, in spite of the fact that their scholastic background contains some ten years of English. As for the construction discussed here, less than 10% manage to get both *there* and (orthodox) concord right in a translation that should result in *There seem to be very few people here today*. See in particular Olofsson (1989) and Köhlmyr (2005), but also Thagg Fisher (1985) and Karlsson (2002:39–80).

Karlsson, Monica. 2002. *Progression and Regression*. Lund: Lund Studies in English 104.

Köhlmyr, Pia. 2005. “Raw recruits revisited: 1. Introduction. A follow-up study on first-term students’ English proficiency.” *LMS-Lingua* 2005/4. 36-39.

Levin, Magnus. 2001. *Agreement with Collective Nouns in English.* Lund: Lund Studies in English 103.

Martinez Insua, Ana E. & Ignacio M. Palacios Martinez. 2003. “A corpus-based approach to non-concord in Present-day English existential *there*-constructions.” *English Studies* 84/3. 262-283.

Meechan, Marjory & Michele Foley. 1994. “On resolving disagreement: linguistic theory and variation—*There’s bridges*.” *Language variation and change* 6. 63-85.

Olofsson, Arne. 1989. “Raw recruits. A summary of diagnostic test results 1987–89” (mimeo available from the author).

Olofsson, Arne. 2007. “Just how wrong is *There seems to be* + a plural noun phrase?” *Moderna Språk* 101:1. 2-8.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Thagg Fisher, Ulla. 1985. *The Sweet Sound of Concord*. Lund: Lund Studies in English 73.

Recurrent word-combinations in English student essays[1]

Signe Oksefjell Ebeling, University of Oslo

Abstract

The point of departure for this study is Stubbs and Barth's article from 2003 where they describe how recurrent phrases can be used as text-type discriminators. A subset of the *British Academic Written English* (BAWE) corpus—a corpus of student writing—consisting of English Studies essays is compared with two text-types taken from BNC-Baby in order to test and validate Stubbs and Barth's claims about recurrent word-combinations. Then the most salient combinations occurring in the BAWE student essays are subjected to a functional analysis, based on Moon (1998), to better be able to say something about their main functions in this text-type. Finally, the functions of recurrent phrases of the BAWE essays will be compared to the functional characteristics of the academic prose part of BNC-Baby.

Keywords: recurrent word-combinations, ngrams, lexical bundles, functional analysis, text-type discriminator, corpus-based analysis

> In a corpus of texts from different speakers and writers, they [recurrent word-chains] can be studied (a) as a predictable characteristic of different text-types, and (b) as evidence of units of routine language (Stubbs and Barth 2003:62)

1. Introduction and aims

Recurrent word-combinations have in recent years been used to characterise genre and text-type; they have even been described as "text-type discriminators" (Stubbs and Barth 2003). This relationship between recurrence and text-type will be explored in the present article. I will follow Altenberg (1998:101) in defining recurrent word-combinations as "any continuous string of words occurring more than once in identical form".

This paper sets out to give an overview of which word-combinations English Studies students tend to resort to in their essays, what these combinations typically express in functional terms, and how they compare with word-combinations typical of other text-types.

[1] I am grateful to Hilde Hasselgård and two anonymous reviewers for comments on an earlier version of this paper.

The first part of the study focuses on recurrent patterns found in essays written by UK undergraduate students in English Studies, with the ultimate aim of establishing their functional characteristics. In order to establish which patterns typically recur in English essays, quantitative corpus methods will be applied. The recurrent patterns identified for English essays will be compared with recurrent word-combinations in other text-types. For this purpose, the text categories "academic prose" and "written fiction" taken from BNC-Baby were chosen. By comparing three potentially very different text categories, Stubbs and Barth's (2003: 62) claim that "text-types are repetitive in different ways and to different extents" will be examined and validated.

The second part, and the bulk of this study, takes the most salient recurrent word combinations in two of the text types, viz. academic prose and the English essay, as its starting point, with the aim of establishing and comparing the functional qualities of these text types. This will ensure a more balanced approach towards the data, complementing quantitative findings with a qualitative, functional analysis. The functional analysis follows Moon's (1998) model of classifying expressions into organizational, informational, situational, evaluative and modalizing (cf. section 6.1).

Where other studies have focused on how novice writers compare with professional writers of the same domain (e.g. Shaw 2009 on student vs. professional writing in literary studies) or how they try to come to terms with the epistemology of their discipline (e.g. Charles 2006, Ebeling & Wickens Forthcoming), this study is more concerned with the functional roles the recurrent word-combinations play in the student essays and how these compare with academic prose in general. These functional roles may in turn point to the epistemological properties within different discourse communities, as "phraseology and epistemology are indissolubly interlinked" (Groom 2009: 125).

2. Material and method

The main data for this study are taken from the *British Academic Written English* (BAWE) corpus.[2] It is a corpus of proficient student writing

---

[2] The British Academic Written English (BAWE) corpus was developed at the Universities of Warwick, Reading and Oxford Brookes under the directorship of

covering a wide variety of student assignment types across four disciplinary groupings, viz. Arts and Humanities, Life Sciences, Physical Sciences and Social Sciences. (See Nesi et al. (2005) and Heuboeck et al. (2007) for a more detailed account of the corpus.)

One of the largest homogeneous groups of assignments included in the corpus is the English essay, and the sample examined in this study comprises 89 essays from English studies, amounting to approximately 205,000 running words.[3]

Although homogenous in the sense that they are all classified as essays,[4] they differ in other respects; they are written by different students of different genders, different ages and different language backgrounds and at different levels of study. Nevertheless, the essays will be studied here as a single unit making up the data termed BAWE essays (i.e. English Studies essays in UK higher education), mostly concerned with English literature.

Much the same variation with respect to authors, authors' gender and written style also applies to the data taken from BNC-Baby, which will be used for comparison with the BAWE material. BNC-Baby is a subset of the *British National Corpus*, containing "four one-million word samples, each compiled as an example of a particular genre: fiction, newspapers, academic writing and spoken conversation."[5] Only two of these—academic prose and fiction—will be used for comparison with the BAWE data.

Hilary Nesi and Sheena Gardner (formerly of the Centre for Applied Linguistics [previously called CELTE], Warwick), Paul Thompson (Department of Applied Linguistics, Reading) and Paul Wickens (Westminster Institute of Education, Oxford Brookes), with funding from the ESRC, http://www.coventry.ac.uk/researchnet/d/505, accessed 28 Oct. 2009.

[3] Footnotes and endnotes were ignored for the purpose of this study. The main reason for this was that most notes in the English essays were references and not running text.

[4] "Essay" is one of thirteen genre families that have been identified in the BAWE material according to the learning function they serve (see Nesi & Gardner Unpublished MS). Incidentally, the essays studied here were also labelled essays by the students themselves.

[5] http://www.natcorp.ox.ac.uk/corpus/index.xml.ID=products#baby, accessed 28 Oct. 2009.

Of the two text-types chosen from BNC-Baby, fiction is perhaps the more clearly defined. Academic prose, on the other hand, is a more loosely defined category and, in addition, it represents a number of different scientific disciplines. Although both academic prose and fiction are potentially very different from the English essay in terms of recurrent word-combinations used, the academic setting for the BAWE material will probably have an impact on the results. It can be expected that there will be more overlap between the BAWE essays and academic prose than between the essays and fiction.

The method applied in this study follows Stubbs and Barth (2003), where the most frequent recurrent word-combinations are compared across the selected text-types, in order to investigate the degree to which the word combinations can be said to work as text-type discriminators.

The two BNC-Baby sub-corpora contain 5 times as many words as the sub-corpus of BAWE essays (ca. 1 mill. vs. ca. 200,000 words) and it may be argued that such a discrepancy in size may distort the findings as a more varied set of recurrent combinations can be expected in a larger material. This made me want to find out if a direct comparison of frequent word combinations in the different-sized corpora would distort my findings significantly, or whether I could use the 1 mill. word corpora for comparison. I therefore took a random sample of 200,000 words from academic prose and fiction, and compared the top 30 three-word combinations in that sample with the top 30 three-word combinations in the 1 mill. word original corpora. In both fiction and academic prose there was only 63% overlap between the shorter sample and the full corpus. As the reason for this might have been my selection process of picking 6 BNC-Baby files at random, I tried a different sampling method of about 6,500 words from each of the 30 academic prose files and about 8,000 words from each of the 25 fiction files in order to reach 200,000 words. This proved to be more successful in terms of stability in the recurrent combinations, with 80% overlap for academic prose and 73% for fiction. The latter 200,000 word samples quite clearly are more representative of the text-types as a whole than the 6-file sample. The fact that the difference between the 200,000 word samples (when using the second sampling method) in both fiction and academic prose had a fairly high rate of overlap with the full million word sub-corpora, I decided, for convenience, to use the full-size ones for further comparison with the BAWE data.

3. Background

For the purpose of this study, the terms *text-type* and *genre* have been conflated. Both are hard to pin down, and as pointed out by Swales "[t]he word [genre] is highly attractive—even to the Parisian timbre of its normal pronunciation—but extremely slippery" (Swales 1990: 33). The BAWE corpus manual identifies a range of genres, which in turn have been grouped into 13 genre families, or classes of genres, sharing functional and structural properties (cf. Heuboeck et al. 2007: 7 & 45ff). Following the BAWE classification scheme,[6] influenced by Swales (1990) and Martin (1992, 1997), genre is defined as "a staged, goal-oriented social process realised through register" (Martin 1992: 505), and members of a genre family "may share a central function, or they may have evolved in the same disciplinary context" (Nesi & Gardner Unpublished MS). In the following, I will use the term *text-type*, since my point of departure is Stubbs and Barth's investigation of *text-type* discriminators.

The data for this investigation are from three text-types identified by the BNC-Baby team as academic prose and fiction and by the BAWE team as essay, one of the "making sense" genre families (cf. Nesi & Gardner Unpublished MS). The essay in the BAWE corpus is characterised by writing where the evidence to support an argument has to be sought widely and is open to debate; essays "expect students to express a viewpoint, e.g. 'Is it worthwhile to test intelligence?', which suggests students should be gathering evidence and forming their own thesis in response" (Nesi & Gardner Unpublished MS).

Previous research has shown that the use of ngrams, or clusters, as text-type discriminators is rewarding "because they give insights into important aspects of the phraseology used by writers in different contexts" (Scott & Tribble 2006: 132). This is not to say that clusters alone can identify a text-type, but they can point to typical formulations used in particular text-types, and thus be used for instructive purposes for novice writers. Also, for the purpose of the present study, they can be used to point to differences between student writers in one discipline and how they compare to other writers with "the potential to enhance our appreciation (and that of learners) of what works in particular kinds of text" (ibid.).

[6] See Nesi & Gardner (Unpublished MS).

This technique has not been applied to the BAWE material before, and it will be interesting to see what ngram statistics will tell us about the English essay as a text-type compared to the other text-types investigated here.

4. Ngrams

Although this research focuses mainly on clusters of words, a brief look at the 10 most frequent word forms in the BAWE essays show that English essays do not deviate significantly from English in general as regards the most frequent single-word frequencies (cf. Sinclair's overview of frequent word forms in the *Bank of English* (Sinclair 1999). This underlines Sinclair's point about frequent words forming a large proportion of any text (ibid.:157). Typically these frequent words are function words, which do not easily distinguish text-types, at least not on the same level as content words may do. However, in combination with other function words or content words, function words are important building blocks in the phraseology of a language. Phraseology, as pointed out by Altenberg (1998:101), involves "various kinds of composite units and 'pre-patterned' expressions such as idioms, fixed phrases, and collocations". After examining the phraseology of spoken English on the basis of recurrent word-combinations, Altenberg admits that, although not all such combinations are of phraseological interest, they are "a useful starting point for an examination of the phraseology of spoken English" (ibid.:102). In the current setting recurrent word-combinations will provide a good starting point for investigating in what way word patterns in English essays reflect the text-type in which they are used.

In this section we will identify ngrams, i.e. sequences of types, in the BAWE essays; we will only be concerned with those that occur frequently.[7] Such recurrent word-combinations differ from idioms and collocations, since they are not necessarily complete syntactic constituents or phrases. They are similar to what Biber and Conrad have termed *lexical bundles*, viz. "the most frequent recurring lexical sequences; however, they are usually not complete structural units, and

[7] To extract the ngrams, or clusters, for this study WordSmith Tools (WS 5.0) were used (Scott 2008).

usually not fixed expressions" (1999:183).[8] They view lexical bundles as extended collocations typically occurring in sequences of three or more words. In this paper we will be mainly concerned with three- and four-word combinations. The formal and functional distinction that may exist between idioms, collocations and recurrent word-combinations will not be taken into account here, since all word-combinations will be identified on the basis of frequency alone. Thus, some of the recurrent word-combinations associated with English essays could also be termed collocations. It is less likely that there will be idioms among the recurrent word-combinations; these are comparatively rare in actual writing (cf. Sinclair 1999, Biber et al. 1999).

Earlier studies have shown that bigrams have little impact as text-type discriminators, partly due to the fact that the most frequent ones rarely form full structural units, e.g. *of the*, *is the*, *to a* (cf. Stubbs & Barth 2003, Altenberg 1998, Biber and Conrad 1999). This is also the case in the BAWE essays. The complete phrases that were identified among the top 50 bigrams typically reflect the content of the essays, e.g. *the reader*, *the poem*, *the novel, the text*. The relative overuse of content bigrams such as *the poem* in BAWE vs. academic prose (1,356 occurrences per million words vs. 26 occurrences per million words) can be easily explained by the fact that the BAWE essays are more specialized (in terms of both topic and text-type) than the academic prose corpus from BNC-Baby. A similar observation is pointed out by Pecorari (2008: 16), in her article on repeated language in academic discourse, where the corpus she used was much more specialized than the one she compared it with, viz. the academic corpus in Biber et al. (1999). In other words, it is important to keep in mind the narrowness of the BAWE material, i.e. the essays are all written within the framework of a limited number of modules by fewer authors than is the case in BNC-Baby. The four two-word noun combinations that were found among the 50 most frequent two-word combinations in the academic prose material from BNC-Baby are of a much more general kind than the ones recorded for the BAWE essays: *the same*, *per cent*, *for example* and *the other*. These

---

[8] In WordSmith Tools these are called *clusters*, i.e. "a *group of words which follow each other in a text*. The term *phrase* is not used here because it has technical senses in linguistics which would imply a grammatical relation between the words in it" (http://www.lexically.net/downloads/version5/HTML/?proc_definitions.htm, accessed 28 Oct. 2009).

noun combinations do not have the same ability to identify subject matter as the ones mentioned for the BAWE essays. The fact that *per cent* and *for example* are high up on the list of bigrams in academic prose may indeed point to a characteristic trait of academic writing as regards lexical choice. Still, there seems to be too few defining features, including content words, present in the most frequent bigrams for either text sample. Thus, bigrams will not be discussed any further in the present paper. Let us instead turn to trigrams and see whether our findings match the findings of Stubbs and Barth (2003:71): "With three-word chains, constituent content words appear earlier in the lists and are more frequent".

4.1 Trigrams

Using a similar approach to that of Stubbs and Barth (2003), I will compare the most frequent trigrams across three sub-corpora, listing the ones that are found in all three, followed by those that occur in two only. Finally, I will proceed to the trigrams that are most frequently found in only one of the sub-corpora.

Of the 50 most frequent trigrams, there are only four that occur in all three sub-corpora: *the end of, to be a, one of the,* and *end of the* (see Appendix I for an overview of the top 50 trigrams in the three text-types). As a three-word chain containing a content word, *the end of* also occurred among Stubbs and Barth's frequent trigrams featured in their FICTION and LEARNED categories. Their source material was taken from the Brown family of corpora, viz. Brown, LOB, Frown, and FLOB.[9] In addition, *the end of* was also recorded among the top 30 in their third genre BELLES (including belles lettres, biography, memoirs).[10] This trigram, then, seems to be a fairly general one in English overall.

In our context of looking at recurrent word-combinations that characterise specific text-types, it will therefore be potentially more rewarding to look at trigrams occurring in two, or only one, of our sub-corpora. First, let us move to trigrams that occur in two of our categories.

---

[9] Cf. ICAME corpus manuals, available at http://khnt.aksis.uib.no/icame/manuals/index.htm.

[10] Cf. the Brown Corpus Manual, available at http://khnt.aksis.uib.no/icame/manuals/brown/INDEX.HTM.

Table 1 Trigrams occurring among the 50 most frequent trigrams of two sub-corpora

Trigrams occurring in academic prose and BAWE essays	**Trigrams occurring in fiction and BAWE essays**	**Trigrams occurring in academic prose and fiction**
as well as	at the end	part of the
in order to		
in the first		
in which the		
it is a		
it is not		
it is possible		
on the other		
that it is		
the fact that		
the other hand		
the use of		
there is a		
there is no		

Table 1 shows that there is only one trigram that occurs in both academic prose and fiction and one that occurs in both fiction and the BAWE essays, while there are 14 that occur in both academic prose and English essays. This supports the view that English essays are closer to academic prose than to fiction. Indeed, if we look at the trigrams that are found both in academic prose and essays, quite a few of these are structuring devices that reflect a particular style, e.g. *the fact that*, *in order to, it is possible*. If they occur with a certain frequency in both academic prose and essays, we must assume that these are important building blocks of the text-types in question and in that sense point to a similarity between academic prose and essays.

Another observation that can be made on the basis of the trigrams is that the fiction ngrams often include a past tense form of a verb while both academic prose and the BAWE essays include a present tense form (cf. column 1 in Table 1). This tendency is pointed out by Biber et al. (1999: 456), where the corpus findings show that academic prose shows "a strong preference for present tense forms. Fiction shows the opposite pattern, with a strong preference for past tense verbs"; among the top 50

trigrams for BNC-Baby fiction, we find: *there was a*, *it was a*, *it was the*, *that it was*, etc. (cf. Appendix I).

The trigrams that are found among the 30 most frequent ones in one of the sub-corpora only are listed in Table 2.

Table 2 Trigrams unique among the 30 most frequent ones in one sub-corpus[11]

BAWE essays	**Academic prose**	**Fiction**
a sense of	a number of	a lot of
can be seen	and so on	back to the
due to the	in relation to	for a moment
heart of darkness	in terms of	going to be
in the novel	in this case	had been a
in the poem	it has been	he had been
it is the	it may be	he was a
of the novel	likely to be	in front of
of the poem	per cent of	it had been
the good soldier	some of the	it was a
the idea of	terms of the	it was the
the importance of	that there is	it would be
the reader is	the effect of	on to the
the reader to	the number of	out of the
the way in	the basis of	she had been
to the reader		shook his head
way in which		side of the
		that he had
		that he was
		the back of
		the rest of
		there was a
		there was no
		was going to
		would have been

Table 2 reveals quite a few things about the three sub-corpora under discussion. In fiction, as we would expect, personal pronouns are commonly found as part of trigrams; both *she* and *he* are used as part of the 30 most frequent trigrams. Another aspect worth noting is that

[11] Capital letters have not been included in the tables throughout, although some ngrams are capitalised in the original texts, e.g. *Heart of Darkness*. However, if part of the current discussion, these ngrams will be capitalised.

trigrams containing verbs are much more frequent in fiction than in the other two categories, and as mentioned above, the verbs are commonly found in their past tense form. Fiction also has the largest number of trigrams not found in the other two categories, 27 vs. 14 for academic prose and 17 for essays. This suggests that of the three categories under discussion, fiction is the one that least resembles either of the other two.

The trigrams found among the top 30 in English essays undoubtedly reveal more of the actual content of the essays than is the case of either fiction or academic prose. Although the academic prose trigrams also contain content words such as *terms* and *basis*, the content words in the essays have more specific content, e.g. *poem*, *reader*, *text*. In fact, from the list in Table 2 it can be deduced that the essays are concerned with literary analysis, drawing on a narrower pool of lexis than in a less specialized corpus (cf. Pecorari 2008). In such cases it may therefore be more rewarding to look at patterns as text-type discriminators rather than exact lexical matches (e.g. *of* + det. + noun).

Before we move on to examine the function of the ngrams in our sample, let us examine quadrigrams in the material.

4.2 Quadrigrams

Four-word combinations, or quadrigrams, show similar patterns to those of trigrams across the three text categories. Of the top 50 quadrigrams, only four occur in all three categories: *the end of the*, *the rest of the*, *at the end of*, *at the same time* (see Appendix II for a complete list). It should be noted here that some of the quadrigrams include some of the trigrams, e.g. *at the end of* ◄ *at the end*, or as Biber et al. (1999:990) put it "[s]horter bundles are often incorporated into more than one longer lexical bundle".

Again, a combination featuring the noun *end* is central and seems to be an item that commonly occurs across text categories.[12] Also worth noticing is the sequence "(prep. +) definite article + noun + preposition (+ def. art.)" as a common prefabricated and recurrent pattern, which has

[12] Incidentally, this has also been noted across a variety of other apprentice writer corpora (including foreign learners), cf. Brook O'Donnell and Römer's presentation at the 2009 ICAME conference on "Proficiency development and the phraseology of learner language".

previously been identified one of "the top-5-POS-grams in the whole BNC (down to a cut-off of 3 for n-grams)" (Stubbs 2004).

As with the trigrams, recurrent quadrigrams among the top 50 show most overlap between academic prose and the BAWE essays, with 11 shared quadrigrams, in addition to the four mentioned above. Three of these are of the type DET + N + PREP + REL. PRON. (e.g. *the way in which*). The remaining eight are: *on the other hand*, *it is possible to*, *the fact that the*, *that there is no*, *in the form of*, *as well as the*, *it is clear that*, *that there is a*. These reflect the same tendency as the trigrams; most are formulaic devices representing a particular writing style. This shared feature may point to the fact that the literary criticism performed in the BAWE essays is a subset of the wider academic prose genre, triggered by the academic setting in which the essays have been produced.

As regards quadrigrams among the top 30 that are unique to one sub-corpus, Table 3 shows that most are found in fiction (26), with academic prose (20) and the BAWE essays (18) not far behind.

Table 3 Quadrigrams unique among the 30 most frequent ones in one sub-corpus

BAWE essays	**Academic prose**	**Fiction**
allows the reader to	a wide range of	a cup of tea
as can be seen	as a result of	as if he was
at the beginning of	as shown in fig	at the back of
by the use of	at the time of	at the top of
can be seen in	can be used to	did not want to
could be argued that	in terms of the	for the first time
in the winter's tale	in the case of	he was going to
in the good soldier	in the context of	he shook his head
in heart of darkness	in the united states	i don't want to
it could be argued	in rylands v fletcher	in the middle of
of the poem the	in the form of	in front of him
the beginning of the	in relation to the	in front of the
the image of the	in the absence of	in one of the
the importance of the	it is important to	in the first place
the death of the	on the basis of	it would have been
through the use of	one of the most	on the edge of
to the fact that	per cent of the	on the other side

way in which the	the house of lords	the edge of the
	the house of commons	the top of the
	the size of the	the back of the
		the side of the
		the other side of
		the middle of the
		the centre of the
		was going to be
		what do you mean

The most specific combinations with content nouns are found in the essays, but notably there are also three very specific noun combinations in academic prose: *the House of Lords*, *in the United States*, *in Ryland v Fletcher*, and *the House of Commons*, all being more topic-specific than the other combinations (including trigrams).

Both academic prose and fiction show clear tendencies with regard to patterns that may help define them as text-types. The most common pattern among the 20 quadrigrams found in academic prose is PREP + (DET) + N + PREP + (DET), e.g. *on the basis of*. Although the fiction texts commonly make use of a similar pattern, there is a difference in the choice of noun, e.g. *the edge of the*. A similar tendency was found by Stubbs and Barth in their learned vs. fiction categories (Stubbs & Barth 2003: 72). Furthermore, the fiction patterns typically have an element (N) indicating position: *top*, *middle*, *back*, etc. (cf. Stubbs and Barth 2003:72). Another characteristic feature of the fiction quadrigrams are the combinations with personal pronouns + VP, e. g. *I don't want to*. Based on the trigrams and quadrigrams examined from BNC-Baby fiction, it can be concluded, with Stubbs and Barth, that it has a verbal style with vocabulary from the lexical fields of knowing and wanting.

On the basis of recurrent word-combinations, academic prose can be summarised as being nominal in nature; 17 out of the 20 quadrigrams listed in Table 3 contain a noun, most commonly an abstract noun. Our findings correspond well with what Biber et al. (1999:992) recorded for lexical bundles in academic prose: "most lexical bundles in academic prose are building blocks for extended noun phrases or prepositional phrases". Furthermore, only two lexical verbs have been recorded among the quadrigrams unique to academic prose, viz. *show* and *use*. Whether it is on the basis of facts like these that Stubbs and Barth state that their

learned category is characterised by "a lack of stylistic variation" (2003:79) is hard to determine, but it is certainly one factor pointing in that direction.

The BAWE essays show less uniform patterns than either of the other two categories. They display some similarities with academic prose in that nouns are predominant; however, the noun patterns are more similar in their form and function in the academic prose texts, with more instances of preposition preceding the noun/noun phrase as instances of formulaic language. (This will be discussed in more detail in the following sections.) Furthermore, the essays show very few similarities with fiction, and I would hesitate in characterising essay as somewhere between academic prose and fiction in terms of text type. The picture seems to be more complicated than that. Nevertheless, ngrams have been shown to differ in the text-types discussed above, which suggests that they may be used for text-type discrimination, at least to some extent, and they will serve as a good starting point for the following functional analysis of the ngrams.

5. Recurrent word-combinations characterising English essays vs. academic prose

In the previous sections some similarities have been found between the BAWE essays and academic prose, particularly with regard to nominal patterns. However, differences, also in these patterns, have emerged, thus a functional analysis of the most frequent n-grams in these two text-types is called for, in order to offer a more in-depth comparison.

After introducing the model on which the functional analysis is based, I will first look at the BAWE essays. The most striking characteristic of the BAWE essays involves the relatively high number of specific proper nouns. In order to determine whether there is more uniformity in the most frequent ngrams, this section is devoted to a functional analysis of the ngrams found in the essays. I will concentrate on the recurrent word-combinations that are most commonly found in the BAWE essays. More specifically, I will take a closer look at the trigrams and quadrigrams that were found among the top 30 only in the BAWE material (cf. column 1 in Tables 2 and 3). Bigrams will not be included here, as they were seen to contain too few text defining features (cf. Section 4). A more detailed analysis, featuring a functional analysis, of

trigrams and quadrigrams will hopefully tell us more about the text-type "English essay" than we have managed to reveal so far.

A similar analysis will then be carried out on the top 30 tri- and quadrigrams in academic prose before the functional characteristics of the two text-types are compared.

5.1 Classification and interpretation

In the analysis, I will adopt Moon's classification model which is based on Halliday's model of the three metafunctions: ideational, interpersonal and textual. Moon uses the model to classify fixed expressions and idioms (FEIs), and she states that "[t]he text functions of FEIs may be classified according to the way in which they contribute to the content and structure of a text. The precise contribution is instantial and bound up with context, but it is nevertheless possible to generalize and to chart typical functions" (Moon 1998:217). Figure 1 relates Halliday's ideational and interpersonal functions to Moon's FEI functions.

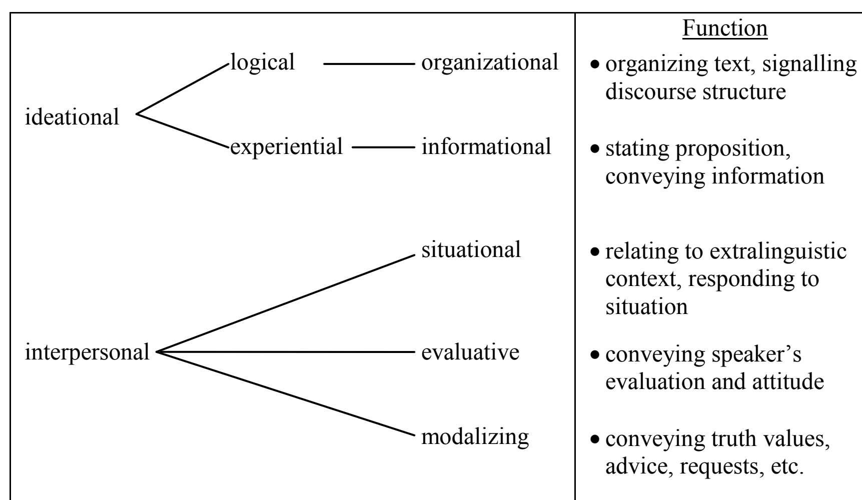


Figure 1: Breakdown of Moon's model (based on Moon 1998:217-218)

Halliday's textual function does not form part of Moon's model; "[t]he textual component [...] is best considered instantially in terms of the ways in which FEIs are placed topically and thematically, or contribute

cohesion to their texts" (ibid:218-219). It should be mentioned that Moon's organizational category has been criticized for being placed within the ideational function. As Culpeper and Kytö point out "[s]ome of the organizational expressions Moon deal with seem to overlap with the textual function […]. Later in her book, Moon refers back to her organizational categories and states that they provide 'grammatical cohesion'" (Culpeper & Kytö 2002: 49). Nevertheless, I will use Moon's categories here, well aware that her organizational category is not exclusively ideational, but also textual in Hallidayan terms.

The main difference between Moon's FEIs and our ngrams is that the ngrams do not necessarily make up linguistic units. FEIs are all fixed expressions or idioms that typically form meaningful phrases or clauses. Nevertheless, we will try to apply the model to our ngrams, only excluding ngrams containing functional words only, e.g. *it is the* (where *is* is condidered to have too little content to be labelled "lexical"). Table 4 gives an overview of the 35 most common trigrams and quadrigrams (unique among the top 30) in the BAWE essays, according to functional class. Again it should be pointed out that trigrams may be subsumed under quadrigrams (see Section 5.3). However, this is not believed to distort the overall functional classification of these ngrams; we simply get some overlap.

Table 4 Functional classification of ngrams in English essays

Organizational	**Informational**	**Evaluative**	**Modalizing**
at the beginning of	heart of darkness	a sense of	as can be seen
beginning of the	in heart of darkness	allows the reader to	can be seen in
by the use of	in the good soldier	due to the	can be seen
in the first the	in the novel	the importance of the	could be argued that
through the use of	in the poem	the importance of	it could be argued
	in the winter's tale	the way in	
	of the novel	to the fact that	
	of the poem the	way in which the	
	of the poem	way in which	
	the death of the		
	the good soldier		
	the idea of		
	the image of the		
	the reader is		
	the reader to		
	to the reader		

In my analysis, most of the ngrams were seen to fall inside the text function "informational", mainly due to the fairly high number of nouns, both proper and common. In some cases it may be difficult to classify the ngrams according to the functional framework specified above, without looking at more context. Thus, an example of each of the categories might be in order.

Organizational
(1) *In the first* stanza he uses hellish imagery: …

Informational
(2) *The reader is* plunged into the mind of the character …

Evaluative
(3) 'the Winter's Tale' is an unusual play *due to the* fact that it cannot be easily categorised into any particular form of play.

Modalizing
(4) *It could be argued* that the society created, and therefore Morvern, is simply amoral, …
Examples 1-4 are all fairly clear examples of their category. However, what about the ngram *in the Winter's Tale*, which has been classified as informational rather than organizational. Typically, this particular ngram has either been used as part of a noun phrase as in (5) or as an Adverbial providing information (6).

(5) The characters *in The Winter's Tale* have a past, …
(6) Shakespeare deliberately ridicules the 'unity of time' *in The Winter's Tale* …

Only in one of the 21 occurrences of *in the Winter's Tale* has it been used in sentence initial position where it could be considered as an organizational device (7).

(7) *In The Winter's Tale* (Shakespeare, 1610), a narrator personifying time is quickly and visually used to describe a shift in time of sixteen years.

The different interpretations of *in the Winter's Tale* are based on the functional role the ngram plays in the context, and it may not always be straightforward.

Returning to Table 4, we see that, proportionally, the ngrams have been grouped as follows: organizational 13.9%, informational 44.4%, situational 0%, evaluative 27.8%, and modalizing 13.9%. This means that more than 58% of the ngrams are, in Hallidayan terms, of an ideational nature, whereas the remaining 42% or so are of an interpersonal nature. On the level of discourse, then, it could be stated that English essays are highly informational, relatively evaluative, and to some extent organizational and modalizing. Not unexpectedly no clear situational word-combinations emerged; they are "typically found in spoken discourse as they are responses to or occasioned by the extralinguistic context: they may also be illocutionary speech acts. They are therefore constrained by real-world sociocultural factors" (Moon 1998: 225). Examples include: "I beg your pardon", "go for it", "it's a small world" (ibid.).

In a study of phraseology and epistemology in academic book reviews within two humanities disciplines, Groom confirms the claim that the epistemology of the knowledge domain of the humanities is essentially reiterative in nature "(i.e. concerned with revisiting perennial questions and reinterpreting previously existing data)" (Groom 2009: 124). His study focused on recurrent combinations around key prepositions (e.g. conceptualization + *of* + phenomenon). Although the approach in the present study differs from that of Groom, I believe the functional analysis of ngrams in English studies essays (as a discipline within the humanities) equally lends support to the humanities as being reiterative, particularly shown through the high proportion of informational and evaluative ngrams.

In order to investigate how the BAWE ngrams compare in functional terms with those of academic prose, a similar analysis of the trigrams and quadrigrams in academic prose was also performed, revealing the functional distribution given in Table 5.

Table 5 Functional classification of ngrams in academic prose

Organizational	**Informational**	**Evaluative**	**Modalizing**
and so on	a number of	as a result of	can be used to
as shown in fig	a wide range of	it is important to	it may be
in relation to the	at the time of	one of the most	likely to be
in relation to	in rylands v fletcher		
in the context of	in terms of the		
	in terms of		
	in the absence of		
	in the case of		
	in the form of		
	in the united states		
	in this case		
	on the basis of		
	per cent of the		
	per cent of		
	some of the		
	terms of the		
	the basis of		
	the effect of		
	the house of lords		
	the house of commons		
	the number of		
	the size of the		

Compared to the BAWE essays, then, academic prose is even more informational in its nature, as seen in the amount of informational ngrams in Table 4 (44.4% for BAWE) vs. Table 5 (66.7% for academic prose). The other major difference is found in the evaluative function, where the ngrams studied in academic prose only reached 9.1%, while the BAWE essays were clearly evaluative with 27.8%. The BAWE material also showed a slightly more modalizing tendency with 13.9% vs. 9.1% for academic prose. The two categories were fairly similar in terms of organizational elements used (15.1% vs. 13.9%).

As was the case with the BAWE essays, most of the ngrams are informational, typically containing nouns, both proper and common. While the common nouns in the BAWE essays are mainly concrete, topic-dependent nouns (e.g. *poem* and *reader*), the nouns featuring in the informational category in academic prose are more universal, and perhaps even abstract, e.g. *terms* and *basis*.

While academic prose is found to be overwhelmingly informational, the BAWE essays are typically informational *and* evaluative in nature. In

Moon's terms, then, it appears that academic prose is typically ideational in nature, while the English essay is both ideational and interpersonal.

Although we have looked in more detail at the ngrams that only figure among the top 30 of one text-type, it should also be noted that a fair amount of the trigrams and quadrigrams shared between BAWE and academic prose are structuring devices, i.e. organizational according to Moon's model, e.g. *and so on*, *at the beginning of*.

In this study, Stubbs and Barth's claim that ngrams are text-type type discriminators has been confirmed. In addition, and perhaps more important, the present study has demonstrated that, by taking ngrams as our starting point in a functional analysis of recurrent word-combinations, we are able to form better conclusions as to how text-types differ. This is true even in cases where the ngrams do not constitute full semantic units, nor full phrases in traditional terms.

6. Concluding remarks

In this study we have followed Stubbs and Barth (2003) in using recurrent word-combinations as a means of highlighting differences between groups of texts. Of the three sub-corpora we investigated, two were taken from BNC-Baby, fiction and academic prose, and one from the BAWE corpus, English Studies essay. Although the method put forward by Stubbs and Barth (2003) seems to work well in many respects, it is obvious that certain noun combinations in the BAWE essays tell us more about the subject matter of the texts than it does about the text-type in general, e.g. combinations such as *Heart of Darkness* and *The Good Soldier*. In this context it is important to stress that the BAWE material is relatively restricted compared to the two other text groups. On the other hand, combinations including words such as *poem* and *reader* are good indicators of English essays as a specific text-type. Furthermore, in the comparison of ngrams across academic prose, fiction and student essays, clear differences emerged. Academic prose and fiction showed more clear tendencies in the use of recurrent ngrams than the BAWE essays. Although it may be argued that this is due to the size of the material analysed here, since clearer tendencies might be expected in a larger material, we saw in Section 2 that size does not play a major role when investigating the top 30 ngrams.

A functional investigation of the most frequent trigrams and quadrigrams in the essays and academic prose was also carried out. It was found that most ngrams unique (among the top 30) to academic prose were overwhelmingly informational, while in BAWE they were informational or evaluative, leading to the conclusion that academic prose is highly ideational, whereas the English essay is both ideational and interpersonal. This confirms what we might have expected to be typical features of essays commenting on or analysing literature, which in many ways is a subjective activity and would call for use of interpersonal language (evaluative), such as personal pronouns, verbs of evaluation or attitude, etc. In addition, there also seems to be an urge to be informative, where the informative perhaps could be seen as a basis for the evaluative.

The overlap between academic prose and English essays is also worth noticing. The academic setting of essay writing triggers certain recurrent word-combinations commonly used in academic writing in general (e.g. *in order to*, *the fact that*, in Table 1), showing that students clearly make an effort to meet the academic writing practices that are expected, and as such, the English essay is a sub-category of academic writing.

The study reported here shows that even a small-scale investigation of recurrent word-combinations can tell us something about text-type. The method explored here paves the way for similar studies on text-types across more and different kinds of corpora. With corpora such as BAWE, and its American counterpart MICUSP,[13] more could be found out about similarities and differences across student writing. This would presumably show more clearly which combinations contribute to distinguish essays from other assignment types. Further studies could also include student vs. professional writing or native vs. non-native student writing in order to investigate how salient ngrams really are and to what extent the same patterns and functions are used by apprentice and professional writers, or by learners and native speakers.

---

[13] Michigan Corpus of Upper-Level Student Papers, http://micusp.elicorpora.info/, accessed 28 Oct. 2009.

References

Primary Sources—Corpora

BNC-Baby v1.0, designed and constructed by Y. Berglund and M. Wynne, Oxford Text Archive. CD edited and compiled by L. Burnard, Oxford University's Research Technologies Service, Aug. 2004, <http://www.natcorp.ox.ac.uk/corpus/baby/manual.pdf>, accessed 28 Oct. 2009.

British Academic Written English Corpus, <http://www.coventry.ac.uk/researchnet/d/911>, accessed 28 Oct. 2009.

Secondary Sources

Altenberg, B. 1998. "On the phraseology of spoken English: The evidence of recurrent word-combinations." In A.P. Cowie (ed.), *Phraseology. Theory, Analysis, and Applications*. Oxford: Oxford University Press, 101-122.

Biber, D. & S. Conrad. 1999. "Lexical bundles in conversation and academic prose." In H. Hasselgård & S. Oksefjell (eds.), *Out of Corpora. Studies in Honour of Stig Johansson*. Amsterdam / Philadelphia: Rodopi, 181-190.

Biber, D., S. Johansson, G. Leech, S. Conrad, and E. Finegan. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Longman.

Charles, M. 2006. "The construction of Stance in Reporting Clauses: A Cross-disciplinary Study of Theses." *Applied Linguistics 27/3*: 492-518.

Culpeper, J. and M. Kytö. 2002. "Lexical Bundles in Early Modern English dialogues: A window into the speech-related language of the past". In T. Fanego, B. Méndez-Naya, and E. Seoane (eds.), *Sounds, Words, Texts and Change*. Selected Papers from 11 ICEHL, Santiago de Compostela, 7–11 September 2000. Amsterdam/Philadelphia: John Benjamins, 45–63.

Ebeling, S.O. & P. Wickens. Forthcoming. "Interpersonal themes and author stance in student writing." To appear in *Proceedings from the 30th ICAME Conference*, Lancaster 2009.

Groom, N. 2009. "Phraseology and epistemology in academic book reviews: A corpus-driven analysis of two humanities disciplines." In

K. Hyland & G. Diani (eds.) *Academic Evaluation. Review Genres in University Settings*. Basingstoke: Palgrave MacMillan, 122-139.

Heuboeck, A., J. Holmes, and H. Nesi. 2007. *The BAWE Corpus Manual*. Available under ‘BAWE documentation at http://www.coventry.ac.uk/researchnet/d/505/a/5160, accessed 28 Oct. 2009.

Martin. J.R. 1992. *English Text. System and Structure*. Philadelphia / Amsterdam: John Benjamins Publishing Company.

Martin. J.R. 1997. “Analysing genre: functional parameters.” In Christie, F. and J.R. Martin (eds.), *Genre and Institutions: Social Processes in the Workplace and School*. London: Continuum, 3-39.

Moon, R. 1998. *Fixed Expressions and Idioms in English. A Corpus-Based Approach*. Oxford: Clarendon Press.

Nesi, H. and S. Gardner. Unpublished manuscript. “Families of genres of assessed writing.”

Nesi, H., S. Gardner, R. Forsyth, D. Hindle, P. Wickens, S.O. Ebeling, M. Leedham, P. Thompson, and A. Heuboeck. 2005. “Towards the compilation of a corpus of assessed student writing: An account of work in progress.” In Danielsson, P. and M. Wagenmakers (eds.) *Proceedings from The Corpus Linguistics Conference Series*, Vol. 1, No.1 <http://www.coventry.ac.uk/researchnet/external/content/1/c4/30/09/v1192628796/user/towards_compilation.pdf>.

Pecorari, D. 2008. “Repeated language in academic discourse: the case of biology background students.” *Nordic Journal of English Studies*, vol. 7 No. 3. 9-33.

Scott, M., 2008. WordSmith Tools version 5, Liverpool: Lexical Analysis Software.

Scott, M. and C. Tribble. 2006. *Textual Patterns. Key Words and Corpus Analysis in Language Education*. Amsterdam / Philadelphia: John Benjamins.

Shaw, P. 2009. “Linking adverbials in student and professional writing in literary studies: What makes writing mature.” In M. Charles, D. Pecorari, and S. Hunston (eds.) *Academic Writing: At the Interface of Corpus and Discourse*. London: Continuum, 215-235.

Sinclair, J. 1999. “A way with common words.” In H. Hasselgård & S. Oksefjell (eds.), *Out of Corpora. Studies in Honour of Stig Johansson*. Amsterdam / Philadelphia: Rodopi, 157-180.

Stubbs, M. 2004. "On Very Frequent Phrases in English: Distributions, Functions, and Structures." Plenary lecture at ICAME 25, Verona, Italy, 2004. http://web.archive.org/web/20070828004603/http://www.uni-trier.de/uni/fb2/anglistik/Projekte/stubbs/icame-2004.htm, accessed 20 April 2010.

Stubbs, M. & I. Barth. 2003. "Using recurrent phrases as text-type discriminators. A quantitative method and some findings." *Functions of Language*, 10 (1), 61-104.

Swales, J.M. 1990. *Genre Analysis*. Cambridge: Cambridge University Press.

Appendix 1

Top 50 trigrams in the three text-types in alphabetical order (academic prose, fiction, and BAWE essays)

BAWE essays	**Academic prose**	**Fiction**[14]
a sense of	a number of	a couple of
an example of	and it is	a kind of
as it is	and so on	a lot of
as well as	as a result	and it was
at the end	as well as	as if he
can be seen	because of the	as soon as
due to the	but it is	at the end
end of the	end of the	back to the
heart of darkness	in order to	be able to
image of the	in relation to	but it was
in order to	in terms of	do you think
in the first	in the case	end of the
in the novel	in the first	for a moment
in the poem	in this case	going to be
in which the	in which the	had been a
is able to	is likely to	he did not
it is a	is not a	he had been
it is not	it can be	he was a
it is possible	it has been	i don’t know
it is the	it is a	i want to
nature of the	it is not	in front of
of the city	it is possible	it had been
of the novel	it may be	it was a
of the play	likely to be	it was the
of the poem	many of the	it would be
of the text	on the other	must have been
on the other	one of the	on to the
one of the	part of the	one of the

[14] WordSmith seems to counts e.g. inverted commas as part of ngrams. For the purpose of this study I have not included trigrams or quadrigrams of the type *he said “*, where the inverted comma is looked upon as the third word of the trigram.

such as the	per cent of	out of the
that it is	shown in fig	part of the
the beginning of	some of the	she did not
the end of	terms of the	she had been
the fact that	that it is	shook his head
the good soldier	that there is	side of the
the idea of	the basis of	that he had
the image of	the case of	that he was
the importance of	the effect of	that it was
the other hand	the end of	that she had
the poem is	the fact that	that she was
the reader is	the form of	the back of
the reader to	the house of	the end of
the use of	the number of	the first time
the way in	the other hand	the rest of
the winter's tale	the terms of	there was a
there is a	the united states	there was no
there is no	the use of	to be a
to be a	there is a	to do with
to the reader	there is no	to have a
use of the	this is not	was going to
way in which	to be a	would have been

Appendix II

Top 50 quadrigrams in the three text-types in alphabetical order (academic prose, fiction, and BAWE essays)

BAWE essays	**Academic prose**	**Fiction**
allows the reader to	a wide range of	a cup of tea
and the good soldier	as a result of	as if he was
as can be seen	as shown in fig	at the back of
as well as the	as we have seen	at the bottom of
at the beginning of	as well as the	at the end of
at the end of	at the end of	at the same time
at the same time	at the same time	at the top of
be seen in the	at the time of	did not want to

by the use of
can be seen in
could be argued that
gender and gender roles
importance of being earnest
in contrast to the
in heart of darkness
in the form of
in the good soldier
in the winter's tale
is an example of
it could be argued
it is clear that
it is possible to
mise en scene
of the poem the
on the other hand
rest of the poem
tess of the d'urbervilles
that there is no
that there is a
the beginning of the
the death of the
the end of the
the extent to which
the fact that the
the image of the
the importance of the
the importance of being
the nature of the
the rest of the
the role of the
the structure of the
the use of the
the use of language
the way in which
the ways in which

by the fact that
can be used to
in relation to the
in rylands v fletcher
in terms of the
in the absence of
in the case of
in the context of
in the course of
in the form of
in the united states
in this case the
is likely to be
it is clear that
it is difficult to
it is important to
it is necessary to
it is possible to
more likely to be
on the basis of
on the other hand
on the part of
one of the most
per cent of the
regions ii and iii
rule in rylands v
that there is a
that there is no
the basis of the
the court of appeal
the end of the
the extent to which
the fact that the
the house of lords
the house of commons
the nature of the
the rest of the
the rule in rylands

for the first time
from time to time
got to his feet
he shook his head
he was going to
I don't want to
in front of him
in front of the
in one of the
in the direction of
in the first place
in the middle of
it had been a
it was as if
it would have been
nothing to do with
on the edge of
on the other side
other side of the
out of the window
she shook her head
she was going to
that sort of thing
the back of the
the back of his
the bottom of the
the centre of the
the edge of the
the end of the
the middle of the
the other side of
the rest of the
the side of the
the top of the
the two of them
there had been a
there was no sign

through the use of
to make sense of
to the fact that
way in which the
with the use of

the size of the
the structure of the
the way in which
the ways in which

to be able to
was going to be
was one of the
what do you mean
what do you think

Spanish EFL undergraduate students' perceptions of learning styles

Joana Salazar Noguera, University of the Balearic Islands

Justin Wageman, North Dakota State University

Abstract
This paper presents the findings of research on learning styles carried out in a Spanish EFL higher education context. The specific objectives of the study were to investigate students' perceptions of their learning styles and to determine students' responses to the identification of their learning styles using the Kolb Learning Style Inventory (LSI) adapted by McCarthy (1980). Participants included 53 students from an English Philology degree program. The results showed that most of the students had a Diverging style and an Accommodating style. The findings also suggested that students had a positive reaction to the identification of their learning styles, despite its novelty in the Spanish EFL university context. Most of the students' perceptions of their own learning styles concurred with the results obtained from the Kolb LSI. More research on educational practices and university instruction that have to do with learning style concepts is needed.

Keywords: learning styles, Kolb Learning Styles Inventory (LSI) adapted by McCarthy (1980), English as a Foreign Language (EFL), students' perceptions

Introduction

Since 1980 the concept of learning styles has received a lot of attention in research, principally in the field of psychology and education (Cano-Garcia & Hughes, 2000; Coffield, Mosely, Hall, & Ecclestone, 2004). Moreover, learning styles research has also been applied to the field of Second Language Acquisition (SLA) (Reid, 1995; Ehrman, 1996; Ehrman & Oxford, 1990; Oxford, 1990; Oxford & Ehrman, 1993; Oxford, Ehrman, & Lavine, 1991; Chapelle, 1992; Chapelle & Roberts, 1986). Learning styles and diversity in the foreign language classroom have been considered crucial factors mainly due to their close relationship to learning strategies. As Reid (1995) points out, "my own classroom research in learning styles began when I recognized the diversity of learning styles in my ESL classes…[T]here are substantial individual differences among students' preferred styles and their selected use of strategies" (p. 300).

Within the scope of SLA, many learning style classifications have been developed, such as field dependence/independence, tolerance, perception, reflection or impulsivity, and multiple intelligences (Ehrman, 1996; Griffiths & Sheen, 1992; Reid, 1995; Skehan, 1998). A plethora of definitions also exist for the concept of learning styles (Dunn & Griggs, 2000; Lemire & Gray, 2003; Center on Disability Studies, n.d.; Sims & Sims, 1994). For instance, Keefe and Ferrell (1990) define style as "a complexus of related characteristics in which the whole is greater than its parts. Learning style is a gestalt combining internal and external operations derived from the individual's neurobiology, personality, and development and reflected in learner behavior" (p. 59).

Essentially, learning styles refer to the way information is processed, and each individual processes information in a unique way. The existence of different learning styles shows not only that there are various ways of processing information, but also that each learning style may have strengths and weaknesses. This view is supported by Dörnyei and Skehan (2003) who claim that "different styles may be equally valid and advantageous. It is possible to envision all styles as making contributions, even if in different domains" (p. 602).

Studies by researchers such as Marton or Carroll are crucial to the literature of learning styles, as they conclude that the knowledge of learning styles may predict "academic achievement" and may also influence the "improvement of teaching-learning processes" (Zywno, 2003, p. 12). Carroll was the first researcher to clearly show the influence of individual learning styles on academic achievement (Carroll, 1963; Henson & Borthwick, 1984; Claxton & Murrell, 1987; Bedford, 2006). Up to that point, the preeminent belief was that Intelligence Quotient (IQ) was the only predictor of academic achievement. More recently, Gardner (2005) reported how focusing on individual differences in second language acquisition is most appropriate when analyzing the degree of association between motivation and achievement. This conclusion was supported by Horwitz (1995) who posited that student affective variables represent the willingness to engage in the activities necessary to enhance second language attainment.

Van Zwanenberg, Wilkinson, and Anderson (2000) and Zywno (2003) considered the existence of other variables, such as the variety of learning and teaching styles or academic performance expectations, which could also influence student output. This notion was shared by

Ellis (1989) who reported on a study of two adult learners of L2 German and how their learning styles vary to cope with the instruction method provided.

In SLA higher education contexts, instruments to assess learning styles have not been sufficiently developed with a high degree of validity and reliability. An exception to this was Oxford and Burry-Stock (1995) who described applications of the Strategy Inventory for Language Learning, an instrument found to have high reliability with many different cultural groups learning a foreign language. More promising work was done by Wintergerst, DeCapua, and Verna (2003) who published the positive results of the testing conducted on a learning styles instrument to assess foreign language students. DeCapua and Wintergerst (2005) have contended that a triangulated approach using a questionnaire, semi-structured oral interviews, and participant observations to measure learning styles in EFL contexts would present a fuller picture of instrument validation. It is clear that more extended use of these assessment tools is needed.

As a consequence of the need for further instrument development, planning for all students' learning styles in the EFL context becomes a complex task for foreign language instructors. Indeed, Ellis (1992) points out that "matching is best achieved by the teacher catering for individual needs during the moment-by-moment process of teaching" (p. 188). Moran (as cited in Bedford, 2006) asserts that it is important that teachers' own learning styles be appropriate for their students' learning style, arguing for the "desirability of achieving a match between the learning styles of instructors and those of their students" (p. 28). Therefore, matching students' learning styles to instructional techniques may influence learning significantly. However, identification of learning styles is not a common educational practice yet. In some educational contexts, the use of learning style inventories to detect different students' learning styles has only been done in the interest of matching the diverse cultures in the classroom (Hickcox, 2006) or examining the role culture plays in the way individuals learn (Joy & Kolb, 2009).

As reviewed so far, the topic of individual learning styles in language learning is complex and has provided little conclusive knowledge (Ehrman, Leaver, & Oxford, 2003; Peterson, Rayner, & Armstrong, 2009). Even if learning style identification and matching theories have helped to focus learning style research in recent years

(Sims & Sims, 2006), the practice of learning style identification among Spanish EFL practitioners is still rare. This may be due to a certain reticence among instructors in using teaching methods that consider affective variables, like learning styles, or it may be due to a lack of familiarity with learning style inventories. The Kolb (1985) Learning Style Inventory (LSI) adapted by McCarthy (1980) is an instrument that has traditionally been widely used in psychological educational contexts; however, it has not been utilized in EFL higher education contexts. Therefore, this study uses the Kolb LSI in a Spanish EFL university classroom to observe undergraduate students' perceptions of their identified learning styles. The research questions were as follows:

1. What are the Spanish EFL undergraduate students' learning styles according to the Kolb Learning Style Inventory?
2. What are the Spanish EFL undergraduate students' reactions to using the Kolb Learning Style Inventory in the EFL classroom?

To answer these questions, a study was conducted at the University of the Balearic Islands (UIB) in Palma de Mallorca, Spain.

Method

Participants

The participants in this study were all first-year students (N=53) enrolled in the English Philology degree program at the UIB. There were 42 female and 11 male students. The average age of the students was 19 years.

The English Philology program takes a traditional approach to teaching and assessment. Instruction includes a primary emphasis on traditional pedagogy (Kauchak & Eggen, 2008), although a segment of the curriculum does include the theories and principles of communicative methodologies in the teaching of a foreign language (Cook, 2000). In addition to guided lectures and explicit instruction, activities such as oral presentations, research panels, and interdisciplinary study are also included in each course.

Students are evaluated by means of diverse performance tasks; in most courses, there are three to four traditional objective examinations, a comprehensive final essay examination, and a written project.

Instruments

All students completed two research instruments. The first was the Kolb LSI, an instrument selected because it gathers data on the way individuals receive and process information. Moreover, it has been found to be suitable in tertiary education settings (Hickcox, 2006). During administration of the Kolb LSI, the researcher checked whether the test items were understood by the students and taught students how to calculate their scores.

Four learning styles are distinguished in this inventory: Diverging, Assimilating, Converging and Accommodating. Divergers tend to be creative, and they like to learn in settings that favor affective variables and group work. Individuals with a predominantly Assimilating style learn by organizing ideas in a logical way, and they like to learn abstract concepts. Students with a Converging style are good at practical applications of concepts and ideas. Individuals with an Accommodating style are good at learning with others. Their dominant learning modalities are active experimentation and concrete experience.

The second instrument was a follow-up questionnaire (See Appendix 1) administered to the students to determine their perceptions of the value of identifying their learning styles. The questionnaire was specifically developed for this study and was used to observe student reaction to having worked on learning styles in the EFL classroom. The questionnaire consisted of eight questions. Each of the first six questions included two different parts. The first part was a Yes/No question, which mainly aimed at gathering feedback on the students' reactions to the Kolb LSI. The second part was an open-ended question where students had to state the reason for having chosen yes or no in the first part. Question 1 asked whether or not the students liked having to identify their own learning style using the Kolb LSI. Questions two and three requested that respondents indicate if they had taken the Kolb LSI before, or if they had taken another learning styles inventory. Questions four and five solicited the students' opinion on the usefulness of the Kolb LSI for students and teachers. Question six asked the informants about

the extent of their agreement or disagreement with their Inventory scores. Finally, the questionnaire included two additional open-ended questions (i.e., 7 and 8) in which respondents listed the characteristics they believed most represented themselves as learners, as well as commented on their experience and perceptions in using the Kolb LSI.

The responses to all the open-ended questions were analyzed for patterns and categorized. The categories that emerged include all the various answers the informants provided. Open-ended questions were used in this study since the purpose was to investigate the students' overall reactions to having used the Kolb LSI without directing their responses through previous wording. As the respondents' answers were not limited, open-ended questions allowed informants to include more information related to their attitudes and feelings towards the use of the Inventory in the EFL classroom.

The Kolb LSI and the questionnaire were administered during the second semester of the academic year 2007. Students in the present study worked in four one-hour sessions on the identification of their own learning styles and on the awareness of the importance of identifying learning styles. After test and questionnaire administration, general group discussions on learning styles ensued.

Results

In this section, both the quantitative and qualitative results are presented. The results from the Kolb LSI show that 33% of the sample population appear to have a predominantly Diverging learning style, 25% have an Accommodating learning style, and 23% show a predominantly Converging learning style (see Figure 1 below). About 20% appear to have an Assimilating learning style. Thus, a tendency is observed favoring the Diverging, Accommodating and Converging learning styles, with the Diverging learning style showing a slightly stronger tendency among these three styles. Individuals with a Diverging learning style have Concrete Experience (CE) and Reflective Observation (RO) as their dominant learning skills and are generally interested in people and culture. The Assimilating learning style appears to have the poorest presence.

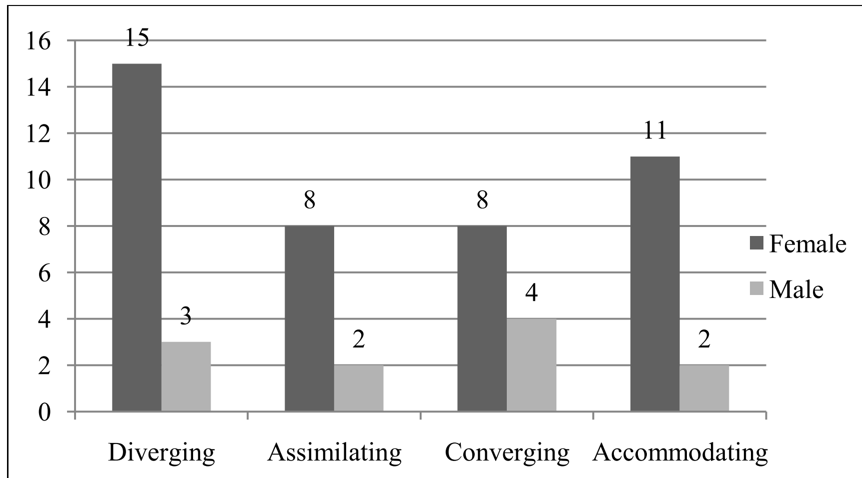


Figure 1. EFL Spanish Undergraduate Students' Results on the Kolb LSI

Table 1 below shows the proportion observed for each one of the four samples: Diverging, Assimilating, Converging and Accommodating. In this case, there are no significant differences among the samples using a confidence level of 95%.

Table 1. ANOM Report

95% Decision Limits

UDL = 0.52
Central line = 0.26
LDL = 0.00

Number of samples outside the limits = 0

* = Outside the limits

Sample	Size	Proportion
1	18	0.339623
2	10	0.188679
3	12	0.226415
4	13	0.245283

Table 2 below includes the StatAdvisor, a procedure that tests the hypothesis that all the mean proportions in the four samples are identical. It also generates a graphic analysis of means to determine which samples are significantly different from the global mean. Given a p-value equal to or higher than 0.10, there are no significant differences between the samples using a confidence level of 90% or higher.

Table 2. The StatAdvisor

Analysis of Means—Binomial proportions

Data/Variables: Col_2

Number of samples = 4
Mean sample size = 13.25
Mean proportion = 0.262371

Chi-square comparison

Chi-square	GL	P-Value
0.94	3	0.8169

A chi-square test was run to test the hypothesis that the row and the column selected are independent. Since the p-value is equal to or higher than 0.10 (see Table 3 below), we cannot reject the hypothesis that the rows and columns are independent. As a result, the row observed in a particular case may not be related to its column.

Table 3. Chi-square comparison

Chi-square	GL	P-Value
1.57	3	0.6665

Note: The frequency of some cells is less than 5

Figure 2 below shows that there are no significant differences between male and female respondents in relation to the four learning styles.

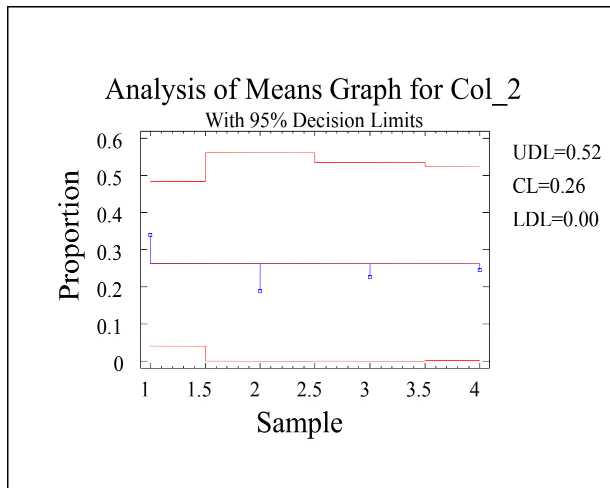


Figure 2. Analysis of Means

Students provided a wide range of responses to the questionnaire used to determine their perceptions about the identification of individual learning styles (Figure 3 below). Questionnaire results include the responses to both close-ended as well as open-ended questions. The open-ended responses were categorized and frequencies were displayed with graphs. The categories are not exclusive as respondents may have provided answers that could be included in more than one category. Representative quotes of the qualitative data are included in this section.

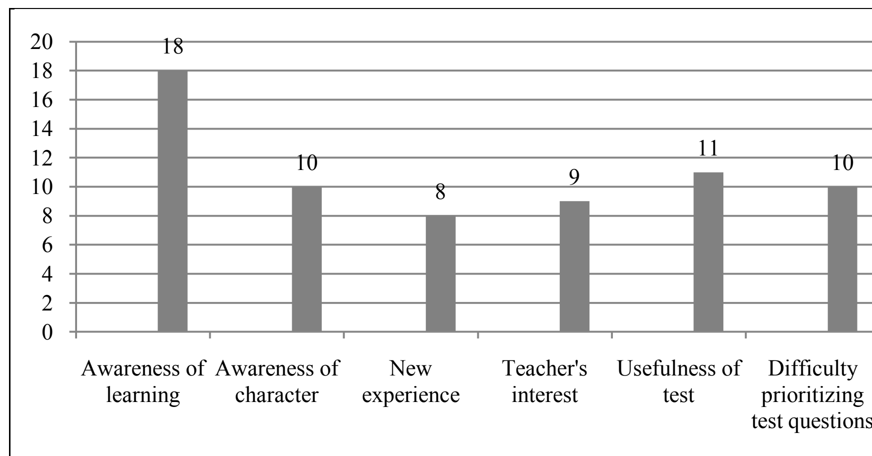


Figure 3. Students' Perceptions of the Use of the Kolb LSI

More than one quarter of the students (i.e., 27%) responded that having identified their learning style through the Kolb LSI helped them become aware of their own way of learning. In fact, many pupils commented that the Kolb LSI was a self-awareness learning tool. A student remarked, "I think it has helped me to understand a little bit more what my way of learning is."[1]

Other respondents referred to the test's potential for developing awareness of how to improve learning. As one student pointed out, "I can say that this activity has helped me to better see how I study and how I could continue to improve studying." A second student said, "I think that it is a good way of knowing how we can make the most of our studying time," while another said, "I think that doing this activity on the questionnaire helps people who need to create a good study plan for themselves."

Some students referred to the utility of the Kolb LSI as a way to become more aware of the importance of a learning style in the learning context. In particular, a student mentioned, "It is a good way to create a method of how to study in relation to both who we are and our way of studying, or even better to be able to face our studies." Another affirmed, "I think that it was an interesting way to analyze myself and to analyze

[1] Students' comments, which were written in English, the students' target language, have been included in their original version.

my way of working." Similarly, one student said, "Having to answer those questions has helped me to reflect on how I study and learn, and it has made me think that maybe I don't do it in a way that would be better."

Some students added that the Kolb LSI had not only been useful for the knowledge gained about their own way of learning, but also for the awareness of other ways of learning. For example, one student commented, "I think that this questionnaire is very complete because it makes you rethink what your learning system is. Because the truth is that we all follow one system in particular without stopping to think about other possibilities we have and that we miss." Another student commented, "I have learned other learning techniques, which I didn't know."

Some students also pointed to the usefulness of the Inventory in determining one's learning behavior (17%) (see Figure 3 above). This idea is illustrated with the following student statement: "I like this type of test because it is useful in determining learning styles and behavior." Yet another explained, "I think that the questionnaire is quite useful because it makes us see how we act in certain situations."

As Figure 3 shows, various students believed that having completed the Kolb LSI in class demonstrated that the teacher was interested in them as students (14%). One respondent reflected, "I think that these questionnaires are a good idea as they make us realize how we study and how we can learn in a better way. Moreover, this shows that the teacher is interested in the pupils' needs."

Many students also thought that the information included in the LSI was useful for the teacher. The information compiled from the results may be able to inform teaching methodology and help improve actual teaching practice. In this respect, one student concluded, "The experiences I have in participating actively in class are good, and I think that the "Learning Styles questionnaire" is very interesting and positive for the students as well as for the teachers because students become aware of and realize how they have to study and how they should do it, and teachers can see what strategies are more effective when teaching their subject, and how they have to do it to improve their classes day by day."

Some students also revealed that knowing about their learning styles through the Kolb LSI had not only been useful for them as students, but

had also helped them to know more about themselves, mainly about their character and personality (14%) (see Figure 3 above). In this regard, one student noted, “I think these types of activities should be done in order to better know ourselves.” Another stated, “It is like a test which reflects certain aspects of your personality and your way of being.” Still another student added, “I have learned many useful things that not only apply to other participants, but I can also apply them to other things. For instance, I now dare to watch TV series in their original version. I have also found out what things I am more interested in.” One respondent not only referred to the usefulness of knowing about oneself, but also to the fact that the students had enjoyed themselves by engaging in the LSI activity. In that respect, a student said, “The truth is that I think it is a great idea. It has a very good purpose and a fun way of learning about ourselves.”

For many students, completing the Kolb LSI helped them to become self-aware of certain areas that they had previously not reflected on. For example, one student said, “I noticed after reading some of the answers from the questionnaire that I should try to get to know myself even better, because that is clearly going to affect the way I learn.”

In spite of all these positive comments about the Kolb LSI, students also pointed out some drawbacks. As Figure 3 shows, some students responded that it was somewhat difficult for them to establish a priority among all the questions included in the questionnaire (15%). One pupil made the following statement, “Some questions were a bit difficult but I like it.”

Besides some difficulties that they had found in answering the Kolb LSI, students acknowledged that completing a learning styles inventory was something new to them (12%) (see Figure 3 above). A large number of students made statements such as the following: “I had never questioned myself on many of the things included in the questionnaire” and “I have thought about things I don’t usually think about.” Others said, “It was something new and nothing similar to what I had ever done before” and “I had never done a questionnaire of this type. I think it is a good idea to determine what style is better than another one when studying.” Lastly, one student added, “When I did the styles questionnaire I had to pay attention to the way I used to learn. It was something that I hadn’t paid attention to before because I suppose that it is an accepted fact that everyone does it.”

Students in general complained about the limited number of activities on learning styles in the course. In particular, one student opined, “I think that it is a very good idea to complete the questionnaire, and I think that we should do more throughout the course to know our ways of doing things and to know how to improve.”

Furthermore, many students complained about the fact that tasks on learning styles could have been done earlier so as to achieve better academic results. Among the students, one respondent posited, “If I had used the techniques of the questionnaire I would have passed this subject; so I think that it is somehow useful.”

By using the Kolb LSI in class, most students perceived that a completely new teaching approach was being used. One student remarked, “I think that these new trends in teaching have to be welcomed, and a change in the field of teaching and education will always be productive and beneficial.”

On the second question of the perceptions questionnaire, students were asked which characteristics they thought most corresponded to them as learners. They provided a wide range of responses (see Figure 4 below). Fourteen percent of the students mentioned that they preferred learning by observing; thus, these are individuals who perceive information concretely, a finding which concurs with the one third (33%) of the population interviewed who reported a Diverging learning style (see Figure 1). Ten percent of the students indicated a preference for learning by getting involved in activities, a characteristic also portrayed in people with the Diverging learning style who generally need to be personally involved in tasks. Seven percent of the informants considered themselves to be responsible, and seven percent also asserted that they take their time before acting. This would also seem to correspond to the characteristics of seeking commitment and of valuing insightful thinking, both of which are common traits in Divergers. Three percent of the students affirmed that they accept people or situations as they are, which is a characteristic of Divergers. Some students said they liked learning things through feelings, and many students also added that they usually had lots of questions (6%).

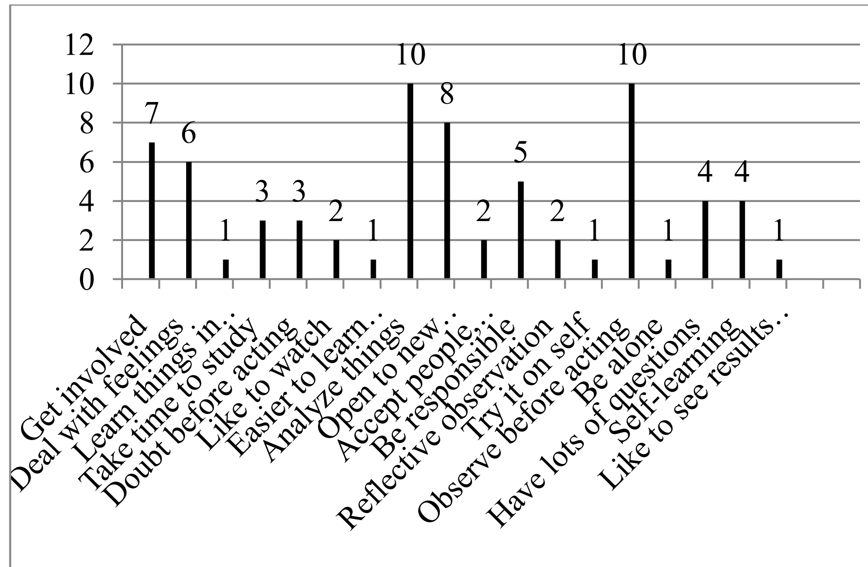


Figure 4. Students' Characteristics that Correspond Most to Them as Learners

Twelve percent of the students considered themselves open to new experiences, which would correspond to the 25% of students who showed an Accommodating learning style (see Figure 1) and who therefore like to "involve themselves in new and challenging experiences" (Kolb & Kolb, 2006, p. 50). Some students indicated that they prefer to work by trying things out for themselves, which also points to a trait of Accommodators who prefer to learn by trial and error.

Fourteen percent of the students said that they preferred learning by analyzing things. This would correspond to the 23% of students who showed a Converging learning style.

Some students (3%) thought of themselves as users of the skill of Reflective Observation (RO), which would correspond to either Divergers or Assimilators. In addition, others commented on being energetic and enthusiastic, or seeing results as a consequence of their own efforts.

By answering the questionnaire, some students realized that they showed more than one pattern of learning behavior and therefore could be included in more than one learning style category. As one student pointed out, "I think that liking to break things down into pieces and

always evaluating the possibilities makes me more of an AC person, although the other type, RO, has some characteristics that are most like me too."

All in all, the results in Figure 4 confirm the presence of three main learning styles (viz., Diverging, Accommodating and Converging) among the students and correspond in general to the results presented in Figure 1.

Finally, as to student reaction to the results of the Kolb LSI (see Figure 5 below), most of the students (91%) agreed with their scores, which reveals that the information obtained from the Kolb inventory coincided with the students' previous perceptions of their own learning style. In this regard, one student pointed out, "I think that the Learning Styles questionnaire is quite interesting and in my case it is quite accurate and right with my learning style, which is concrete and reflective."

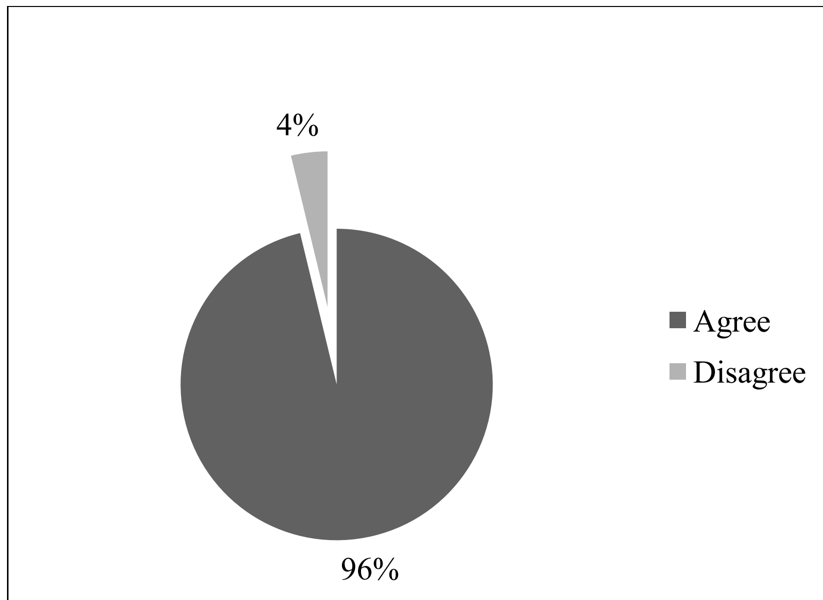


Figure 5. Students' Agreement with the Kolb LSI Results

1. Student comments, which were written in English, the students' target language, have been included in their original version.

Discussion

The objective of the present study was twofold. First, it aimed at investigating the learning styles of Spanish undergraduate students by means of the Kolb LSI, and, second, it aimed at determining students' perceptions of their own learning style identification by means of an opinion questionnaire.

As for the first research question, the results on the Kolb LSI indicated that most undergraduate students studying for an English Philology degree at the University of the Balearic Islands, Spain, had either a Diverging, Accommodating or Converging learning style. Among these, the Diverging style had the highest percentage. Among the sample of undergraduate learners, female informants were a decided majority, comprising about 79% of the total sample population, while the male presence in the study did not reach the one-quarter mark (21%). Yet, the results also revealed that most female students had a Diverging learning style while most male students had a Converging learning style. It could be that the males in this group were skill-oriented and learned by problem-solving and decision-making. Our results differ from two other studies (Contessa, Ciardiello & Perlman, 2005; Mammen et al., 2007) with participants who were surgical students. Those results revealed Converging as the most commonly occurring style and suggested that individual learning styles may be constant throughout residency training.

In the present study the learning style which registered the lowest percentage among students was the Assimilating learning style, which could be due to the fact that Assimilating type of learners are "less focused on people and more interested in ideas and abstract concepts" (Kolb & Kolb, 2006, p. 49). Assimilators, therefore, tend to specialize in the natural sciences, maths, or research. These are subjects that do not pertain to this study's participants. This would be corroborated by the findings of Terrell (2002) whose results indicated that most doctoral students majoring in Computing Technology in Education fell into Kolb's Converger and Assimilator categories. Similarly, Demirkan and Demirbaş (2007) observed that the distribution of freshman design students was concentrated in the assimilating group.

As for the second research question, the results showed that the Spanish EFL undergraduate students generally had a positive reaction to using the Kolb LSI in the EFL classroom. More specifically, the qualitative findings obtained through the questionnaire showed that the

Kolb LSI was an appropriate tool to promote self-awareness of learning, identification of learning style, and the importance of its use in the classroom. This agrees with the contention that the Kolb LSI is "the most widely used adult-oriented inventory emphasizing information processing and cognitive personality style models" (Hickcox, 2006, p. 5). The findings further support those by Lashley and Barron (2006) who strongly suggest using the Kolb's experiential cycle as a way of approaching the learning needs of students, as well as encouraging the development of balanced learning strategies that lead to reflective practice.

The qualitative findings also show that the results obtained from the Inventory were in accordance with the students' previous perceptions of their own learning style. The application of the Kolb LSI in an EFL context also made for an increase in the knowledge of learning styles among the students. Some of the learning styles may have differed from a student's own particular learning style and thus favored an awareness of how to improve learning.

Additionally, the qualitative findings from the undergraduate students' perceptions also suggest that the application of the Kolb LSI in an EFL context may positively influence the praxis of the teacher and allow for teaching method selection based on the student learning style landscape. The findings of Thomas and McKay (2010) provide evidence of the improvement of learning outcomes when instructional material is matched to students' cognitive styles.

The qualitative results suggest just how novel it is to use the Kolb LSI in the foreign language classroom. Unfortunately, traditional teaching methodologies still dominate in the Spanish EFL university context in which little attention is given to learning styles. Terrell (2002) advised that educational institutions should be prepared to address learning style issues when developing and offering formative programs. That has not happened yet.

Undergraduate students encountered some difficulties in categorizing their styles when completing the Kolb LSI. This may be due to the students' lack of familiarity with these types of tools in particular, and the rarity of tackling issues related to the learning process in the EFL classroom in general. Similar challenges were described by Arthurs (2007) among nurse educators in a nursing program. Still, the results of

the present study show a generally positive reaction to the use of the Kolb Inventory in the EFL classroom.

Conclusions and Recommendations

The results obtained in this study are generally supportive of using tools, such as the Kolb LSI, to assist in identifying student learning styles in the second language classroom and to help especially in enhancing student awareness of various learning styles. More qualitative research should be conducted to provide a more accurate picture of EFL educational practices that address various learning style approaches (or the lack thereof) in the curricula. It would also be useful to control for the possible effect of teacher acknowledgment of the student learning style diversity, as well as for the effect of various differentiated instructional methods (Elhe, 2007). The information which may be gathered from student learning styles may "enable educators to be more constructively responsive to individual differences amongst students, and to design instruction that accords with the instructor's purposes in deliberately striving to achieve either a match or a mis-match of the instruction with individual students' learning styles" (Claxton and Murrell 1987, p. iii). Cavanagh and Coffin (2009) also believe that matching students' learning styles preferences with teaching styles is important for maximizing learning.

This study has shown that the majority of the informants responded positively to the use of the Kolb LSI, which would warrant further use of such inventories in EFL tertiary education settings. Efforts to generalize a common practice of identification of learning style should be made to promote improvement in the educational quality of EFL contexts. As Hickcox posits (2006) "It is highly recommended that educators are offered opportunities to connect learning styles to a variety of teaching approaches" (p. 14). Finally, teacher education programs should also address both theoretical and practical aspects of learning styles.

There are certain limitations of this study, such as the small number of participants, as well as the fact that they were all drawn from one group of students in a particular setting at a single educational level. Further research should consider the diversity of learning styles not only in the field of instruction, but also in educational assessment, to avoid

systematically favoring students with one particular learning style in higher education academic contexts.

References

Arthurs, J. B. 2007. "A juggling act in the classroom: Managing different learning styles." *Teaching and Learning in Nursing, 2*(1), 2-7.

Bedford, T. 2006. "Learning styles: A review of English-language literature." In R. R. Sims & S. J. Sims (Eds.), *Learning styles and learning*. New York, NY: Nova Science Publishers. 19-42.

Cano-Garcia, F., & Hughes, E. H. 2000. "Learning and thinking styles: An analysis of their interrelationship and influence on academic achievement." *Educational Psychology, 20*(4), 413-430.

Carroll, J. B. 1963. "A model of school learning." *Teachers College Record, 64*, 723-733.

Cavanagh, S. J., & Coffin, D. A. 2009. "Matching instructional preference and teaching styles: A review of the literature." *Nurse Education Today*, *14*(2), 106-110.

Center on Disability Studies. (n.d.). *Effective practices brief: Social and emotional learning*. Retrieved from University of Hawaii, Center on Disability Studies website: http://www.sig.hawaii.edu/products/briefs/sel_brief.pdf.

Chapelle, C. A. 1992. "Disembedding 'Disembedded figures in the landscape…': An appraisal of Griffiths and Sheen's 'Reappraisal of L2 research on field dependence/independence'." *Applied Linguistics*, *13*(4), 375-384.

Chapelle, C. A., & Roberts, C. 1986. "Ambiguity tolerance and field independence as predictors of proficiency in English as a second language." *Language Learning*, *36*, 27-45.

Claxton, C. S., & Murrell, P. H. 1987. *Learning styles: Implications for improving educational practices* (Report No. 4). Washington, DC: Association for the Study of Higher Education.

Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. 2004. *Learning styles and pedagogy in post-16 learning: A systematic and critical review*. Retrieved from Learning and Skills Research Centre website: http://www.hull.ac.uk/php/edskas/learning%20styles.pdf

Contessa, J., Ciardiello, K. A., & Perlman, S. 2005. "Surgery resident learning styles and academic achievement." *Journal of Surgical Education, 62*(3), 344-347.

Cook, G. 2000. *Language play, language learning*. Oxford, England: Oxford University Press.

DeCapua, A., & Wintergerst, A. C. 2005. "Assessing and validating a learning styles instrument." *System, 33*(1), 1-16.

Demirkan, H., & Demirbaş, Ö. O. 2007. "Focus on the learning styles of freshman design students." *Learning and Instruction, 17*(3), 345-359.

Dörnyei, Z., & Skehan, P. 2003. "Individual differences in second language learning." In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589-639). Oxford, England: Blackwell.

Dunn, R., & Griggs, S. 2000. *Practical approaches to using learning styles in higher education.* Westport, Connecticut: Bergin and Garvey.

Ehle, M. J. 2007. "Enhancing equity and diversity in teacher education through assessment and accommodation of pre-service teachers' learning and teaching styles." Paper presented at the ATEE 32nd Annual Conference, University of Wolverhampton, Telford Campus.

Ehrman, M. E. 1996. *Understanding second language difficulties*. Thousand Oaks, CA: Sage.

Ehrman, M., & Leaver, B. L. 2003. "Cognitive styles in the service of language learning." *System, 31*(3), 393-415.

Ehrman, M. E, Leaver, B. L., & Oxford, R. L. 2003. "A brief overview of individual differences in second language learning." *System, 31*(3), 313-330.

Ehrman, M. E., & Oxford, R. 1990. "Adult language learning styles and strategies in an intensive training setting." *Modern Language Journal, 74*(3), 311-327.

Ellis, R. 1989. "Classroom learning styles and their effect on second language acquisition: A study of two learners." *System, 17*(2), 249-262.

Ellis, R. 1992. *Second language acquisition and language pedagogy*. Clevedon, Philadelphia: Multilingual Matters.

Gardner, R. C. 2005. “Motivation and Attitudes in Second Language Learning.” In Brown, K. (Ed.), *Encyclopedia of language and linguistics* (2nd ed., Vol. 8, pp. 348-355). Oxford, England: Elsevier.

Griffiths, R., & Sheen, R. 1992. “Disembedded figures in the landscape: A reappraisal of L2 research on field dependence/independence.” *Applied Linguistics, 13*, 133-148.

Henson, K. T., & Borthwick, P. 1984. “Matching styles: A historical look.” *Theory into Practice, 23*, 3-9.

Hickcox, L. K. 2006. “Learning styles: A review of the inventories—1960s-2000s and the question of their actual uses inside and outside of the classroom.” In R. R. Sims & S. J. Sims (Eds.), *Learning styles and learning* (pp. 1-17). New York, NY: Nova Science Publishers.

Horwitz, E. K. 1995. “Student affective reactions and the teaching and learning of foreign languages.” *International Journal of Educational Research, 23*(7), 573-579.

Joy, S., & Kolb, D. A. 2009. “Are there cultural differences in learning style?” *International Journal of Intercultural Relations, 33*(1), 69-85.

Kauchak, D., & Eggen, P. 2008. *Introduction to teaching: Becoming a professional* (3rd ed.). Boston, MA: Pearson Education.

Keefe, J. W., & Ferrell, B. G. 1990. “Developing a defensible learning style paradigm.” *Educational Leadership, 48*(2), 57-61.

Kolb, D. A. 1985. *Learning style inventory.* Retrieved from http://www.haygroup.com.

Kolb, A. Y., & Kolb, D. A. 2006. ”Learning styles and learning spaces: A review of the multidisciplinary application of experiential learning theory in higher education.” In R. R. Sims & S. J. Sims (Eds.), *Learning styles and learning*. New York, NY: Nova Science Publishers, 45-91.

Lashley, C., & Barron P. E. 2006. “The learning style preferences of hospitality management students: Observations from an international and cross cultural study.” *International Journal of Hospitality Management, 25*(4), 552-569.

Lemire, D., & Gray, J. 2003. ”An introduction to learning styles for developmental educators.” *Journal of College Reading and Learning, 33*(2), 231-239.

Mammen, J. M., Fischer, D. R., Anderson, A., James, L. E., Nussbaum, M. S., Bower, R. H., & Pritts, T. A. 2007. “Learning styles vary

among general surgery residents: Analysis of 12 years of data." *Journal of Surgery Education, 64*(6), 386-389.

McCarthy, B. 1980. *The 4MAT system: Teaching to learning styles with right/left mode techniques*. Barrington, IL: Excel.

Moran, A. 1991. "What can learning styles research learn from cognitive psychology? *Educational Psychology, 11*(3/4), 239-246.

Oxford, R. L. 1990. *Language learning strategies: What every teacher should know*. Boston, MA: Heinle and Heinle.

Oxford, R. L., & Burry-Stock, J. A. 1995. "Assessing the use of language learning strategies worldwide with the ESL/EFL version of the Strategy Inventory for Language Learning (SILL)." *System, 23*(1), 1-23.

Oxford, R., & Ehrman, M. 1993. "Second language research on individual differences." *Annual Review of Applied Linguistics, 13*, 118-205.

Oxford, R., Ehrman, M., & Lavine, R. 1991. "Style wars: Teacher-student style conflicts in the language classroom." In S. S. Magnan (Ed.), *Challenges in the 1990s for college foreign language programs*. Boston, MA: Heinle and Heinle, 1-25.

Peterson, E. R., Rayner, S. G., & Armstrong, S. J. 2009. "Researching the psychology of cognitive style and learning style: Is there really a future?" *Learning and Individual Differences, 19*(4), 518-523.

Reid, J. M. (Ed.). 1995. *Learning styles in the ESL/EFL classroom*. Boston, MA: Heinle and Heinle.

Sims, R. R., & Sims, S. J. (Eds.). 1994. *The importance of learning styles: Understanding the implications for learning, course design, and education*. Westport, CT: Greenwood Press.

Sims R. R., & Sims, S. J. (Eds.). 2006. *Learning styles and learning*. New York, NY: Nova Science Publishers.

Skehan, P. 1998. *A cognitive approach to language learning*. Oxford, England: Oxford University Press.

Terrell, S. 2002. "The effect of learning style on doctoral course completion in a web-based learning environment." *Internet and Higher Education*, 5(4), 345-352.

Thomas, P. R., & McKay, J. B. 2010. "Cognitive styles and instructional design in university learning." *Learning and Individual Differences, 20*(3), 197-202.

Van Zwanenberg, N., Wilkinson, L. J., & Anderson, A. 2000. "Felder and Silverman's Index of Learning Styles and Honey and Mumford's Learning Styles Questionnaire: How do they compare and do they predict academic performance?" *Educational Psychology, 20*(3), 365-380.

Wintergerst, A. C., DeCapua, A., & Verna, M. A. 2003. Conceptualizing learning style modalities for ESL/EFL students. *System, 31*(1), 85-106.

Zywno, M. S. 2003. *A contribution to validation of score meaning for Felder-Soloman's Index of Learning Styles*. Paper presented at the American Society for Engineering Education Annual Conference & Exposition, Nashville, TN.

Appendix 1

Questionnaire to determine the perceptions on identifying one's learning style through the Kolb (1985) Learning Style Inventory

(Students)

MALE/FEMALE:
AGE:
DEGREE:
COURSE:

1. Did you like to identify your learning style through the Kolb LSI?

YES	NO

Why____

2. Was it the first time that you identified your learning style in class with the Kolb LSI?

YES	NO

3. Had you identified your learning style before with another type of test?

YES	NO

If you gave a positive answer, which test did you use?

4. Do you think that it is useful for you to know what learning style you have?

YES	NO

Why____
____
____
____

5. Do you think that it is useful for your teacher to know what learning style you have?

YES	NO

Why____
____
____
____

6. Did you agree on the learning styles results obtained through the Kolb LSI?

YES	NO

Why____
____
____
____

7. From the Kolb LSI, which characteristics seemed to be the most connected to who you believe you are as a learner?

____
____
____
____
____
____

8. Write about your experience and perceptions using the Kolb LSI?

____
____
____
____
____
____

Variation and language engineering in Yoruba-English code-switching

Emmanuel Adedayo Adedun, University of Lagos

Mojisola Shodipe, University of Lagos

Abstract
This study deals with the identification and characterization of the variable features of code-switching as used by Yoruba-English bilinguals in Lagos, Nigeria. Against the background of the domestication of English in Nigeria and the reality of language variation across individual, social and generational perspectives, we explore the unique linguistic strategies with which the Yoruba-English speaker reinvents urban speech to meet the challenges of contemporary dynamics of communication in a fast changing world.

1.0 Introduction

The concept of language variation has remained a prominent theme in sociolinguistic enquiry by virtue of its centrality to the explication of the social context of language use. Since no speech community can be said to be completely homogenous, the fact of language variation remains a glaring reality as exemplified in everyday uses of language in different societies. Firth (1951: 78) had stressed the fact that language must be as varied as the groups who use it and the multiplicity of functions to which it is applied. Similarly, Coates (1990: 24) in delineating the domain of sociolinguistics as the social context of language use, argues that "the study of language in its social context means crucially the study of linguistic variation".

Consequently, sociolinguistic studies have been largely characterized by the exploration of the systematic relationship between language and socio-cultural organization of speech communities. The basic assumption behind this is that speakers functioning as members of a particular speech community, and within the ambit of a particular culture, have internalized not only the rules of grammar but also the rules of appropriate speech usage. These rules which are broadly shared by other members of the speech community are applied daily in speech behaviour (Sankoff, 1989). To this end, Chambers (1995: 15) defines

sociolinguistics as the study of the social uses of language, encompassing a multitude of possible enquiries. These include questions about personal, stylistic, social and sociocultural patterns of language use in society. In this regard, sociolinguistics can be said to share the goals of the ethnography of communication (Saville-Troike, 1982) which takes language as a 'socially situated' cultural form. This direction gives prominence to the analysis of the code and the cognitive process of its users.

Furthermore, sociolinguistics and the ethnography of communication are united in their focus on the speech community, the systematic organization and patterning of communication within it, as well as the interaction of these communicative events with other systems of culture. The scope and focus of ethnography of communication underlie the significant contribution of these perspectives to sociolinguistic research, particularly in the description and analysis of naturalistic speech in various social contexts. Thus, while the sociolinguistic perspective of this study explores the dynamics of language variability, the ethnographic dimension deals with the socio-cultural organization of communicative events such as ways of speaking, social norms and values (Saville-Troike, 1982, 1989). Our ethnographic approach in this study will also reveal how socialization processes influence the social differentiation of language behaviour from one generation to another. In this regard, there is need to emphasize the relevance of the internal structure of the urban Lagos Island speech community to the linguistic material which emanates from it. The internal structure of the Lagos speech community includes the various modes of socialization (e.g. greetings, forms of address, dressing, and inter-group relations) as well as cultural world view as expressed at both individual and communal levels of social interaction.

Against the background of the notion of communicative competence (Hymes, 1964, 1974), naturalistic data of language alternation emanating from the Lagos Island speech community are used in this study to account for differential performance at the individual and group levels. Moreover, in view of contemporary directions in sociolinguistic studies of bilingual behavior (Auer, 1991, 1990, 1984; Sankoff 1989; Myers-Scotton 1993;), the social context of speech usage finds prominence in this study in terms of the patterns of bilingual speech production in the daily interactions of Yoruba-English bilinguals. Therefore, as a

sociolinguistic study, this paper explores the systematic interrelations between the language users, the linguistic form and the social context of speech usage. Our sociolinguistic perspective in this study thus deals with the significance of the social context as a crucial component of language alternation in this cosmopolitan speech community.

The term 'language engineering' refers to the potential of a language to express new and emerging ideas, notions or concepts. Capo (1990) defines language engineering as:

> that domain of applied linguistics concerned with the design and implementation of strategies (i.e. conscious and deliberate steps) toward the rehabilitation and optimal utilization of individual languages (1990:1).

Language engineering is therefore conceived as a conscious attempt to influence the form of a language and this implies three phenomena that are related to lexical change (Ammon, 2005: 26). These are: 1. Standardization of pronunciation, spelling, and the meaning of words. 2. Creation of new names from organizations whose acronyms create easily pronounceable words and are semantically related to the organization's aims. 3. Public use of language. However, language engineering in this article involves both the conscious and unconscious use of two bilingual languages for communicative interactions and the resultant lexical change arising from everyday communication.

Thus, languages are constantly engineered to meet the challenges of everyday communication often necessitated by changes in the social, economic or political life of a speech community. Dadzie (2004: 68) notes that every human language is subject to change and several factors responsible for this may range from the historical to the cultural and the linguistic. The English language underwent significant changes as a result of successive invasions of English territories by Saxons, Normans, Danes and the French. So great is the influence of these incursions that the English which was spoken in the 9th century bears no resemblance to the present day English. For example, Old English *seo eaxl* is a far cry from its modern equivalent *shoulder* (Dadzie, 2004).

The Nigerian situation typifies what obtains in many Anglophone West African countries where English gained ascendancy over the numerous ethnic languages as an official lingua franca. The people acquire it as a means of responding to several sociolinguistic needs which include the use of English as a medium of education and as the

language of politics, commerce and even religion. Naturally some localization must occur, since the language must reflect its new environment and portray ideas which did not exist in its original home.

The Nigerian experience is thus characterized by the development of a variety of English which has unique local colouring in the form of the infusion of ideas and concepts from the indigenous languages. Thus Nigerian English (Dadzie 2004a, 2004b; Okoro 2004; Adetugbo 1976, 1980; Adegbija 1987, 1988, 1989, 2004) has become established as a variety of "English as used by Nigerians" (Okoro, 2004). Okoro (ibid: 169) further categorizes the features of Nigerian English into four as follows:

(1) Common-core features: These are features shared with other English varieties worldwide and include syntactically and semantically neutral sentences like "Good morning', "I am going home", etc.

(2) Peculiar Nigerianisms: These include loan words such as 'agbada', 'iroko', 'garri'; coinages such as 'bride-price', 'boys-quarters', 'cash-madam', 'head-tie' and 'area-boy'; category shifts such as 'to flit a room' (to spray with insecticide), to Xerox a document' (to make a photocopy of), 'to tipp-ex an error' (to cover with correcting fluid), and meaning broadening in words such as 'customers' (referring to both buyer and seller in Nigerian English whereas in native Speaker English, it is restricted to the buyer).We must however point out here that usages like 'to tipp-ex' and 'to xerox' are also common in British and American English and are therefore not peculiar to Nigerian English usage. We can however argue that these usages represent examples of the influence of Americanisms in contemporary social interactions of the Nigerian speaker.

(3) Local Idioms: (including modifications of existing native-speaker idioms). These include the following:

– *Don't put sand in my garri.*

(don't ruin my chances)

– *You met me well/your legs are good.*

(inviting someone who has just arrived to join in a meal)

– *She used long leg to obtain the job*

(she obtained the job through undue influence and favouritism)

– One tree can't make a forest.

(BrE: One swallow cannot a summer make)

– Cut your coat according to your size

(BrE … according to your cloth).

(4) Characteristic breaches of the code (i.e. characteristic errors): Here, Okoro distinguishes between *random errors* (those that occur as part of an individual learner's interlanguage and are not necessarily shared by other users) and *characteristic errors* (those that are so regular and so widespread that they have come to be identified as part of the unique features of the language variety being described).

Examples:

Random Error

My father told me to told you to come.

This is considered random because the pattern is not widespread and the speaker is not likely to be consistent in his/her faulty double marking of tense, and can be easily corrected.

Characteristic Error

Buy your stationeries here.

This is characteristic in the sense that it displays the peculiar Nigerian English feature of pluralizing non-count nouns. Okoro observes that the numerous Nigerians who commit this kind of error have remained impervious to correction, thus such errors have become characteristic features of Nigerian English. Other examples of characteristic errors are observed in the following features:

a. redundancies e.g. night vigil, wake keeping, new innovations, funeral ceremony, can be able, secret ballot, etc.

b. omission of determiners before singular nouns e.g.

He came to the city to do ϕ assignment.

He asked me to have ϕ seat

c. stative verbs used dynamically e.g.

…we are not hearing you!

…you are still owing me two thousand naira

… who is having my book?

d. Use of redundant prepositions e.g.

He requested for our assistance

Olu contemplate on what to do.

Similar categorizations of Nigerian English features (Jowitt, 1991) would classify these examples as ‘standard forms’, ‘variants’ and ‘errors’ where ‘variants’ correspond to Okoro’s ‘Nigerianisms’ and ‘local idioms’. Both Jowitt (ibid) and Okoro (ibid) agree that the usage of every Nigerian is a mixture of standard forms and a myriad of errors and variants, otherwise referred to as Popular Nigerian English forms (PNE) and are clearly distinguishable from Standard English (SE) forms by virtue of their inherent local colouring at the lexical, syntactic and semantic levels of usage.

Similarly, many African languages have had their fair share of colonial influence occasioned by contact with European languages like French and English. In the Nigerian situation, Yoruba along with Hausa and Igbo has the status of a national language. The history of contact between English and Yoruba accounts for the process of assimilation and acculturation (Akere, 1987). This led to a 'reinventing' of the linguistic repertoire of the Yoruba-English bilingual speaker to include English loan words and assimilated forms.

Ufomata (1991) in her article 'Englishization of Yoruba Phonology' observes that the adoption of certain loanwords from English has effected a fundamental change in the phonological system of Yoruba. This includes the violation of the restriction on the occurrence of high tone on the first syllable of Yoruba vowel-initial words as in the examples:

Agent	[édʒenti]	[éjenti]
Engine	[endʒini]	énjìnì
Iron	[ájoonu]	ayóònù
Officer	[ofisa]	ófísa

In these examples, it is evident that stress in English words is converted to a corresponding set of tonal patterns when borrowed into Yoruba. Another feature of the influence of English on Yoruba in the use of loan words is the establishment of pitch and segment correspondences between the two languages. According to Ufomata (2004), in most instances, loans simply take on these correspondences while consonant clusters which are absent in Yoruba phonological system are resolved by epenthesis or deletion as in the examples: barber (*bábá),* soldier (*sójà*), half penny (pronounced /eipni/) (*éékpìnì*), street (*títì*), kettle (*kétù*), bicycle (*báísíkù)* (2004:580-581).

At the semantic level however, there are notable exceptions to the correspondences discussed above. In such cases, Yoruba tonal patterns actually keep meaning apart in homonymous English loans e.g.

‘baby’	[bèbí]	pretty young lady
	[bébì]	baby
‘cocoa’	[kókò]	cocoyam
	[kòkó]	cocoa
‘party’	[pátì]	political party
	[patí]	party, social gathering
‘father’	[fádà]	male parent
	[fadá]	reverend father
‘sister’	[sístà]	reverend sister
	[sistá]	older female

From the foregoing, it is shown that the connection between language variation and language engineering in bilingual situations is essentially a function of the socio-cultural context of language use. In this paper, we are concerned with the various ways in which the Yoruba English bilingual in the Lagos cosmopolitan setting explores the linguistic potentials of the two codes in the task of reinventing bilingual speech to accommodate contemporary nuances of everyday communication.

2.0 Background

Much has been written about the forms and functions of English and indigenous languages in Nigeria. Scholars from both literary and linguistic realms have expressed differing views about the status of indigenous Nigerian languages vis-à-vis the overwhelming influence of English (Osundare, 2004; Ufomata, 2004; Bamgbose, 2004). Against the background of this important linguistic principle of language equality, many scholars have lamented the hegemony of English which according

to Bamgbose (2004) is characterized by, among other things, increased power and prestige of English at the expense of other languages; the spread and domination of Anglo-American culture and positive attitudes and preference for English at the expense of one's own language. Bamgbose (2004: 2) further explains that all languages cannot possibly possess equal status in view of differences in language function. He then poses the question: if all languages are equal, why are some languages used in a wider range or domain? According to Bamgbose, differential values may be assigned to languages depending on a combination of sociolinguistic and economic factors**.** In order to consider these differential functions our study of Yoruba-English bilingualism explores the relative facility of the two languages to use linguistic innovation to meet the demands of contemporary usage. In this regard, Bamgbose (ibid) posits that:

> by concentrating on language structure and potentiality of language use, the linguist emphasizes language equality, while by concentrating on language function and language attitudes, the educationist and sociolinguist emphasizes language inequality. (2004:2)

Thus the situation in Nigeria (like many African countries which are former British colonies) is characterized by the retention of English as an official language and the language of post-primary education. While the indigenous languages are restricted to primary education and the cultural domain, Yoruba, alongside Igbo and Hausa, has the status of a national language as prescribed in the 1999 Constitution of the Federal Republic of Nigeria. Along with Igbo and Hausa, Yoruba is also prescribed in the National Policy on Education (1977) as a medium of instruction in primary schools and to be studied as a second language in junior secondary schools in Nigeria. The Yoruba language is the mother tongue of a substantial number of speakers in South West Nigeria and is also spoken outside Nigeria in places like Republic of Benin (where it has the status of a national language) and Togo. Moreover, the Yoruba language retains its presence in the oral literatures of Yoruba descendants now domiciled in Brazil, Cuba and parts of the West Indies, Trinidad and Tobago. In Nigeria, where a majority of speakers live, it is spoken mainly in Lagos, Ondo, Kwara, Ogun, Ekiti, Oyo and Osun states as well as in parts of Edo and Kogi states. According to the 2006 census, the population of Yoruba speakers within Nigeria was estimated at thirty

million, close to one fifth of the population of Nigeria. The same 2006 census put the population of Hausa speakers at about fifty four thousand and the speakers of Igbo at about seventeen thousand.

It is pertinent to say at this juncture that despite the entrenched functions of English and its subsequent preeminence in Nigerian polity, the Yoruba language has had an interesting history of growth and literary development of instruction at all educational levels. Yoruba has been explored in rigorous academic research at the tertiary level. Literary works abound in Yoruba in the three genres of poetry, drama and prose while all the genres of oral literature have equally been documented in the language. Furthermore, the use of Yoruba language in publishing, journalism and broadcasting has been a major boost to indigenous communication in Nigeria and beyond. Presently, there are efforts towards the development of a Yoruba language based computer system.

Like English, the Yoruba language has however had a chequered history of development as a dynamic medium of communication. Babalola (1972) had observed that one of the major problems of Yoruba is that of expressing new items or ideas introduced into the language through other languages in contact situations such as English, French, Arabic and Hausa. He suggested the use of neologisms to solve this problem as in the examples of: minute (*ìséjú*), lesson (*èkó),* glass (*ife),* matches (*ìsáná),* pencil (*léèdì),* etc.

Yusuff (2008) however observes that it is part of the natural developmental process for the speakers of a language to devise means of expressing ideas and concepts which are alien to one's culture. He further notes that apart from deliberate efforts at lexical developments for formal use, the Yoruba language has the grammatical resources to create lexical items when faced with the challenge of innovativeness. Part of the dynamism of the Yoruba language is the shift from the original Oyo dialect base into what is now regarded as standard Yoruba (SY), the variety which is being presently used as a medium of instruction in schools, in literature and broadcasting.

According to Yusuff (ibid) SY is not a regional dialect but it is needed for harmony among the regional dialects. This variety can easily be described now as Common Yoruba (CY), that is, the spoken form which is moving away from the norms of Standard Yoruba. Common Yoruba is the variety being widely spoken in cosmopolitan settings like Lagos, Ibadan, Abeokuta where users possess native speaker intuitive

knowledge of Yoruba and where linguistic challenges exist in everyday communication. Thus, Common Yoruba is widely spoken among students, artisans, teachers, broadcasters, writers, musicians, housewives, etc.

From the foregoing, it is evident that the urban variety (Common Yoruba) which operates alongside English (Standard English forms or Nigerian English variety) can be described as a variety of Yoruba which has the potential for being reinvented in contemporary usage. Moreover, the use of an urban variety of any language with its characteristic feature of marking the speaker as civilized or sophisticated and willing to embrace the dynamics of an evolving world, clearly places the Yoruba-English bilingual in Lagos as an innovative language user.

3.0 Methodology

The thrust of this study is the exploration of the linguistic resources at the disposal of the Yoruba-English bilingual and the various ways these resources are used to reinvent the two languages for the expression of contemporary dynamics of urban existence. In this regard, this study relies on a corpus of naturally occurring bilingual speech of Lagos city dwellers in different social settings. The speakers cut across all strata of society but are all united by the common trait of being Yoruba L1 and English L2 speakers with common history of acquisition of English at an early age. However, levels of proficiency in both languages vary based on level of education and length of exposure to both languages. Using the spontaneous interviews and the non-participant observation method, the researcher isolates forms (lexical, morphological or semantic) which exhibit the features of language engineering in the use of both Yoruba and English.

It has however been argued (Wolfson, 1976) that in the absolute sense, there is no such thing as natural speech such as the type obtained from tightly controlled structured interviews and the spontaneous interviews. Wolfson's argument is that the respondent is usually constrained to operate within the performance context of the question-answer pattern of the interview format. Therefore, the respondent's responses in terms of content and style of delivery are considerably influenced by the cues given by the interviewer.

Some scholars have also observed that Wolfson's position appears to take little cognizance of the great potency of the natural setting as a determinant of language or variety choice. In this kind of situation, emotions such as joy, fear, tension, anger or anxiety (which are normally exhibited in speech acts such as exclaiming and interjecting) are usually demonstrated in the more naturally occurring of the two codes or varieties in a speaker's repertoire. This kind of language use can be said to underlie not only patterns of codeswitching in bilingual speech, but also the choice of the elements used by the speaker.

4.0 Conceptual Consideration

The description of language, whether spoken or written, is the primary business of linguistics. Similarly, studies of language use in human societies belong in the realm of sociolinguistics. This paper explores the relationship between language variation and language engineering within the framework of bilingual behaviour in a non-native English environment. Our exploration of these areas also embraces the issue of code-switching and language change since bilingual behaviour, code-switching and borrowing represent some of the crucial features of the sociolinguistic influences which promote language change worldwide.

By the term 'code-switching', we refer to any kind of discourse in which words originating in two different language systems are used side-by-side. Gumperz (1972, 1982) had described the phenomenon of codeswitching as 'the juxtaposition within the same speech exchange of passages of speech belonging to two grammatical systems or subsystems'. In the same vein, Li Wei (2003) describes bilingual code-switching as the alternation of languages in the same interactional episode. Gafaranga and Torras (2002) however argue that the concept of code-switching has traditionally been understood to mean any occurrence of two languages within the same conversation. They further identify code-switching as a specific type of language alternation which they describe as '*interactional otherness*', a term which refers to a kind of internally motivated deviance (that is, serving a specific interactional function) (Gafaranga and Torras 2002).

In contemporary sociolinguistic research however, perhaps the most dominant feature of the literature on the subject is the diversity of opinions about what constitutes an adequate definition of the

phenomenon. While some linguists (cf. Myers-Scotton, 1993) consider the term 'code switching as being synonymous with 'language switching', others like Romaine (1989) insist on using the term in the same sense as Gumperz (ibid) initially used it. Poplack (1980) however defines code-switching not only in relation to discourse but as also being inclusive of the phenomenon of code-mixing. According to Poplack, "code-switching refers to the mixing, by bilinguals, (or monolinguals) of two languages in discourse, often with no change of interlocutor or topic". While asserting that such mixing may take place at any level of linguistic structure, Poplack notes that considerable linguistic attention has however been focused on the occurrence of code-switching within the confines of a single sentence, constituent or even word. Similarly, other sociolinguistic scholars have defined code-switching as the alternative use of two languages at the word, phrase, clause or sentence level and have also called for a distinction among the different language phenomena. For instance, Auer (1984, 1991, 1998) distinguishes code-alternation from what he refers to as the 'new code' and in turn distinguishes code-switching from transfer within the ambit of code alternation.

Among scholars who use the term code-switching to describe any instance of language alternation, there has been the need to establish its different types based on the dynamics of its usage. To this end, Gumperz (1982) identifies situational, metaphorical and conversational code-switching while Myers-Scotton (1993) distinguishes between marked and unmarked code-switching. From a discourse perspective, Romaine (1989) also identifies tag-switching, inter-sentential code-switching and intra-sentential code-switching and states that all three types of code-switching may be found within one and the same discourse.

These various perspectives underscore the fact that the diverse postulations of scholars on the appropriate definition of code-switching should be viewed beyond the realm of differing terminologies. These different views and categorizations of code-switching are actually based on different theoretical orientations and on different views about the notion of language and that of the code in social interactions. Gafaranga and Torras (2002) assert that these viewpoints reflect different epistemological orientations. To this end, they describe the views of Gumperz and Myers-Scotton as 'identity-related explanation' since they are interested in the social values of language and the social motivations

for code-switching, respectively. Auer's view on the other hand represents what is described as a sequential perspective, a viewpoint which is based on Auer's conversation analysis approach to the subject.

At this point, it is worth noting that linguists have observed some note-worthy situations in the social interactions of bilinguals worldwide. It is quite possible for instance, to find that social interaction among bilinguals may not always be conducted in two languages. In fact, many researchers have observed that talk among bilinguals may be conducted in one language only. This has been variously described as *preference for same* l*anguage talk* (Auer, 1984) and as the preference *for same medium talk* (Gafaranga and Torras, 2002). When this occurs, the interaction is usually not considered worthy of notable accounting. Rather, it is seen as the norm among both monolinguals and bilinguals. This has been referred to as the *monolingual bias* (Gafaranga and Torras, 2002) which has been accounted for in terms of the *monolingual medium*.

The second situation which has been observed involves cases where bilinguals use both of their languages without any obvious motivation (Labov, 1972). This differs from monolingual language use in the sense that it is usually noticed by both researchers and community. Linguists commonly refer to this practice as 'a new code' (Auer, 1984), as 'codemixing' and more technically as 'unmarked codeswitching' (Myers-Scotton, 1993). This kind of talk has also been recognised by members of speech communities as a 'stable' practice (Garfinkel, 1972) and thus speakers have been known to assign labels like Spanglish, Chinglish, Franglais and Yorubanglish to such speech patterns as a recognition of the specific nature of these different types of language alternation. This takes us back to the issue of the monolingual bias as the underlying factor for speakers' recognition of language alternation. The phenomenon of language 'switching' or 'mixing' is noticed because it is generally assumed that the norm is to speak in only one language. In this regard, Auer notes:

> ….In many bilingual communities, there is a preference for same language talk, codeswitching runs counter to this preference which of course only heightens its signalling value … (1991: 28-29)

Our exploration of bilingual behavior in this paper also embraces the issue of code-switching and language change. It has been observed that bilingual behaviour, code-switching and borrowing represent some of the

crucial features of the sociolinguistic influences which promote language change worldwide. Backus (1996: 27) defines ‘language change’ as ‘contact-induced structural change’. In other words, changes in the structure of a language as a result of language contact. One of the major structural changes in contact situations can be observed in word order. Thus it operates on the hypothesis that word order change in bilingual situations is caused by frequent code-switching (Backus 2005) as in the example:

A.	I	saw	that	tall	man	yesterday
	NP	verb	det	Adj	noun	Adv

Yoruba-English code-switching:

	Mo	rí	man	gíga	yen	lánàá
	(I	saw	man	tall	that	yesterday)
	NP	verb	noun	Adj	det	Adv

B.	He	wore	a	red	outfit
	NP	verb	det	Adj	noun

Yoruba-English Code-switching:

	O	wo	aso	red
	NP	verb	noun	Adj
	He	wore	outfit	red

C.	We	saw	thirty	people	there
	NP	verb	det	noun	Adj.

Yoruba-English Code-switching

	A	rí	èèyàn	thirty	níbè
	NP	verb	det	noun	Adj
	We	saw	people	thirty	there

However, some linguists believe that formal changes in a language can be stated at various levels of abstraction. Backus (ibid) thus argues that it is rather imprecise to say that “word order” has changed in a given language. We do not know for instance whether it is “basic word order” (cf. Whaley 1997) that has changed or word order in a specific construction, such as topicalization or the change from declarative to interrogative. Therefore, before accepting a form as ‘new’, it is necessary to show that it was not part of the language all along. According to

Backus, the clearest evidence of this of course would be the absence of the feature in the pre-contact variety, but this type of evidence is not always easy to obtain in syntax. Moreover, Nettle and Romaine (2000) observe that many languages which are currently heavily dominated by another language (such as Yoruba) do not have monolingual speakers anymore who can act as a yardstick, while pre-contact records have often disappeared or are non existent.

Many scholars have noted that all theories of language change are skeptical about what can be achieved by way of predicting the future of languages (Croft 2000, Weinreich 1953). To this end, Matras (2000) poses the question: "how predictable is contact-induced change in grammar? Matras' answer is that one can predict for certain subsystems of grammar what course change will take, or at least make 'an intelligent guess'. Johanson (2002) however argues that the idea of looking for causes of language change has its pitfalls chiefly because causation of language contact takes place at various levels and many factors interact. Thus, while some description of change address structural factors (e.g. gaps in the system), others address the type or mechanisms of the change (e.g. reanalysis), yet others the social context in which the change arose (e.g. language contact or desire for upward mobility).

Backus (ibid) further observes that a relevant consideration in this regard is the distinction between ultimate and proximate causes. Ultimate causes of language change are likely to be socio-cultural (Kontra, 2001; Thomason and Kaufman, 1988) while proximate causes include cognitive, attitudinal, motivational and probably, purely structural factors.

Similarly, Thomason and Kaufman (1988) make a distinction between 'predictors' (social and linguistic factors that drive change), 'mechanisms' (the processes through which change is effected) and 'results' (the visible effects on the changing language. Backus (ibid) observes however that the distinction between predictors and mechanisms though useful, is not always easy to make.

In view of all the factors involved, it is evident why it is difficult to arrive at a generally accepted theory. In addition to this there is the issue of whether contact induced change should be separated from a general theory of change, mainly because bilingualism is one of the more important social factors promoting change and internal and external factors often conspire in this type of change.

Furthermore, Backus (ibid) notes that in addition to direct importation (borrowing), there is also indirect change induced by the circumstances of contact, but this does not involve actual borrowing. Indirect changes are especially likely to involve a combination of internal and external causes (Romaine 1989; Thomason 2001).

5.0 Features of language engineering in Yoruba English code-switching

5.1 Lexical

The Yoruba-English bilingual situation in cosmopolitan Lagos represents the kind of speech where speakers alternatively use Yoruba and English words or phrases for a variety of reasons. This kind of linguistic ability may not incessantly be available to speakers with less proficiency in both languages. At the lexical level, the Yoruba English bilinguals adapt their urban speech for the achievement of the following:

1. **Personal identity**

   Personal or individual identity implies identification of the self within the large space of the society, at specific period and within groups to which people belong. This is most exemplified in the personal naming system whereby these bilinguals reinvent themselves through the prevalence of code-mixed personal names, a trend which cuts across age and gender divides. For instance, people are generally identified by the use of a Yoruba first name followed by an English appellation which indicates professional or occupational identification. Thus the common sociolinguistic pattern of personal naming in this community values names like: *Ibrahim Casket, Taiye Vegetable, Sule Mandilas, Wale Teacher, Musibau Escort, Shamsideen Developer, Azeez Councillor, Musiliu Coach, Tunde Parking, Nurudeen Sergeant, Fausat Cellular, Kunle Entertainment.*

Apart from indicating an acknowledgement of the prestige norms associated with English as an international language, this naming pattern

captures the ways of life of speakers in terms of their specific preoccupations. Thus, the code-mixed names capture the essence of contemporary usage where speakers attempt to attach some level of prestige or sophistication for themselves through personal naming.

However, some code-mixed names also serve different functions for the users. Certain members of the community come to have certain names or nomenclatures attached to them by others by virtue of their individual personal traits, behaviour or idiosyncracies, whether positive or negative. Such distinctive personal traits are usually embedded in the English component of these bilingual names. They include:

Gbolahan Computer: a smart person, a cunning smooth operator who always seems to have everything figured out (like a computer).

Tunde Banana: a slippery, deceitful person. (slippery as a banana peel)

Mustapha Emotion: a gentle, lovable and genial person.

Bisi Dollar: a high flying society lady who travels abroad frequently for business and is well-known for her penchant for spending foreign currency. (e.g. dollars)

Fatai Always: a man who is ever so resourceful and willing to explore any means to make money.

Names such as these operate on the principle of social acceptance or rejection, approval or disapproval as the case may be. But implicit in all these is the fact that the naming system provides insights into individual character in relation to the social norms of a specific period and in some cases, neighbourhood dynamics. The practice of identifying people by catchy, trendy or symbolic names also includes the prevalence of nicknames. These include:

a. Relexicalised versions of personal names such as: *Rosco* for Rasaq, *Owoblow* for Owolabi, *China* for Shina and *Murphy* for Mufutau.

b. Initializations such as: *I.D.* for Idowu, *R.S.K* for Rasaq, *S.K* or *Eskay* for Sikiru, *B.G* for Bode George, *T.J.* or *Tee Jay* for Tajudeen.

c. Acronymisations e.g. Samuel Adewale Maja (SAM), Ganiyu Olawale Solomon (GOS), Femi Akintunde Johnson (FAJ), Bushura Alebiosu (BUSH) etc.

d. Clipping: This involves using the shortened forms of personal names as in the examples of FASH for Fashola, KUSH for Kushimo, LAI for Olayiwola and TOKS for Tokunbo.

Generally, these coinages serve many communicative functions which include the quest for social identification, group solidarity or the need to disguise one's identity. They represent the various ways in which individuals and social identities are mediated by bilingual behaviour.

2. Group Identity

Language features and the various identities they portray generally imply a boundary function. The peculiar socio-linguistic structure of the Lagos speech community means that individuals find their relevance in group identity. This means that people perceive themselves and are perceived by others as more relevant individually when they are also recognized as part of a group. Therefore everyone strives to be defined within a group structure. The bilingual communal norm is thus also reflected in the naming of social groups to indicate sophistication and a willingness to follow the trend of contemporary usage. Thus, we have the preference for Yoruba-English group names against monolingual Yoruba or monolingual English names as in the examples: *Ricca Gents, Inabiri Ladies, Waka Club, The Great Offin Gents, Balogun Yuppies, Fila Connection, New Generationext of Olowogbowo, Upper Olowogbowo Gents, etc.* Some group names are however retained in English but with phonological and/or orthographic modifications as in the examples: *Sunday Skool* for Sunday School, *Faya One* for Fire I.

5.2 Semantic

At the semantic level of usage, the Yoruba-English bilinguals of Urban Lagos attempt to reengineer their speech with the following strategies:

5.21. Devernacularisation

This involves the use of lexical items in code-switching discourses in forms which deviate from their original meaning and usage in the indigenous Yoruba language. Devernacularisation in this case is a strong indicator of generational variations in bilingual behaviour whereby the members of the younger generation exhibit a preference for inventive, creative usages while the older generation prefers more conservative usages. Thus, devernacularisation represents a feature of outer marking (Mdiase Tham, 1990) or in-group marker for youths. It takes the form of a restricted code or in-group slang which is used to exclude others not considered members of the group. For example, Yoruba words like "isu" (yam) and 'eran' (meat) are devernacularised when used as pseudonyms for money in expressions like:

(i) Awon *boys* yen se isu *seriously* ni *council.*

literal translation:
(Those boys make yam seriously in the council)

Meaning:
Those boys are making a lot of money in the council.

(ii) A ni ki *leader* ya eran fun awon *party members*.

literal translation:
(We asked the leader to cut some meat for party members)

Meaning:
We asked the leader to give money to the party members.

Such expressions however retain their original meaning and usage in the speech of the older generation of Yoruba-English bilinguals—e.g.

Awon *boys* yen ri owo l'eni.
(Those boys made money today)

A ni ki *leader* gbe owo fun *party members*.
(We asked the leader to give money to the party members).

Other examples of devernacularised usages in the bilingual speech of youths include the following:

Example 1:

Ojà[1]: (Youth: hard drugs)

Ó gbé ojà lo si London ni wón bá mu ni airport.

(He/she travelled to London with hard drugs and was arrested at the airport).

Ojà[1]: (adults: merchandise /goods)

Ó gbé ojà lo si London ni wón bá mu ni airport.

(He/she took some merchandise/goods to London and was arrested at the airport)

Ojà[2]: youths: (illicit business)

a. Sé ojà wà l'énìí?

Translation: Is there any (illicit) business today?

b. A ń lo sí ojà.

Translation: We are about to engage in some illicit business.

Oja[2]: Adults (market place)

a. Sé ojà wà l'énìí?

Translation: Is the market open today?

b. A ń lo sí ojà

Translation: We are going to the market.

Example 2:

Èjìré: Youths (police officers)

e.g. Awon èjìré ti wá *arrest* è.

Meaning: He/she was arrested by the police.

Ejire: Adults (twins)

Awon èjìré ti wá *arrest* è.

Meaning: He/she was arrested by the twins.

Example 3:

Iná: youths (chaos, trouble)

A ma sí iná fún àwon *people* yen.

Meaning: We shall open fire on those people.

Iná: Adults (fire/fireplace)

A ma sí iná fún àwon *people* yen.

Meaning: We shall switch on the light (electricity) for those people.

5.2.2. Relexification

This occurs when youths jettison original usages and replace them with new expressions in relexified form. This represents a form of divergence from adult speech which is usually a reflection of community norms. Like devernacularisation, the use of relexification in the bilingual speech of younger people serves as an in-group marker and is motivated by factors like secrecy, group solidarity, self protection and peer influence. It occurs in either of the two codes. e.g.

Yoruba	**English**	**Relexified form**	
i. àdá	cutlass	pa'na:	'put out fire'

e.g. Won kan ni *slight disagreement*, l'o ba yo pa'na.

They just had a slight disagreement and he wielded a cutlass!

ii. fóònù :	phone	aago:	bell

e.g. O lè gbá mi l'áago l'ójó *Monday.*

You can give me a ring (call) on Monday.

iii. bàtà: shoes: ìtìlè “stepping tool”

e.g. Itilè esè rè yen *serious!*

The shoes he/she wore were awesome!

iv. ìmúra: dressing: ìleéfòò “floating”

e.g. Won *arrive* pèlú ilefòó àgbàlagbà

Meaning: He arrived the venue with power dressing.

5.2.3. Nativisation

The speech of many educated young bilinguals often exhibit the use of certain code-mixed expressions which have the form of nativised English usages. The original English expressions usually have some of their elements translated to Yoruba to create an entire new code.

e.g.	**English**	**Yoruba**
	Have a nice day/weekend	(e) ni nice day/weekend
	Are you alright/ok?	Se o wa alright/ok?

Some of these nativised expressions have a surface structure of Yoruba and an English deep structure having being derived from core English fixed expressions or idioms. Otherwise known as slang, these usages are common features of the speech of the younger generation e.g.

	Yoruba	**English L1 Source**
a.	*Disclosure:*	
	Slang: Má síi/má be	To open a can of worms.
	(do not disclose)	
b.	*Escape*:	
	Slang: O já (he ran away)	To break loose.
c.	*Brashness:*	
	Slang: O ń bé (he is too brash)	To jump the gun.

As in relexification and devernacularisation, these usages retain their vernacular meanings in the speech of older literal translation of Yoruba-English speakers as in the examples:

Má si/ Má be: Do not open.

ó já: It broke loose.

Ó ń bé: He /She jumped/ took a leap forward.

5.3. Morphological

This involves the adaptation of English loan words to the morphology of the borrowing language (i.e. Yoruba) e.g:

(i) Yoruba *lai* (meaning negative) + English verb as in:

A ko ni lo *lai* discuss oro naa.

We shall not leave without discussing the matter.

(ii) Yoruba *ol (oni)* (owner of) + English noun.

e.g. E lo discuss pelu oloja (ol(oni) + oja). Go and discuss with the owner of the business.

Conversely, we also observe some cases of the adaptation of Yoruba words and expressions to the morphology of English whereby an English morphological feature is affixed to a Yoruba word to form a new Yoruba-English expression. Some speakers have described such usage as 'Yorubanglish' (Garfinkel, 1972) as in the examples:

a. Báwo ni gbogbo *pre-Ileya* arrangements lódoò yín?

(How are you handling all the *pre*-Ileya arrangements at your end?) (Ileya is the name of a muslim festival usually held annually to coincide with muslim pilgrimage to Mecca).

b. Bayo sick díè. Gbogbo *post-igbéyàwó* fatigue yen ló ma fàá.

(Bayo has been a little ill. I assume it is a symptom of the post-wedding fatigue) (*Igbéyàwó* = marriage /wedding)

c. Daddy mi wa rather *Ijebu*-ish tó bá di òrò owó.

(My dad is rather stingy when it comes to money matters)

(Ijebu is the name of a Yoruba tribe in South West Nigeria. The Ijebu people are widely believed to be either stingy or prudent with money, hence the evolution of the slang word *Ijebu* as a metaphorical reference to a stingy/prudent person)

d. Bi àwon olè se dá mótò rè dúró, ó rora bó sílè *jéjé*-ly, ó surrender key fún won.

(As the robbers stopped his car, he stepped out of the car meekly and surrendered his car keys to them).

Most of these expressions are often confined to slang usage, in-group registers or relexicalisations for the purpose of solidarity or to show emotive content of bilingual speech.

5.4. Phonological

Some lexical features of Yoruba-English code-switching also exhibit considerable linguistic integration at the phonological level. Established loan words are marked by phonological integration mostly in the speech of the educated and non-educated speakers. This is essentially a function of speakers' bilingual ability or the peculiar sociolinguistic situation e.g.

Words	**Educated**	**Non-educated**
School	school	skúùlù
Ball	ball	bóòlù
Teacher	teacher	tísà
Station	station	tésòn
Shovel	shovel	shóbìrì
Railway	railway	reluweéè

This is normally characterized by the speaker's superimposition of indigenous pronunciation patterns on the phonic representation of English loan words.

5.5 Error Features in Yoruba-English Engineering

It is worth noting that most of the examples cited above represent bilingual language use of educated speakers. Speakers of English in Nigeria have been classified into three groups on the basis of their linguistic competence. Speakers are classified along a language continuum as acrolectal, mesolectal and basilectal respectively. This categorization is important for the assessment of the linguistic content of bilinguals' social interaction in terms of the notion of communicative competence highlighted earlier. While many Yoruba-English bilingual of urban Lagos fall into the category identified as acrolectal Nigerian speakers, some can be classified as belonging to the mesolectal group. Acrolectal speakers are those with tertiary education while mesolectal speakers generally possess school leaving certificate. The group identified as basilectal are at the extreme of the continuum and are characterized by relatively low educational attainment such as primary school education or in the extreme cases, no education at all. Yet everyone acquires the second language, English, and use it communicatively as part of the general bilingual speech norm of the community. Differential levels of usage can however be established when the notion of communicative competence is applied.

It has been observed that different categories of speakers draw from the community speech repertoire based on their linguistic competence. As a result, aspects of the naturalistic speech of our respondents exhibited varying features of basilectal speech. Using Okoro's (2004) typology, Some of these can be categorized as follows:

a. Characteristic errors:

- A n lo si *night vigil* lola.

  (we are going to a night vigil tomorrow)

- O ye kí á ti máa rí orísirísi n*ew innovations* ni government alágbádá yìií.

(We ought to have witnessed many new innovations by this democratic government by now).

- A insist pe ki awon soja *return back* si barracks won. Democracy la fe!

  (We insist that soldiers return back to their barracks. We prefer democracy!).

- Njé o lè borrow mi ní some cash? Màá fún e lóla.

  (Can you borrow me some cash? I will pay back *tomorrow).*

- *If you late again*, oò nísisé níbí mó.

  (If you come late again, you shall be relieved of your job)

- Landlord ti da *properties* wa jade. A need accommodation badly bayii.

  (The landlord threw out our properties. We desperately need accommodation now)

b. Random Errors:

This is exhibited in varying ways. They include wrong use of words as in the example:

- Eelo ni e lè avoid (afford) láti san fún aso yen? (How much can you afford to pay for the dress?)
- Mo try best mi lati convince (persuade) obinrin naa lati pada wa. (I tried my best to persuade the woman to come back)
- Nígbàtí àwon olópàá dé, anybody (everybody) bé dànù ni! (When the police arrived, everybody disappeared!)

6. Conclusion

In this paper, we have examined some of the salient sociolinguistic features of language engineering by Yoruba-English bilinguals at the lexical, semantic and morphological and phonological levels. Essentially, we have demonstrated that both English and Yoruba exert considerable influence over one another and that the various strategies for language engineering are supported by factors in the speech usage. Moreover, our

examples have shown clearly that Yoruba English bilingual engineering finds a place in the explication of the unique sociolinguistic components of language change in contemporary communication.

References

Adegbija, E. 1987. "Speech Act Functions: Dimensions of Communication and Meaning in the English Language in Nigeria". *ITL Review of Applied Linguistics* vol. 76. pp. 43-62.

Adegbija, E. 1988. "My Friend, Where Is Anini?: Decoding the Meaning of Utterances". *Journal of Pragmatics*. vol. 12. pp. 151-160.

Adegbija, E. 1989. "Lexico-semantic Variation in Nigerian English". *World Englishes* vol. 8, No. 2. pp. 165-177.

Adegbija, E. 2004. "The domestication of English in Nigeria". Babalola, E.A. and Awonusi, V.O. (eds.), *The Domestication of English in Nigeria: A festschrift for Abiodun Adetugbo at 65*. Lagos: University of Lagos Press. pp. 20-44.

Adetugbo, A. 1976. "Nigerian English: Fact or Fiction?" *Lagos Notes and Records.* vol. 6, Nos. 1 & 2. pp. 128-141.

Adetugbo, A. 1980. "Appropriateness and Nigerian English". Proceedings of 9th NESA conference and WAJE vol. xxi, No. 3. pp. 73-91.

Akere, F. 1987. "Sociolinguistic Assimilation in Lagos". Adefuye, O; Agiri, D. and Osuntokun, A. (eds.), *History of the Peoples of Lagos State*. Lagos: Lantern Books. pp. 160-186.

Ammon, U. 2005. *Sociolinguistics: An International Handbook*. London: de Gruyter.

Auer, J.C.P. 1984. *Bilingual conversation.* Amsterdam: John Benjamins

Auer, J.C.P. 1998. *Code-Switching in Conversation: Language, Interaction and Identity*. London: Routledge.

Auer, J.C.P. 1991. "Bilingualism in/as Social Action". Papers for the symposium on Code-switching in *Bilingual Studies: Theory, Significance and Perspectives,* vol. 2. Strasbourg: The European Science Foundation.

Babalola, A. 1972. "The Growth of the Yoruba Language: Problems and Solutions". Staff seminar paper and subsequent discussion. Lagos: School of African and Asian Studies, University of Lagos.

Backus, A. 1996. *Two in One: Bilingual Speech of Turkish Immigrants in the Netherlands*. Tilburg: Tilburg University Press.

Bamgbose, A. 2004a. "Negotiating English Through Yoruba: Implications for Standard Nigerian English". Owolabi, K. and Dasylva, A. (eds.), *Forms and Functions of English and Indigenous Languages in Nigeria.* pp. 612-630.

Bamgbose, A. 2004b. "English and Inequality: An African Perspective". Babalola, E.A. and V.O. Awonusi (eds.), *The Domestication of English in Nigeria: A Festschrift for Abiodun Adetugbo at 65.* Lagos: University of Lagos Press, pp. 1-14.

Capo, H.B.C. 1990. "Comparative Linguistics and Language Engineering in Africa". Emenanjo, E.N. (ed.), *Multilingualism, Minority Languages and Language Policy in Nigeria*. Agbor: Central Books Ltd, pp. 64-78.

Chambers, J.K. 1995. *Sociolinguistic Theory: Linguistic Variation and Its Social Significance*. Oxford: Blackwell.

Coates, J. 1986. *Women, Men and Language*. London: Longman.

Coates, J. 1990. "Introduction". Coates, J. and Cameron, D. (eds.) *Women in Their Speech Communities: New Perspectives on Language and Sex*. New York: Longman.

Croft, W. 2000. *Explaining Language Change: An Evolutionary Approach*. Harlow: Pearson Education.

Dadzie, A.B. 2004a. "The Concept of Nigerian English". Dadzie, A.B.K. and Awonusi, V.O. (eds.) *Nigerian English: Influences and Characteristics*. Lagos: Concept Publications. pp. 85-99.

Dadzie, A.B.K and Awonusi, V.O. (eds.). 2004b. *Nigerian English: Influences and Characteristics*. Lagos: Concept Publications.

Federal Ministry of Education. 1977. *National Policy on Education*. Lagos: Federal Government Press.

Federal Republic of Nigeria. 1999. *Constitution of the Federal Republic of Nigeria*. Lagos: Federal Government Press.

Federal Republic of Nigeria. 2006. National Population Census, Abuja.

Firth, J.A. 1951. *Papers in Linguistics (1934-55).* London: Oxford University Press.

Gafaranga, J. & Toras, M.C. 2002. "Interactional Otherness: Towards a Redefinition of Code-Switching". *The International Journal of Bilingualism*, 6 (1), 1-22.

Garfinkel, H. 1972. "Remarks on Ethnomethodology". Gumperz, J. J. and Hymes, D. (eds.), *Directions in Sociolinguistics*. New York: Holt, Rinehart and Winston, pp. 301-24. *International Journal of the Sociology of Language,* volume 1989, Issue 78. pp. 43-56.

Gumperz, John, J. 1982. *Discourse* S*trategies*. Cambridge University Press

Gumperz, John. 1972. "The Communicative Competence of Bilinguals: Some Hypotheses and Suggestions". *Language in Society* vol. 1, issue 1. pp. 143-154.

Hymes. 1964. "Introduction: Toward Ethnography of Communication". *Amenial Anthropology,* vol. 6. pp. 12-25.

Johanson, L. 2002. *Structural Factors in Turkish Language Contact*. London: Curzon.

Jowitt, D. 1991. *Nigerian English Usage: An Introduction*. Lagos: Longman.

Kontra, M. 2001. "Hungarian Verbal Puzzles and the Intensity of Language Contact". *Journal of Sociolinguistics,* vol. 5. pp. 163-179.

Labov, William. 1972. *Sociolinguistic Patterns,* Philadelphia, PA, University of Pennsylvania Press.

Li Wei. 2003. "Mother Tongue Maintenance in a Chinese Community School in New Castle upon Tyne". *Language and Education,* 7. pp. 199-215.

Matras, Y. 2000. "How Predictable Is Contact Induced Change in Grammar?" In C. Renfew, A. Mcmahon and L. Trask (eds.), *Time Depth in Historical Linguistics*. Comlnoge: McDonald Institute. pp. 563-583.

Myers-Scotton, Carol. 1993. "Building the Frame in Codeswitching: Evidence from Africa". In *Topics in African Linguistics*, Mufwene, Salikoko S. and Lioba Moshi (eds.).

Nettle, D & Romaine, S. 2000. "Vanishing Voices". *The Extinction of the Worlds Languages*. Oxford University Press.

Okoro, O. 2004. "Codifying Nigerian English: Some Practical Problems and Labelling". Babalola, E.A. and Awonusi, V.O. (eds.), *The* D*omestication of English in Nigeria:* A *Festschrift for Abiodun Adetugbo at 65.* Lagos: University of Lagos Press. pp. 166-181.

Osundare, N. 2004. "What is the Nationality of Your Idioms?: African Writers and the Language Question". Owolabi, K. and Dasylva, A. (eds.), *Forms and Functions of English and Indigenous Languages in*

Nigeria (A Festschrift in Honour of Ayo Banjo). Ibadan: Group Publishers. pp.56-79.

Poplack, N. 1980. "Bilingualism Competence: his interferences or practical integrity". In L. Elias-Olivares (ed.) *Spanish in the USA: Beyond the Southeast*. Arlingn, Urjune: National Clarifying House for Bilingual Education. pp. 107-31.

Romaine, S. 1989. "Pidgins, Creole Immigrants and Dying Languages", Dorian, M. (ed.) *Investigations of Obsolescence*: *Studies in Language Attrition and Death*. Cambridge: Cambridge University Press. pp. 369-383.

Sankoff, D. (ed.). 1978. *Long Variation: Models and Methods*. New York: Aindeemne Press.

Sankoff, G. 1989 "A Quantitative Paradigm for the Study of Competence". Bauman, R. and Sherzer, J. (eds.), *Explorations in the Ethnography of Speaking. (*2nd edn.). Cambridge: Cambridge University Press. pp. 18-49.

Saville-Troike, M. 1982. *The Ethnography of Communication.* Oxford: Blackwell.

Saville-Troike, Mariel. 1989. *The Ethnography of Communication Publishing*.

Thiam, N. 1990. "L'évolution du wolof véhiculaire en milieu urban Sénégalaise: le contexte dakarois". *Plurilinguismes*, vol. 2. pp. 10-37.

Thomason, S.G. 2001. *Language Contact: An Introduction*. Washington: Georgetown University.

Thomason, S.G. and Kaufman, T. 1988. *Language Contact, Creolization and Genetic Linguistics*. Berkeley: University of California Press.

Ufomata, T. 1991. "Englishization of Yoruba Phonology". *World Englishes,* vol. 10 No. 1, pp. 33-51.

Ufomata, T. 2004. "Tone and Stress in Contact: the example of English loanwords in Yoruba". Owolabi, K. and Dasylva A. (eds.) *Forms and Functions of English and Indigenous Languages in Nigeria, pp. 594-611.*

Weinreich, U. 1953. *Languages in Contact: Findings and Problems*. The Hague: Mouton & Co.

Whaley, L.L. 1997. *Introduction to Typology: The Unity and Diversity of Language.* Thousand Oaks: Sage.

Wolfson, Nessa. 1976. "Speech Events and Natural Speech: Some Implications for Sociolinguistic Methodology" *Language in Society*, 5:189-209 Cambridge University Press.

Yusuff, L.A. 2008. "Lexical Morphology in Yoruba Language Engineering", Ph.D. Dissertation, University of Lagos, Nigeria.

Rosewater, *wheel of fortune*: Compounding and lexicalisation in seventeenth-century scientific texts

Begoña Crespo, University of A Coruña

Abstract
This paper investigates the question of compounding as a productive word-formation process in Scientific English by exploring the concepts of collocation and lexicalisation. The claim is that compounds can exhibit different internal structures, including syntactically ambiguous forms, as is the case with the noun + prepositional phrase. Frequency of co-occurrence and the unique meaning of all elements, together with the phenomenon of technicalisation, argue in favour of such an assumption. Some constraints, however, must be admitted. On occasions, the semantic type of the head noun (abstract, concrete, proper, common) can determine whether a particular construction is to be classified as a compound or not.

1. Introduction

Modern science in the seventeenth century proposed a new type of learning based on mathematical logic, systematic experimentation, and mechanical models. From a linguistic point of view, this new method made any type of scholastic argumentation redundant, demanding instead of dialectic resources a new lexicon for modern technology, a set of accurate vocabulary items, plus structures which served to describe with precision any type of discovery or scientific development (Hard and Jamison: 25).

Science is a sort of micro-cosmos within the Universe of knowledge. Hence, the language of science, though subject to the general contexts and characteristics of a particular language, has its own distinctive features.

The evolution of science and technology fostered the creation of new words, and sometimes even led to the introduction of new morphological patterns into the language, such as the creation of neoclassical compounds under the influence of Latin and Greek (Beal: 13-34). The scientific community, represented by the members of the Royal Society, began a debate over the formal characteristics of scientific discourse which bore some similarities to the "Inkhorn Controversy" of the sixteenth century, since in both cases the discussion revolved around how to improve the lexical capacity of English. Some eighteenth century

scholars used “words or morphemes from the classical languages as building blocks in scientific terminology” (Beal, 14). Others, however, resorted to compounds.[1]

An expansion in the reading public in the late seventeenth and early eighteenth century, encompassing not only professional groups and the rich but also the middle classes, favoured the use of the vernacular to transmit “science”, even though the style employed by writers varied considerably. In some instances it was simple and clear, as demanded by the Royal Society;[2] in others, it was more literary, complex and dense.[3] These two broad styles coexisted in the construction of scientific discourse during the eighteenth century, when a handful of scientific and scholarly societies were created as forums for discussion and co-operative investigation. One of the ways in which scientific English was modified was the creation of new terms. Among these formations, compounds occupied an important position.

In order to contribute to the description of scientific language at the beginning of modern science, I will explore the differences between collocations and these new compound nouns through different degrees of lexicalisation. To this end, the paper will be organised as follows: in Section 1 I will discuss the concepts of collocation and compounding and the process of lexicalisation. Section 2 will present the corpus used in this study. Section 3 will concentrate on the analysis of data, focusing on types rather than on tokens, though some word counts will be also provided. Results will be presented according to the variables of etymology and discipline. Section 4 will deal with some unclear cases, and, finally, in Section 5 I will try to provide some conclusions.

2. From collocations to compounding. The concepts

Crystal (1997: 69) defines collocation as “the habitual co-occurrence of individual LEXICAL ITEMS. […] collocations are then, a

---

[1] As a matter of fact, sixteenth century authors such as Ralph Lever in the Arte of Reason had already underlined the usefulness of resorting to compounding in English (quoted by Foster Jones, 1953: 126).

[2] In accordance with Baconian stylistic patterns.

[3] Like Boyle himself in some of his works.

SYNTAGMATIC lexical relation. They are linguistically predictable to a greater or lesser extent […].”

The very broad aspect of this definition, which derives from Firthian linguistics and was followed by Halliday (1966) and Sinclair (1966, 1991), implies that position and frequency of occurrence are both valid criteria for the characterisation of collocations. In principle, there is not a sense relation between the collocates, at least from the very beginning of the use of juxtaposed forms, and their co-occurrence is not fixed.

Benson et al. (1986) stated that, among the possible lexical combinations, the different degrees of semantic cohesion allow the establishment of the following taxonomy from less to more cohesive structure:

Free combinations

Idioms

(Typical) collocations

Transitional combinations

Compounds

This gradation affords a different perspective on the phenomenon, since the authors here take for granted the existence of semantic coalescence between the lexemes of a collocation.

The frequent co-occurrence of two or more elements can bring about their lexicalisation[4]. The term “lexicalisation” itself can be interpreted as fusion or “univerbation of a syntactic phrase or construction into a single word” (Brinton and Closs Traugott, 2005: 48-91).

Lipka claims (2004: 3) that lexicalisation is the outcome of an increasing unity (or “wordiness”) of the form and concept, and its familiarity, as an item, to the members of a larger or smaller speech community. In this paper I will visualise collocations and compounds as existing at opposite ends of a line (see Diagram 1 below). The line is a depiction of the diachronic process (lexicalisation) in which collocations are involved on their way towards the process of compounding.

[4] Kavka (2009: 21) claims that “compounds can be viewed … as idiomatic expressions”.

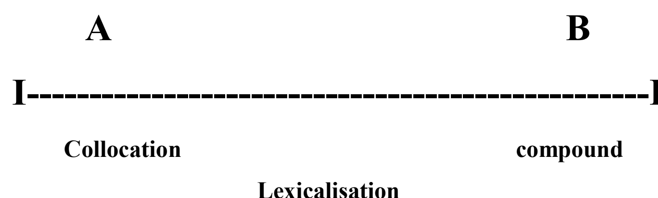


Diagram 1. Diachronic development.

Consequently, when analysing any sort of collocational structures, different degrees of lexicalisation may be found. Compounding is the final stage: the elements forming a compound must have acquired a high degree of lexical and semantic cohesion so as to be regarded as a complex lexeme or lexical unit.

A compound is defined by Bauer (1983: 29) as "a lexeme containing two or more potential stems". This broad definition is then narrowed by the classification of compounds into four different groups according to semantic criteria, that is, the relationship in terms of meaning between the grammatical head and the preceding element. Compounding, then, means "the unification of parts which are no longer independent, there is stress shift to the first syllable and semantic motivation is lost" (Brinton and Closs Traugott, 2005: 34).

Some authors argue that lexicalisation is a process that affects "larger-than-words objects" (Hohenhaus, 2005), such as fixed expressions, idioms and clichés (Lipka, 2005: 40). However, I agree with Lipka (2005: 40) that "lexicalisation is only motivated for units of the lexicon, like simple and complex lexemes and lexical units, but that institutionalization is not restricted in this way". He goes on to say that most cases of institutionalisation (like some collocations—green with envy—or routine formulas—bottoms up—) are culture-specific. In this sense, in the language of science, one may find that some constructions, after having undergone a lexicalisation process, are discipline-specific. Therefore, I could talk about a "technicalisation" process as the basis of a specialised linguistic domain or jargon, particularly when these multi-lexematic combinations are known and accepted by the corresponding discourse community.[5]

---

[5] Ward (2007) also mentions the technicalisation process.

3. Corpus material

The corpus material used in this study was analysed with reference to three parameters: date or time-span, discipline, and lexical category. For the purpose of analysis I have selected texts of two different disciplines, Medicine and Astronomy, these representing text types aimed at different kinds of reading public. The medical text, entitled A Choice Manual of Rare and Select SECRETS IN PHYSYICK AND CHYRURGERY; Collected, and Practised by the Right Honorable, the Countesse of KENT, late deceased, was written by William Shears and published in 1653.[6] Astronomy is represented by samples extracted from Armonicum Coeleste: or, the Coelestial Harmonie of the Visible World, dated 1651, and A Book of Knowledge in three Parts, written in 1663 by Samuel Strangehopes. Both these works are likely to be included in the forthcoming CETA (Corpus of English Texts on Astronomy)[7] which forms part of the Coruña Corpus: a collection of samples for the historical study of English Scientific Writing.[8] The three works, then, were published in the second half of the seventeenth century, a period when the influence of Empiricism on the discourse of science can be said to have begun.

As can be seen in Table 1 below, a total of 36,268 words will be analysed, 20,874 belong to the medical text and 15,394 to the Astronomy texts.

---

[6] Shears' text (1653) has been transcribed by the team compiling the Corpus of Early English Recipes (CoER). This particular sample has been transcribed by Drs Alonso Almeida and Ortega Barrera to whom I am deeply indebted.

[7] I want to thank Päivi Pahta and Irma Taavitsainen for their counselling in the compilation of CETA which, in turn, has made possible, to a certain extent, the writing of this paper.

[8] The Coruña Corpus of English Scientific Writing is a current project at the University of A Coruña (Spain) by the Research Group for Multidimensional Corpus-based Studies in English (MuStE). The main interest of the group is the study of language change and variation in scientific texts. One of the subcorpora being compiled is CETA (Corpus of English Texts on Astronomy). More information about the research group can be found at http://www.udc.es/grupos/muste

Table 1. Corpus material

Title, author	Date	Discipline	Number of words
A Choice Manual of Rare and Select SECRETS IN PHYSYICK AND CHYRURGERY; Collected, and Practised by the Right Honorable, the Countesse of KENT, late deceased by William Shears.	1653	Medicine	20,874
Armonicum Coeleste: or, the Coelestial Harmonie of the Visible World, by Vincent Wing	1651	Astronomy	6,650
A Book of Knowledge in three Parts, written by Samuel Strangehopes	1663	Astronomy	8,634
TOTAL NUMBER OF WORDS			36,268

Nouns are the chosen lexical category here. Sager, Dungworth and McDonald (1980), and, more recently, Nevalainen (1999), among others, have pointed out that they are the most relevant lexical category in scientific writing. In addition, the majority of compounds contain N+N (Quirk et al., 1985: 1567; 1570). However, not all members of the nominal class found in texts have been considered in the analysis. Since this paper seeks to explore differences between compounds and collocations, only those structures that are (or appear to be) compounds have been taken into account. Derivatives and simple nouns, then, have been disregarded.

Place names and proper nouns have also been disregarded, as well as cardinal points, nouns denoting seasons, days of the week, months and zodiac signs, except when they form part of the compound as in:

(1) northeast wind (Strangehopes, 1663: 30)

On the contrary, nominalisations of two types, -ing (grafting, moistening) and adjectival (riches, contraries) have both been taken into consideration.

4. Analysis and results

1,013 different nouns or types were found corresponding to 9,681 tokens, more or less equally represented across the two disciplines (ast: 2,654 + 2,010; med: 5,017). Of all those types, 121 (11.94%) correspond to compounds or compound-like structures. Although in the seventeenth century compounding does not seem to be as productive a process of word-formation as it was in previous periods, especially in Old English, it stands out as a useful mechanism for conveying scientific contents in a simple, clear and concise style, as demanded by contemporary writers of science. The idea of clarity and concision, which was on the minds of seventeenth-century authors, crystallised a century later in Margaret Bryan[9]'s preface to A compendious system of astronomy in a course of familiar lectures (1797):

> I know that I have no claim to the public suffrage, only on account of the clearness of my illustrations, which, as well as the diagrams, are principally original. As to the phraseology, I fear it is too deficient in ornament to procure me any credit; yet I hope the clearness of elucidations may gloss over the imperfections of the stile in which they are delivered:— Had I copied that of other authors, I might perhaps have rendered these Lectures more pleasing, although less intelligible to my pupils; who, being familiar with my diction, understand my illustrations much better, as I have thence been able to deliver them more naturally and forcibly. (Margaret Bryan, Preface, viii, 1797).

The use of compounds is a way of compressing the message without losing simplicity and precision. In the analysis of compound nouns I will work with two variables, etymology and discipline, which will, hopefully, afford new insights into the use of compounding in scientific discourse.

4.1 Etymology

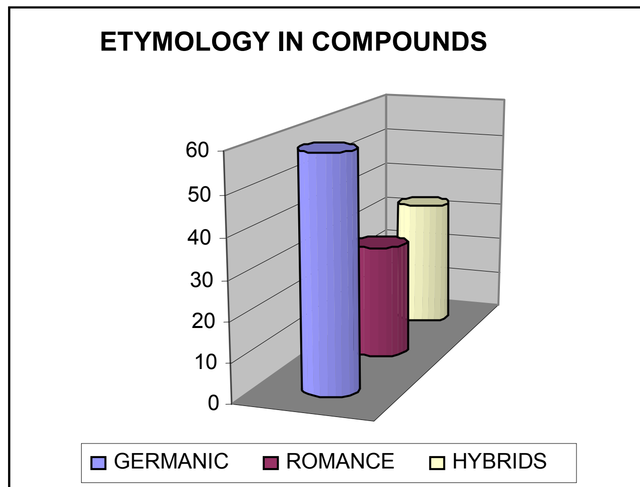
Compounds seem to show some peculiarities regarding etymology. For this study the online version of Oxford English Dictionary (OED) was consulted, and the ultimate origin of each term taken. All the different provenances that were found in it were classified into three groups:

---

[9] She was one of the first women writers on astronomy whose works were published under her own name.

Germanic, Romance and hybrid[10]. Hybrids contain those forms that combine elements of both Germanic and Romance descent.

The data suggests that compound nouns generally stem from Germanic sources, as seen in Graph 1.



Graph 1.Etymology in compounds

The observed tendency here for the creation of compounds from native stock contradicts previous studies that found Romance sources to be preferred for derivative forms (Moskowich, 2008). Some examples from the current data can be found in (2)-(4):

Germanic:

(2) earthquake (Strangehopes, 1663: 47),

[10] I have chosen to use these labels because of their agglutinating nature ('Romance', for instance, covers Latin and other related languages) though I am aware that much controversy surrounds the terms. My intention has been to simplify, so as to see "vernacular origin" as opposed to "others", mainly of Romance provenance. On the subject of etymologies, the OED3 is undertaking a full revision of etymological origins which, when completed, historians of the English language will have to take into account.

(2b) sore feets (Stangehopes, 1663: 35)

Romance:

(3) aqua mirabilis (Shears, 1653: 106),

(3b) orange floures (Shears, 1653: 205)

Hybrids:

(4) fixed stars (Strangehopes, 1663: 25)

(4b) gumdragon (Shears, 1653: 985/6)

4.2 Discipline

According to the second variable, the number of compound types found in both disciplines is approximately the same: 60 in Astronomy texts and 61 in the Medicine text. On closer inspection, though, this numerical similarity belies significant differences, which can be found when examining the etymological origin of compounds in each discipline.

4.2.1 The etymology of compounds and lexicalisations in Astronomy texts

As Table 2 shows, more than half of all the compound nouns found are of Germanic origin (51, 6%), followed by hybrid formations, with more than one third (36, 7%) of the total compound nouns in the Astronomy texts. Instances of Romance provenance represent only 11, 7 % of all these nouns.

Table 2. Compounding and etymology in Astronomy texts

	Romance		Germanic		Hybrids		Total	
Compounding	7	11.6%	31	51.6%	22	36.6%	60	

Examples[11] in (5) illustrate this predominance:

> (5) Earthquakes, fore-head, rainbow, shipwrack, sun rise, sun set, witchcraft. (Strangehopes, 1653)

An example of hybrid formation can be found in (6) below:

> (6) Windchollick (Strangehopes, 1663: 40)

The clear abundance of Germanic elements in the compounds found in the Astronomy samples is, no doubt, due to the fact that at least one of these texts could be classified as non-professional in nature. The work by Strangehopes is a good example of a text in which Astronomy and Astrology are not yet separate disciplines. In addition, this text provides a basic account of certain daily events in relation to other celestial processes. This could explain the use of common, everyday vocabulary in a "specialised text". The popularisation of knowledge was a *locus comunis* of the intellectual climate that pervaded seventeenth-century society. The Humanist trend of the preceding century had first introduced the notion that knowledge was of value to all individuals, regardless of their social status. Moreover, Astronomy was seen as a useful and practical science in fields such as navigation (Inkster, 1992: 119). Clearly, though, not everything written was aimed at the same kind of audience. Texts could range in their degree of informativeness depending on whether they were addressed to scholars, less educated readers, or laymen.[12] I would suggest that "addressee or type of audience" might well condition the lexical patterns used to convey scientific information, however specific and technical the discipline under discussion.

Wing's book, unlike that of Strangehopes, is addressed to a more specialised audience, as can be inferred from the author's own words:

---

[11] The text is scientific, from the 17th century point of view and these are examples of terminology used in scientific texts.

[12] Taavitsainen (2004) has classified Middle English medical texts into three different layers or levels of "informativeness", so that we can find commentaries, compilations and question-answer formulae (Moskowich, 2008).

> The first Book containeth those necessary and immediate Elements of TRIGONOMETRY abstractly propounded, which as the foundation to the superstructure, are laid down in a due Method and compendious manner. (Wing, 1651: 1)

4.2.2 The Etymology of compounds and lexicalisation in the medical text

My analysis of the medical text reveals an abundance of compound types with different origins, as can be seen in Table 3:

Table 3. Word formation and etymology in the Medicine text

	Romance		Germanic		Hybrids		Total	
Compounding	19	31.14%	28	45.9%	14	22.9%	61	

Of all cases, 31.14% have a Romance provenance, as in examples (7) to (9), 22.9% have a mixed etymology, as in (10), and 45.9% have been obtained from native forms, as in (11):

(7) ambergreece

(8) Venice Turpintine

(9) aqua mirabilis

(10) fixed stars (Strangehopes, 1663: 68)

(11) brimstone (Shears, 1653: 1)

The formal characteristics of the Germanic lexicon used in this book may have played a part in the process of compounding. Monosyllabic items of native provenance which are transparent can be juxtaposed to others to obtain new compounds (Gotti, 1996: 22). These formations could equally meet the pragmatic principle of maximum transparency[13], as in example (12):

13 Gotti (1996: 21) defines it as "the principle by which the specialist creates terms in such a way that their forms clearly reflected the concepts to which they refer".

(12) pennyworth[14]

Once more, as in the Astronomy sample by Strangehopes, the text is practise-oriented rather than academic; that is, the intended audience is the average practitioner who needs a clear and simple lexicon to grasp the information contained in the written text.

5. Some unclear cases

The border between compounds and collocations is a fuzzy one. In the data under assessment here there are some instances which do not fit neatly into either of these categories, but which occupy different positions on a lexicalisation scale. Apart from the prototypical elements of the category "compounds" that have been seen before, I have come across certain types that could be viewed as peripheral to the class (Rosch, 1978).

Table 4 below lists these unclear types found in each text:

Table 4. Unclear types

Medicine text (Shears)	Astronomy text (Wing)	Astronomy text (Strangehopes)
Spirit of wine	Line of the Auges	Lord of the Assendants
	Sine of the angle	Northeast wind
	The Sine of the Horizontal Parallax	Angle of the earth
	The Auge of the Epicycle	South angle
	The Auge of the Excentrique	East angle
	The Auge of the Moon	Angle of the west
	Center of the Orbe	North angle
	Center of the Moon	Wheel Of fortune
	Center of the sifnifer	Angle of the south
	Synodicall Variation of the Moon	Fall of the leave
	Center of the Epicycle	Lord of the Eclipse
	The center of the planetary orbits	
	Circle of variation	

[14] This term is not scientific as far as I can see. Perhaps we need the context cited to be sure. In any case, it also has many lexicalised senses, such as a bargain, something of little value, a small quantity of something etc., but it is not the item that is scientific but the text in which it is used.

	Elliptique Orbite	
	Elliptique circle	
	Circumference of the pricked	
	Center of the pricked	
	Centre of the Equant	
	Sine of he Hypothenusall	
	Centre of the world	
	Circle of Altitude	
	The center of the orbite	
	The sine of the angle	
	The tangent of the angle	
	Zodiacall circle	

A total of 44 unclear types (4.4 %) are distributed as follows: only one type in the Medicine text, but 43 in the Astronomy samples (32 in Wing's Harmonicon and 11 in Strangehopes' A Book of Knowledge).

This indicates that there are fewer unclear or peripheral cases in those samples where compound nouns of Germanic origin predominate (Shears' A Choice Manual and Strangehopes' A Book of Knowledge). Though these texts belong to different disciplines, they share a common target audience. They can be included within an informative kind of text-type.

From a structural point of view, these instances can be grouped as follows:

Type I. N+N: noun + noun
Type II. N+A: noun + adjective
Type III. A+N: adjective + noun
Type IV. NPs containing an N+PP: noun + prepositional phrase

Type I, N+N combinations (northeast wind, south angle, north angle, east angle), show the highest degree of lexicalisation. They might, therefore, be regarded as compounds rather than NPs.[15] From a semantic point of view, they apparently convey a single meaning which is predictable from the meaning of the grammatical head, which in turn is modified by the left-hand element. As a result, the compound noun is a

[15] It can be inferred that I consider these elements not as syntactic constituents but as morphological entities.

hyponym (northeast wind) of the nuclear lexeme (wind). It is a sort of endocentric compound. In this sense, these compounds are precise and specialised, though on some occasions the lexical unit that generates the compound is used in common, ordinary speech (wind).

Multi-lexeme constructions containing an N and an A, either pre- or post-posed (zodiacall circle, circle Equant), could also be analysed as endocentric compounds, hyponyms of the corresponding head rather than NPs. N+A compounds are illustrated by examples such as circle excentrique, circle equant. These are cases that correspond to the French type, in which the adjective occurs in post-position (Moskowich, 2002). Could these combinations be understood to be compounds?[16] They should be as the combination of the two simple lexical units generates a third one with a specific meaning within the astronomy discipline.

Less lexicalised are those combinations formed by N+PP (line of the Auges, center of the Orbe, the tangent of the angle). Could they be interpreted, once again, as compounds? For an affirmative answer, the explanation might be that of-phrases as post modifiers imitate French style, transforming English scientific discourse into a more analytic variety of the language[17]. In addition, the semantic cohesion the elements of this structure exhibit might be symptomatic of a certain level of lexicalisation. Hence, their consideration, not as mere collocations, but as compounds, is possible. This is especially the case of those examples which admit the double format: N+N or N+PP, as in south angle or angle of the South. In this sense, I agree that "complex lexemes are nominalizations of the respective collocations" (Lipka, 2005: 40), and that it is only a question of time as to how long it takes for each of these structures to become a compound (or not).

In other examples, however, the sequence N+N is not possible:

---

[16] At least, they were considered by Jespersen (1949: 25) as compound nouns. In fact, in Jespersen's terminology these "chiefly French" constructions are "compounds with post-adjunct adjectives".

[17] As Di Sciullo (2005: 25) has observed "In French, root compounds may include a phrasal constituent. A preposition precedes the phrasal constituent and is generally one of the set of grammatical prepositions, such as de and à, as in serviette de table and table à café."

(13) Lord of the Ascendants	(14) *Ascendants Lord
(15) Wheel of Fortune	(16) *Fortune wheel
(17) Lord of the Eclipse	(18) *Eclipse Lord
(19) Fall of the leaf	(20) *Leaf Fall[18]

Spirit of wine only occurs as N+PP, but in the same text wine glasse is also attested. They are coexisting variants in which the order of the elements differs. Maybe it is only a question before one of the two variants disappears or that they specialise their meanings and acquire different uses.

If criteria at different linguistic levels are to be applied as a means of showing whether these instances can be admitted as compound nouns or not, phonology must first be discounted. Since we are dealing with written material, stress cannot be used as an accurate indicator of compounding. Spelling is not a valid criterion either, since no process of standardisation had been completed at the time (see, for example, the alternative spellings earth quake/earth-quake/earthquake in the same texts here). Neither does the fact that two words appear without a hyphen necessarily indicate they are independent members of the same NP, as was seen above.

From a syntactic point of view, there are three properties that could play a part in determining compounding, namely, the recursivity principle, the right-hand head rule, and premodification by the intensifier very. None of these can be applied to the above-mentioned instances.

Some morphological constraints also act as obstacles in the interpretation of these sequences as compounds. The right-hand element is not always marked for number. In my view, some semantic

---

[18] My special thanks to the anonymous reviewer of this paper who pointed that the *OED* attested *leaf-fall* as a poetic expression. *Leaf-fall* occurs under the entry leave, n[1] as a special combination used in poetry and in Botany. This seems to support the idea that when the referent of the lexical units in the NP (N+PP (P+NP)) form part of ordinary speech and, consequently, are frequently used, the structure is more prone to lexicalise and have an N+N counterpart that is understood as a compound.

impediments are also inherent to the right-hand element (such as being either a proper or an abstract noun).

(21) fall of the leaves (Strangehopes 1663: 42)

(22) *lord of the eclipses (Strangehopes 1663: 69)

(23) *angle of the souths (Strangehopes, 1663: 54)

(24) *wheel of fortunes (Strangehopes, 1663: 56)

A different kind of behaviour regarding the expression of number in examples (21), on the one hand, and (22)-(24) on the other, can be observed. More extreme approaches would consider prepositional phrases in the above examples as phrasal adjectives functioning as post modifiers in a Noun Phrase (Gross, & Miller, 1990) instead of as prepositions embedded in a compound, as in French (Di Sciullo, 2005).

On the contrary, these unclear cases can, from a semantic point of view, be considered compounds, since they are perceived as a single unit and express a single content (Zanvoort, 1972). Semantic narrowing (or a restricted use, at least) seems to be playing a part in the way in which these structures are perceived. Moreover, there seems to be some sort of etymological conditioning also in the sense that a more restricted context of use is often associated with non-Germanic origin in the data from my corpus, for example in Center of the moon vs center of the epicycle. Those terms that seem to be more specific are less frequently used and come from “Latinate” languages.

6. Conclusions

The findings in this paper show that compounds mainly descend from a Germanic language, maybe conditioned here by the fact that two of the three text samples under survey are addressed to a less literate type of audience. The vocabulary had to be within the reach of the reading public interested in scientific matters since scientific texts as a specific-purpose product were widely disseminated, and were carefully attuned to the demands of the audience.

The combination of two potential stems/bases does not always immediately end in the formation of a prototypical compound.

Combinations of N+N, A+N or even N+A and N+PP can form examples of compounding, although this claim comes with some qualifications, especially relating to the provenance of vocabulary items and the influence of French syntactic structures. Is circle equant less a compound than attorney general? No doubt the fact that both items have a functional distribution, and are therefore to be found in particular text-types (Görlach, 2004), explains that both can be considered compounds.

The phenomenon of lexicalisation is measured in terms of semantic cohesion among the items of a collocational/compound structure or, what is the same, through the systematic "association of lexical items that regularly co-occur" (Halliday and Hassan, 1976: 284).

This cohesion appears to be less so when there are some intervening grammatical words, as is the case with the preposition of or any article (either definite or zero). But the weight of the lexical items in the construction contributes to giving it a greater semantic cohesion, so as to consider its structure as compound-like. Although formally speaking they are closer to NPs, I argue that they have undergone a lexicalisation process and a parallel technicalisation phenomenon. As far as the social use of these expressions, they seem to be discipline-specific (the sine of the angle; the circumference of the pricked). With other PPs in which the head of the NP is of Germanic provenance, technicalisation seems to lose strength and, simultaneously, cohesion seems to be of a lesser degree. We can speak, then, of NPs (centre of the Earth, centre of the world).

7. References

Primary Sources

Strangehopes, Samuel. 1663. *A Book of Knowledge in three Parts*.

Shears, William. 1653. *A Choice Manual of Rare and Select SECRETS IN PHYSYICK AND CHYRURGERY; Collected, and Practised by the Right Honorable, the Countesse of KENT, late deceased*.

Wing, Vincent. 1651. *Armonicum Coeleste: or, the Coelestial Harmonie of the Visible World*.

Secondary Sources

Bauer, Laurie. 1978. *The Grammar of Nominal Compounding.* Odense University Studies in Linguistics. Odense: Odense University Press.

Bauer, Laurie. 1983. *English Word-Formation*. Cambridge: Cambridge University Press.

Beal, Joan C. 2004. *English in modern times 1700-1945.* London: Arnold.

Benson, Morton, Evelyn Benson and Robert F. Ilson. 1986. *The Lexicographic Description of English*. Amsterdam/Philadelphia: John Benjamins.

Brinton, Laurel and Traugott, Elizabeth C. 2005. *Lexicalization and Language Change*. Cambridge: Cambridge University Press.

Bryan, Margaret. 1797. *A compendious system of Astronomy in a course of familiar lectures: in which the principles of the science are clearly elucidated so as to be intelligible to those who have not studied mathematics: also trigonometrical and celestial problems, with a key to the ephemeris, and a vocabulary of the terms of science used in the lectures which later are explained agreeably to their application in them*. London, Printed for the author, and sold by Leigh and Sotheby, York Street, Covent Garden; and G. Kearsley, No. 46, Fleet Street.

Crystal, David. 1997. *Dictionary of Linguistics and Phonetics.* 4th Edition. Blackwell Publishers.

Di Sciullo, Anna M. 2005. “Decomposing Compounds”. *SKASE Journal of Theoretical Linguistics,* 2/3: 14-33.

Foster Jones, Richard. 1953. *The Triumph of the English Language*. Stabdford: Standford University Press.

Görlach, Manfred. 2004. *Text-Types and the History of English*. Berlin: Mouton de Gruyter.

Gotti, Maurizio. 1996. *Robert Boyle and the Language of Science*. Milan: Guerini Cientifica.

Gross, Derek & Miller, Katherine. 1990. “Adjectives in Word-Net”. *International Journal of Lexicography*, 3/4: 265-277.

Halliday, Michael A.K. 1966. “Lexis as a Linguistic Level”. In Bazell, Charles E., J.C. Catford, M.A.K. Halliday and R.H. Robins (eds.) *In Memory of J.R.Firth.* London: Longman. 148-162.

Halliday, Michael A.K. & Hasan, Ruqaiya. 1976. *Cohesion in English*. Longman. London and New York.

Hård, Mikael & Jamison, Andrew. (eds.). 2005. *Hubris and Hybrids. A Cultural History of Technology and Science.* New York, Routledge.

Hohenhaus, Peter. 2005. "Lexicalisation and Institutionalisation". In Lieber, Rochelle and Stekauer, Pavol (eds.) *Handbook of English Word-Formation.* Dordrecht: Springer. 353-373.

Inkster, Ian. 1992. "Advocated and Audience—Aspects of Popular Astronomy in England, 1750-1850". *British Astronomical Association*, 3: 119-123.

Jespersen, Otto. 1949. *A Modern English Grammar on Historical Principles*. London: Allen and Unwin.

Kavka, Stanislav. "Compounding and Idiomatology". 2009. In Lieber, Rochelle and Stekauer, Pavol (eds.) *The Oxford Handbook of Compounding*. Oxford: Oxford University Press.

Lipka, Leonard, Handl, Susanne and Falkner, Wolfgang. 2004. "Lexicalization & Institutionalization: The State of the Art in 2004". *SKASE Journal of Theoretical Linguistics.* 1/1: 1-18.

Lipka, Leonard. "Lexicalization and Institutionalization: Revisited and Extended. A Comment on Peter Hohenhaus *Lexicalization and Institutionalization*". *SKASE Journal of Theoretical Linguistics*, 2/2, (2005) 40-42. 20 Oct 2008. <http://www.skase.sk/Volumes/JTL03/ 05.pdf>

McIntosh, Angus and Halliday, Michael A.K. 1966. *Patterns of Language: Papers in General, Descriptive and Applied Linguistics*. London, Longman.

Moskowich, Isabel. 2002. "The Adjective in English. The 'French type' and its Place in the History of the Language". *Folia Linguistica Historica,* 23/1-2: 59-71.

Moskowich, Isabel. 2008. "'To lerne sciences touching nombres and proporciouns': The Proportion of Affixation in Early Scientific Writing". *English Studies*, 89/1: 39-52.

Nevalainen, Terttu. 1999. "Lexis and Semantics". In Lass, Roger (ed.) *Cambridge History of the English Language: 1476–1776.* Cambridge: Cambridge University Press. 332-458.

OED online. Oxford University Press. 1989. Second edition. 05, 02, 2009. http://dictionary.oed.com/.

Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey and Svartvik, Jan. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Rosch, Eleanor. 1978. "Principles of Categorization". In Rosch, E. & Lloyd, B.L. (eds.) *Cognition and Categorization.* Hillsdale: Lawrence Erlbaum Associates. 24-48.

Sager, Juan C., Dungworth, David and McDonald, Peter F. 1980. *English Special Languages. Principles and Practise in Science and Technology*. Wiesbaden (Germany): Oscar Brandstetter Verlag.

Sinclair, John. 1966. "Beginning the Study of Lexis". In C.E. Bazell, J.C. Catford, M.A.K. Halliday, and R.H. Robins (eds.) *In Memory of J. R. Firth*. London: Longman. 410-430.

Sinclair, John. 1991. *Corpus, Concordance, Collocation.* Oxford: Oxford University Press.

Taavitsainen, Irma. 2004. "Transferring Classical Discourse Conventions into the Vernacular". In Taavitsainen, I. and Pahta, P. (ed.) *Medical and Scientific Writing in Late Medieval English*. Cambridge: Cambridge University Press. 37-72.

Verdaguer, Isabel and Garau, María. 2005. "Las combinaciones léxicas en el inglés científico: presentación de una base de datos. Volumen monográfico". *Revista española de lingüística aplicada*. 67-84.

Ward, Jeremy. 2007. "Collocation and Technicality in EAP Engineering". *Journal of English for Academic Purposes* 6: 18-35.

Zandvoort, Reinard Willem. 1972. *A Handbook of English Grammar*. London: Longman.

Student perceptions of English studies in Bulgaria

Milena Katsarska, Plovdiv University, Bulgaria

Donka Keskinova, Plovdiv University, Bulgaria

Abstract
This paper sets out and discusses the findings of a comprehensive questionnaire survey conducted among English Studies (ES) students in three Bulgarian universities: the University of Sofia (SU), The University of Plovdiv (PU) and the University of Veliko Turnovo (VTU) between December 2007 and May 2008. The discussion registers the condition of ES from the point of view of students at various levels of university programmes. The responses received give data about (a) the ES student constituency; (b) student attitudes toward the discipline and programme in the course of engaging with them; and (c) student expectations at the points of entry and departure in terms of content, delivery, their interests and career aspirations.

Keywords: English Studies; non-Anglophone contexts; survey; student perceptions; Bulgaria

1. Introduction and background considerations

This paper sets out and discusses the findings of a comprehensive questionnaire survey conducted among English Studies (ES) students in three Bulgarian universities: the University of Sofia 'St Kliment Ohridski' (SU), The Paisii Hilendarski University of Plovdiv (PU) and the University of Veliko Turnovo 'St Cyril and St Methodius' (VTU). The questionnaires were administered between December 2007 and May 2008 within the frame of the collaborative research project 'English Studies in Non-Anglophone Contexts: East Europe'.[1] The overall aim was to register the condition of ES—understood as comprised of both language/linguistics and literature/culture—from the point of view of students at various levels of university programmes in Bulgaria.

The responses received give data about (a) the ES student constituency; (b) student attitudes toward the discipline and programme in the course of engaging with them; and (c) student expectations at the

[1] Funded by the British Academy and the Leverhulme Trust and supported by the three universities in Bulgaria, together with The Open University, UK and the universities of Bucharest, Cluj and Constanta in Romania.

points of entry and departure at different levels in terms of content, delivery, their interests and career aspirations. Where appropriate I supplement the data with observations on curricula arrangements and content, institutional and educational documentation, and existing statistical data and scholarly publications. Whereas surveys of students' perceptions and attitudes in higher education (HE) are usually geared towards institutional (or, comparatively, across institutions and contexts) programme assessment in terms of access, achievement, career realization, learning outcomes, pedagogical efficacy, and/or academic performance and/or institutional evaluation, this survey was not for those purposes. The primary concern here was to view the discipline as practiced currently from (a) a learner's perspective, and (b) within a distinctively Bulgarian context—and thereby to provide a wide basis for comparisons and discussions.

Prior to detailing the patterns and departures in student views, it is worth outlining some context-specific features of the discipline in Bulgaria, especially in relation to student matters. I touch upon several of those before describing the structure of the student survey as well as the nature of the sample.

1.1. English Studies students in Bulgaria

High school graduates interested in pursuing a HE degree in English are opting for a discipline with a substantial institutional history in Bulgaria as evidenced in the historical overview by Shurbanov and Stamenov in 'English Studies in Bulgaria' (Engler and Haas 2000, 267-292) and as documented by Vesselinov (2008). From 1928 until the early 1980s their choice was limited to two universities (the second one since 1972) and would follow a single subject arrangement in a five-year course of study.[2] From the late 1980s prospective students have had more options in terms of institutions and degree arrangements with the introduction initially of combined subjects 'English and Bulgarian', later to become more diverse and involve English and another foreign language. These were accommodated in philology departments. These double majors

[2] How ES in Bulgaria compares to ES in other European countries in terms of differences and similarities in degree structure, curricular arrangements, personnel and delivery has been outlined to a degree in Kayman and Mesquita (2005).

occur at present in two variants—as philology degrees and as degrees in applied linguistics. Those high school graduates who applied for university in 1997 were the first to complete a four-year BA degree course of study and have the opportunity to continue in an MA programme of their choice.

Since the 1990s students applying for ES would enter an area of particularly dynamic growth in the HE sector in Bulgaria. In1989 they were about to join a recently enlarged community of about 500 ES students spread across four universities (two of them, in Plovdiv and Shumen, just having admitted first year students in a combined subject). By 2008, on the other hand, they had the choice of five state universities offering ES as a single major together with a varied portfolio of ES in a combined subject BA degree. By this time there were over 450 full-time BA level students at each institution engaged in ES. Additionally, ES students can now also be found in several associated colleges, like Smolyan and Kirdzali, which offer combined subjects 'ES and Bulgarian', and in private universities offering both—ES as a single subject and 'ES and other'. At present, high school graduates applying for ES take part in the second largest language-based recruitment area in Bulgarian universities after Bulgarian Studies. Bulgarian Studies has 120 allocations, while ES is next with 100. In comparison, across disciplines in all departments, Law has the largest student number allocation in state universities (250 places), History (130), Management with a Foreign Language (120), and IT (140)[3]. Large admission numbers notwithstanding, graduates wishing to enroll for an ES degree enter a very competitive field. For the 2009 admissions process at SU, 300 students who identified ES as their first choice competed for 100 available places.

Prospective students sit for entrance exams in English, which are currently pitched at the advanced level of linguistic competence. In terms of the *European Framework of References for Language Competences,* students of English are expected to perform at their admission exams at the level of proficient users (C1). The linguistic competence of applicants is tested by a composite exam which comprises a dictation, a multiple choice test, and essay writing, aimed at gauging levels of reading and listening comprehension, knowledge of grammar,

---

[3] Based on SU admission data.

vocabulary and spelling, writing skills and ability to present arguments. Once admitted to the programme ES students are expected to study almost entirely in the medium of English.

1.2. The student survey

Within the above-mentioned project, the student survey discussed here was a key to acquiring a sense of those who become ES students in Bulgaria and how they perceive the current condition of ES. The questionnaire had three distinct, yet interrelated, parts. The first part was designed to gather data regarding the constituency of the ES student body according to programme; their educational background prior to entering university; their linguistic competences; socioeconomic and ethnic/nationality backgrounds. International and local mobility in ES enrolment was also a point of interest. The second part of the questionnaire encouraged students to reflect on their programme of study as a whole—insofar as they had grasped what the programme was seeking to achieve. The idea was to obtain an understanding of coherence/community/identity within the subject along the lines of programme content and programme delivery. In other words, responses to this section gave a sense of student expectations and interests and the extent to which their programmes are meeting these. This section also enabled students to reflect on the definition of ES in the context of Bulgaria, and in comparison to other contexts that they are aware of. The third and final part of the questionnaire probed students' experience of the discipline in relation to employment and professional prospects. Its objective was to track changing expectations as students progressed through different levels of study, and ultimately to relate ES in Bulgaria generally to career aspirations and prospects in Bulgaria today. A final question allowed space for students to describe the impact that ES has had on their lives in their own words.

1.3. The survey sample

A total of 417 BA and MA students in ES (and related) programmes at three Bulgarian universities responded to the questionnaire. The choice of locations was determined by several considerations. First, the place of each of those universities in the establishment and institutionalization of

the discipline: SU introduced the subject in Bulgaria and set up a department in 1928; VTU followed in 1972-3; and PU in 1988-9. Second, these three universities have the largest recruitment share and largest numbers of students currently involved in ES programmes for the country: for PU 591, SU 485, and VTU 440 studying full-time. Within our sample, students of English were attached to a variety of programmes, all leading directly or related tangentially to conventional English Studies (philology) degrees. These frequently involve the so-called double majors in English *and* Bulgarian, French, Russian, Japanese, Italian, Chinese, etc; or are developed within Applied Linguistics in view of the increasing demand for professional translators and interpreters. Among those, 96% of our informants come from BA full-time programmes and 4% MA English programmes. The structure of the sample across university, degree programme and year of study is given in table 1.

Table 1. Distribution by year of study.

Year of study in the degree	**University**		
	Sofia	**Plovdiv**	**V.Turnovo**
BA – 1 year	25%	15%	40%
BA – 2 year	36%	40%	11%
BA – 3 year	22%	26%	28%
BA – 4 year	17%	15%	12%
MA		4%	9%
Count	***125***	***202***	***90***

In order to reproduce the structure of the general sample spread across the three universities, the data are weighted as indicated in table 2. The statistics used hereafter allow us to regard the sample as representative for the three universities taken together.

Table 2. Weight data by University.

	University		
	Sofia	**Plovdiv**	**V.Turnovo**
Real sample	125	202	90
Weight	*1,07*	*0,81*	*1,35*
Weight sample	133	163	121

When interpreting the results it is necessary to take into account the stochastic error in different relative shares:

Table 2a. Maximal error (P=95%).

%	**3 97**	**5 95**	**10 90**	**15 85**	**20 80**	**30 70**	**40 60**	**50 50**
Max Error	**1,4**	**1,8**	**2,4**	**2,9**	**3,3**	**3,7**	**4,0**	**4,1**

2. Student body constituency

As in other contexts (for the UK, see Williams 2002), ES students in Bulgaria are predominantly female: 76% female to 24% male is the extension ratio from the admission quotas (for 2008, for instance, those were, SU 68/30 and PU 70/35). It is worth noting however that the imbalance comes mostly in the final year of study at BA level where women become an overwhelming majority at 89%. Hence the feminization of ES occurs in a cumulative fashion, from the entry point into ES at HE to point of graduation. Overall, among all the philology degrees at PU, for example, ES programmes have the lowest drop out rate.[4] The age of students involved in ES ranges between 19 and 22 years, with 23+ forming only 17% of the student body, which is

[4] In the case of PU, Slavic Studies 89 students in year 1 and 24 in year 5; for Bulgarian Studies 104 and 37, respectively; while for Bulgarian and English those are 81 and 40; for the English Studies major 112 and 65. Based on Zhivko Ivanov, 'Admission Trends and Statistics'; presentation at the Colloquium in Admission Issues, Smolyan, November 2008, slides 9-12, and 15.

consistent with HE sector as a whole. Bulgarian ES programmes register international recruitment of students (here I do not refer to Erasmus exchanges) without specifically holding recruitment campaigns outside Bulgaria. These students come mostly from the region. While 94.5% declare Bulgarian nationality, there are 3.1% Turkish, 1.3% Macedonian, as well as Serbian and Moldovan nationals. Among ES students of Bulgarian nationality 90.5% identify themselves as ethnically Bulgarian, 5.1% as Turkish and 4% as Muslim in response to 'If Bulgarian, ethnicity where applicable'. An almost negligible number of students choose not to respond here—0.4%, together with all other ethnicities they mentioned. There are significant differences in this regard across the three universities: SU comprises 100% ethnic Bulgarian ES students; VTU 98.7% Bulgarian and 1.3% Turkish; and PU 79.2% Bulgarian, 10.4% Turkish, and 9.4% Muslim. PU has the most ethnically diverse student body in ES and reflects the proportions in the Bulgarian population generally—83.9% Bulgarian, 9.4% Turkish, 4.7% Roma, and 2% others, etc. (NSI census data, 2001[5]). Given the regional spread of ethnic diversity in the country, the location of PU may account for this, as may economic factors, since the capital Sofia is the most expensive city in the country. In 2001 ethnic minority students (Turkish and Muslim, or Roma) comprised 0.04% of the overall student body in the humanities (Georgieva 2002, 115), seven years later, at least in ES, access for minority students seems to have improved.

The fee for students in English philology degrees in state universities is between 420-440 BGN per academic year for the humanities and social sciences. This can be put into perspective by noting that fees for economics majors are lowest at 260-267 BGN, sciences 614 BGN, medicine 960 BGN. According to the survey, 48% of ES students come from households with 600-1,200 BGN monthly income, 33% from families with an income below 600 BGN, and only 19% from families earning over 1,200 monthly. With 354 BGN being the average salary for the country in the first trimester of 2009 (NSI) and 240 BGN the minimal salary since January 2009, it is not surprising that students need to rely on a variety of ways to support themselves while studying—see table 3.

[5] NSI data lists the religious diversity in Bulgaria as follows: Orthodox Christian 82.6%, Muslim 12.2%, Catholic 0.6%, Protestant 0.5%, etc.

Since opportunities for stipends and fellowships are rather limited,[6] in all cases students have marked more than one source of support. Since major universities offering ES are located in big towns in-country mobility is inevitable, and 67% of ES students surveyed had moved to Sofia, Plovdiv and Veliko Turnovo from elsewhere.

Table 3. Financial support.

Financial support	**SU**	**PU**	**VTU**	**Total**
I work part-time and study	45%	22%	17%	28%
I work full-time and study	5%	11%	4%	7%
My parents support me	69%	78%	83%	77%
I have a stipend/fellowship	23%	21%	10%	19%
Other	4%	4%	6%	4%

Students usually enrol into ES programmes immediately after completing their High School (HS) education. The largest numbers of ES students come from an English language or foreign language HS, 28 and 26 per cent respectively. 11% are from a comprehensive HS background (without a particular specialization) and 10% from HS with science profiles. In this respect, the profile of ES students has changed significantly since the late 1980s and early 1990s when an overwhelming majority were English language HS graduates, the so-called English-medium HS. These were established in the 1950s and were the elite HS institutions with English medium instruction in a select way across the curriculum. In what are currently known as Foreign Languages HS (often the transformed former Russian medium HS post-1989) students receive a similar level of language instruction but may not experience English-

[6] The most common stipend is 120 BGN monthly on the basis of academic record average above 4.00 (6.00 is the highest grade in Bulgaria) from all exams in the previous academic year. There are also groups of students (with disabilities, single parents, etc.) who are entitled to the subsidy irrespective of grades or on preferential terms.

medium teaching elsewhere in the curriculum. Consequently, about 46% (at least) of students currently pursuing ES in Bulgaria have had no experience of studying literature in English before enrolling for the HE ES degree.

The changes with regard to language education at the secondary level in Bulgaria in terms of government policies, setting up of new schools, introducing English on a mass scale (from grade 1 of the primary level) have been documented extensively in scholarly publications (Thomas and Georgieva 2002; O'Reily 1998). For the purposes of the present paper, it is sufficient to note two trends. First, ES students come from widely varying HS backgrounds, with differing practices and policies for English language teaching and teaching in the English medium. Second, that, Bulgarian aside, students of English would have often studied another foreign language in addition to English.

Thus the linguistic competence 'portfolio' indicated by students in response to the questions 'How many languages do you use/can you communicate in?' and 'Specify level (basic, moderate, good, fluent) for reading, writing, speaking' comes as no surprise. 52% of the students indicate varying degrees of competence in two languages other than Bulgarian and 19% in three. Bulgarian and English (with 100% each) are followed by German (38%), Russian (19%), Spanish (16%), French (13%), Italian and Turkish (5% each). These numbers are irrespective of the type of degree, since with mixed majors or applied linguistics programmes one would expect such a variety, another language being a prerequisite. The variation between ES philology respondents, however, and 'ES and another' is within the 5% or less. Figure 1 illustrates further the linguistic competence of all students in terms of their self-assessed speaking, writing and reading competences in the first five foreign languages by the mean result from their self-grading on a scale from 2 (lowest) to 5 (highest).

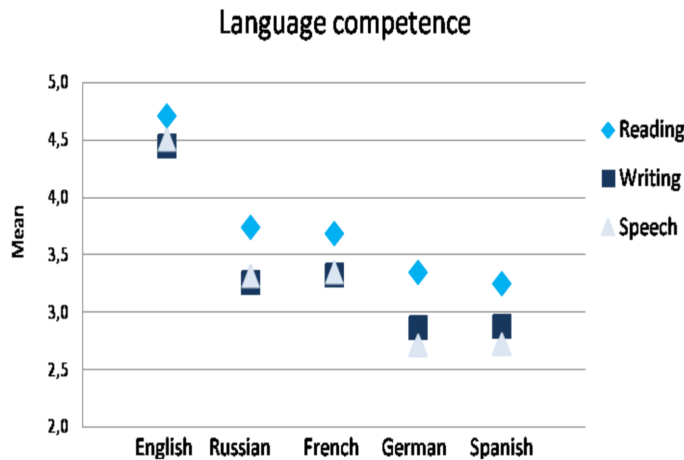


Figure 1

Two related issues emerge from the above observations. First, on the level of curriculum arrangements BA ES in Bulgaria is yet to capitalize on the foreign language competence in other than English languages of its students. Second, with regard to the varied paths by which students reach their required entrance level of English, the programme faces a tension between accommodating these differences or seeking to homogenize the student body in terms of linguistic competence[7] and awareness of ES issues. What the expectations of students themselves are of their programmes of study, to what extent their expectations are met are issues taken up in the next section.

[7] One of the 'instruments' in this respect become the English practice classes in the first year with the introduction in recent years of proficiency level textbooks—such as Leo Jones, *New Progress to Proficiency* (Cambridge: Cambridge University Press, 2002) and Kathy Gude and Michael Duckworth, *Proficiency Masterclass* (Oxford: Oxford University Press, 2003)—coupled with an incentive to produce new local university course books for the practical English classes.

3. Expectations and interests

Prior to actual enrollment in the ES programme prospective students in Bulgaria have access to a range of information sources. These include university catalogues and websites, promotional brochures and consultations with university faculty and/or high school teachers of English.

First and foremost, students enrolling in ES programmes expect to develop their linguistic competence and improve their language skills. Next, students expect to inform themselves about the UK and USA, study of history and culture of English speaking countries and, particularly, study of English and American literature. Figure 2 gives full details of student expectations.

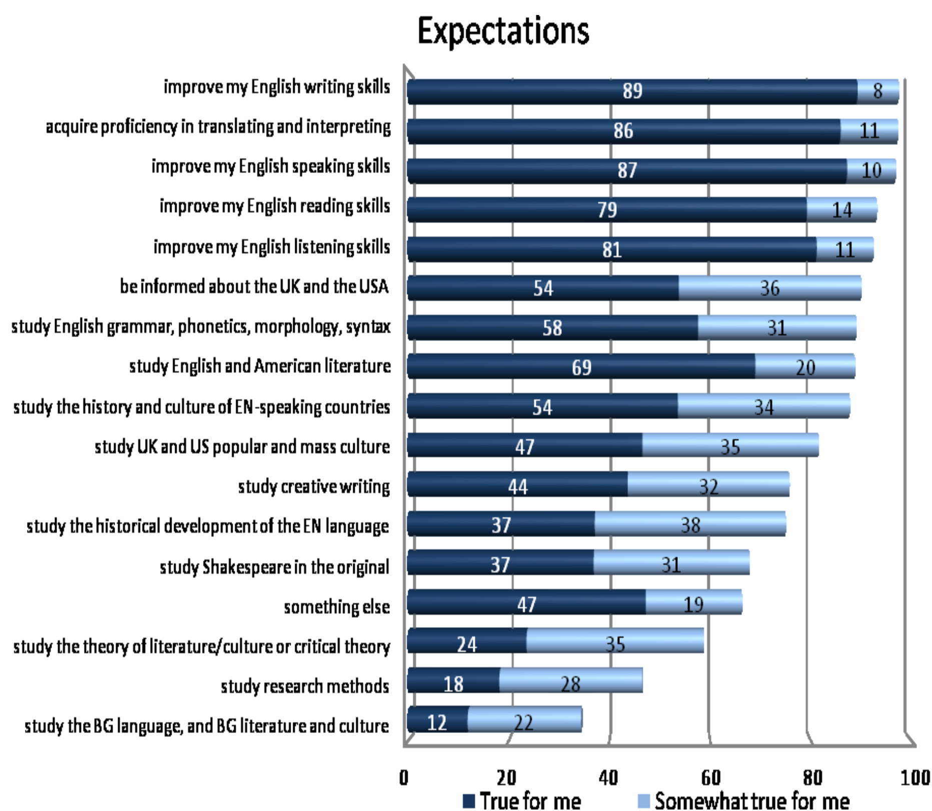


Figure 2

The high expectations to improve language skills and learn linguistics is consistent with the established conventions of ES (philology) programmes in Bulgaria. There are two notable exceptions here. Students seem to expect to engage with 'Historical development of the English language' to a lower degree, and even less to 'study Shakespeare in the original', while both are still significant parts of ES.[8] Student expectations veer toward areas which can be seen as more recent developments in ES, such as 'popular and mass culture' and 'creative writing', which have found a place in the curriculum only since the mid 1990s in Bulgaria. High expectation with regard to the former is naturally fostered not only by the ubiquity of media and popular culture texts, but also by these being part of the educational experience of students (formal and informal) in learning the English language. Areas unfamiliar to students prior to entry in university—e.g. critical theory or research methods, or areas which seem not to be obviously associated with ES—e.g. Bulgarian language and literature, account for low expectations. These are unfamiliar because the former do not merit a particular emphasis in HS curricula or, as is the case with the latter, are less immediately within an envisaged scope of ES programmes because historically ES emphasizes 'immersion' in the respective target language/literature/culture. Also, since students' expectation curve peaks at acquiring foreign language competences, it often seems difficult for them to foresee how the study of Bulgarian language or literature may contribute in this direction. On the whole, few students felt that the programme has changed their expectations (only 8%), 69% said the programme is concordant with their expectations to some degree, and 14% felt that their expectations were fully met.

ES students' stated academic interests resonate with their expectations in several ways, and are generally practice-oriented. Figure 3 shows the percentages of students who acknowledged specific academic interests to greater or lesser degrees:

---

[8] In the case of Historical linguistics (the theoretical study of linguistic change and textual practice with Old and Middle English texts) only SU (30 contact hours) and PU (90 contact hours) have kept it at the BA level, with the former reducing by 50% contact hour allocation after the transformation of the degree to a four-year BA cycle. Shakespeare still occupies a central place in the course in Renaissance English literature.

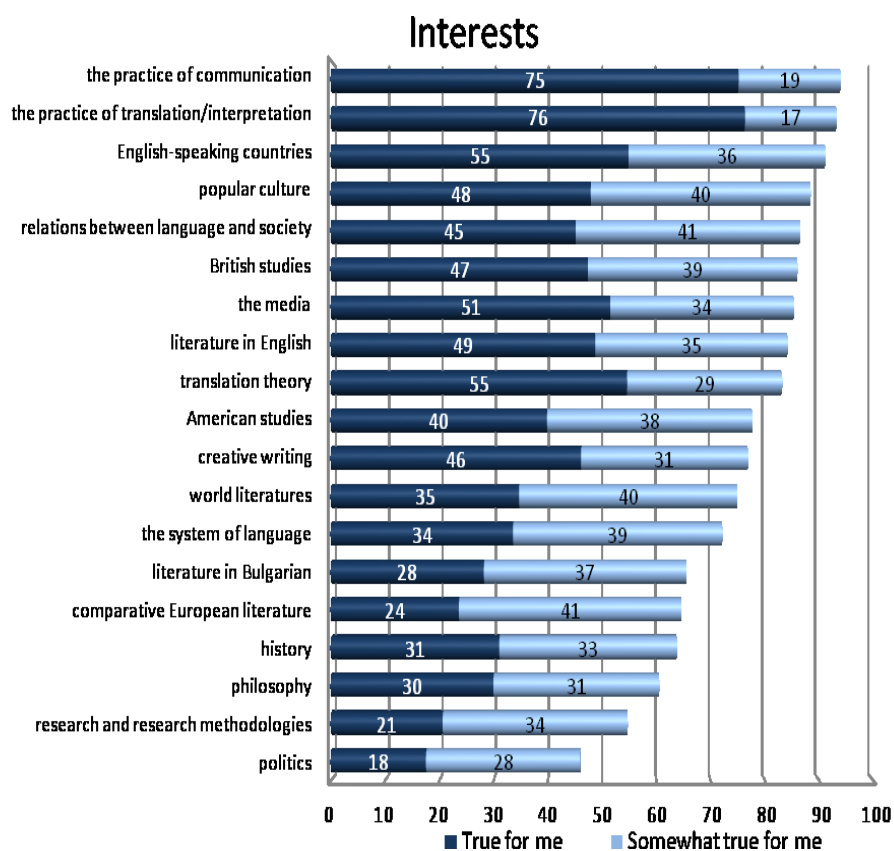


Figure 3

The greatest interest is in improving language skills and being informed about Anglophone contexts, and in the practical business of developing the ability to work as translators and interpreters. Thereafter, popular culture, the media, and sociolinguistics evoke roughly as much interest as 'literature in English'. Within this second group of pronounced interest, 86% declared a particular interest in British studies. Thus it ranks higher than interest in American studies which occupies a top position in the subsequent cluster of declared interest areas together with 'creative writing', 'world literatures' and the 'system of language'. Comparing stated interests between ES single major and 'ES and other', there are

four distinctive areas which differentiate the two categories, with over 10% differences in responses, which is the statistically significant marker in this section. The 'English and other' group state 86% overall interest in 'the system of language' (with 45% positive agreement) as compared to 65% (28% positively) of the ES major group. The most marked interests of students in ES major, as opposed to 'ES and other', are in the spheres of 'literature in English' (positive agreement difference 53% vs. 41%), British studies and American studies. The latter two are especially prominent with over 20 and over 15 % difference in positive identification 'true for me' respectively. It appears then that 'ES and other' has a more pronounced interest in the linguistic side of the discipline, and ES majors favour the literary and cultural studies side. I conclude this part by focusing on the two areas of least interest in the responses of students. Research methods rank low in both expectations and interests. At BA level this is a neglected area in the curriculum. Moreover, with the introduction of the four-year BA degree the final graduation requirement became a composite state exam, and removed the previously existing option of a dissertation submission, which accounts to a great extent for the responses. The fact that there is, in the zone of uncertainty, an overall 55 percent of interest in the area suggests some hesitant awareness of the relevance of research methods for ES.

One of the perhaps most contextually-indicative responses in this section is the categorically lowest area of interest for ES students: 'politics' with 46% overall and 18% positive agreement. In the open-ended questions of the survey, only one of the 417 respondents stated that being a student in ES contributed to an increasing awareness of and interest in politics. These responses derive from the prevailing attitude in Bulgarian ES academia which regards politics as a 'dirty word' which has little to do with serious academic work. There are complex historical developments and current dispositions underlying this. Simplistically put, for students politics involves 'talk about parties, politicians, elections, etc.' permeating the media and their everyday lives which is often associated with corruption, failure of government and other disheartening 'news of the day'. From a historical point of view, politics is usually associated with the rigid ideological framework that existed prior to 1989, when not only society but also academia and particularly ES were 'politicized' in the sense of the subscription to a rigid ideology. Consequently, during the transition period there was a loudly declared

and sweeping move to 'de-politicize' academic discourse and education on all levels. These are superficially sketched observations, but the point is borne out in the survey: politics slips away between two kinds of reductive understanding.

4. Defining English Studies in Bulgaria

The legal framework which informs the current composition of ES functions since the 2002 Higher Education Act, which postulates and controls, among others: a 4-year BA study, the academic hours load min. 2200—max. 3000; ending the degree with a State Exam; the ratio between habilitated and non-habilitated faculty teaching in the degree; the ratio between mandatory and elective courses but without positing prescriptions as to the nature of these core courses or any quotas of classes' allocation.

Given the composite nature of the ES (philology) degree in Bulgaria, which comprises both language/linguistics and literature/culture studies, the curriculum is designed so that courses may address both 'fields' in a balanced way, at least insofar as core obligatory courses go. Courses in English language practice occupy a considerable number of classes and run through the entire four-year course of study in the BA degree.

Students' sense of the definition of ES emphasizes the linguistic orientation of the degree as well as their own practice-bound interests in language. 52% of our respondents describe ES in Bulgaria as 'a study in language and linguistics'. Additionally, 20% of them identify it as 'a study of applied skills in language'. The highest ratio in the latter category represents students who are involved in 'ES and other' degrees at VTU and PU. The identification of Linguistics as the dominant domain is contained in the responses of ES single majors which outnumber the same response from their 'English and other' peers by 14%. Linguistics ranks especially high among SU students (59%), where a closer look at obligatory core courses reveals 690 academic hours allocated to courses in linguistics as opposed to 465 for Literature/Culture. VTU curricular arrangements present a fixed balance of 435 for each in the compulsory corpus. The Literature 'component' is emphasized by 14% of students, who describe the ES degree as 'a course in literary studies'. That percentage is higher in the final two years of the degree, reaching 25% in the fourth year. This shows the influence of

subject distribution in the curriculum, as students engage in more literary studies as they progress. When reflecting in general on the impact ES has had on them, over one third of the students bring to the fore the literary/cultural studies. In their words:

> Positively my engagement with the programme has influenced my interests in Literary Theory and Latin. [...] They broadened my perspective in areas I did not know were interesting (Literary Theory)…(BA ES SU, year 1)

And

> The education has influenced my interests in literature and introduced me some works and writers I'd hardly pay attention to on my own. Also gave me the chance to study and research in detail. Authors gave me broader background knowledge on American & British history & culture. The teachers should be more open and enthusiastic give the students opportunities to present works, authors, etc. of their own choice—more contemporary literature and so on. I feel the greatest benefit so far has been the elective courses and the foreign lecturers. (BA ES PU, year 3)

Often, as evident in the quotation above and in many of the other responses, when given a chance to express their views in open questions, students of English attribute higher value to their interactions with 'native' speakers and courses by guest lecturers in the programme. A number of responses express a recommendation to bring 'more native speakers' to teach on the programme. Students feel that being able to have 'native' communicative skills would improve their prospects, and put an extra value on foreign tutors. That in itself is worthy of further discussion in view of, among others, Holliday's (2008) analysis of the 'native speaker phenomenon' vis-à-vis politics of inclusion. The following observations on some of the ways in which the non-Anglophone context of Bulgaria interacts with the target subject and context(s) has a bearing on that.

5. Context engagement

The extent to which ES in Bulgaria registers the context in which the discipline is practiced within the teaching and learning paradigm is the question addressed here via student responses. Students were asked to comment on the presence and balance of the target and home languages in classroom practice and in the readings that accompany their ES

courses. Figures 4 and 4a give the summary of their responses broken down according to institutional location.

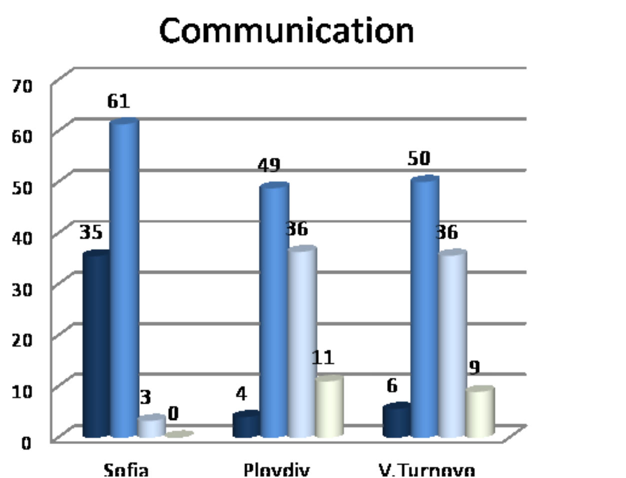


Figure 4

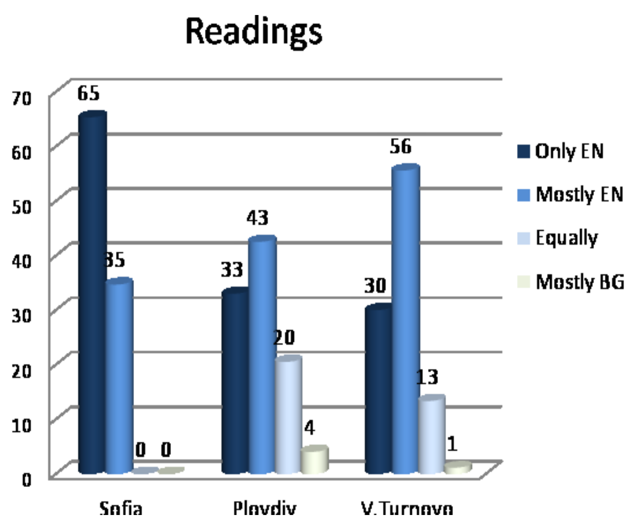


Figure 4a

Even a cursory look at the figures above indicates the dominant presence of the English language both in terms of classroom communication practice and in the assigned classroom and independent readings or bibliographies accompanying various courses. Communicative practice makes more use of the 'home' language, while texts remain overwhelmingly within the domain of the target language. This is also evidenced by a survey of syllabi and course descriptions where Bulgarian texts (required readings or assigned texts) appear only occasionally for the few courses which are taught in Bulgarian[9] or target Bulgarian-into-English translation practice. Within the current legal framework for HE there do not exist prescriptions either as to the language of instruction or of texts. In formal institutional documentation, such as degree and qualification descriptions, written in the Bulgarian language for programme accreditation purposes there is no specific reference to language either. If at all, language is mentioned in educational documentation pertaining to ES if it is not English, the general understanding being that the language degree is in the target language on a number of levels and focuses mostly on the target context.

Variations within the context of Bulgaria are registered through the statistically relevant distinctiveness of SU, where 65% of ES students state that their readings are 'only in English', compared to more mixed readings registered in PU and VTU. PU and VTU variations come mostly from the programme 'Bulgarian and English' and the joint subjects for the two universities respectively, which are developments since the late 1980s. But largely, a philosophy of 'total immersion into the target language' is followed in Bulgarian ES. Students regard this in a positive light. According to one student:

> English is a vocation—studying, teaching whatever. I catch myself sometimes that I think in English—so the programme in a way helped in the development of my second mother tongue—an optional language to think on if I'm tired of Bulgarian. (BA ES PU, year 4)

---

[9] In the ES degree currently there are 2 to 5 courses taught in Bulgarian, such as the Introduction to linguistics and Literary theory (for PU and VTU), the strand for Pedagogical qualification (courses in Psychology, Audio-visual methods, etc), which is elective at SU and VTU, and those classes in which the Bulgarian into English translation is practiced. In objective terms that would be ¼ of English practice classes would be assigning texts in Bulgarian, and 10% of the compulsory core courses at the BA level for ES at VTU, for instance.

A further brief point to be considered along the lines above is the presence of the ‘home’ context in assigned readings. On the whole, course syllabi rely mostly on current critical texts that are being used in ES in the target context(s). Locally produced texts are fewer but nonetheless present, especially in terms of course books and textbooks for practical English classes and descriptive grammar courses, where the latter are often organized on a contrastive/comparative basis. Scholarship by Bulgarian scholars in literature in English and culture studies courses (especially across universities) is rarely placed in required readings sections.

Besides establishing the dominance of the target language and context in view of communicative and reading practices for ES in Bulgaria, it is worth looking at the ways in which ‘English’ as a disciplinary academic focus and ‘Bulgarian’ as a ‘home’ context ‘interact’ with each other in practice within the educational space of ES. Figure 5 presents the summary of student responses to statements that aim to tap into some of the processes of context engagement along the lines of statements related to course design, pedagogic and analytical strategies, classroom practice and assessment procedures.

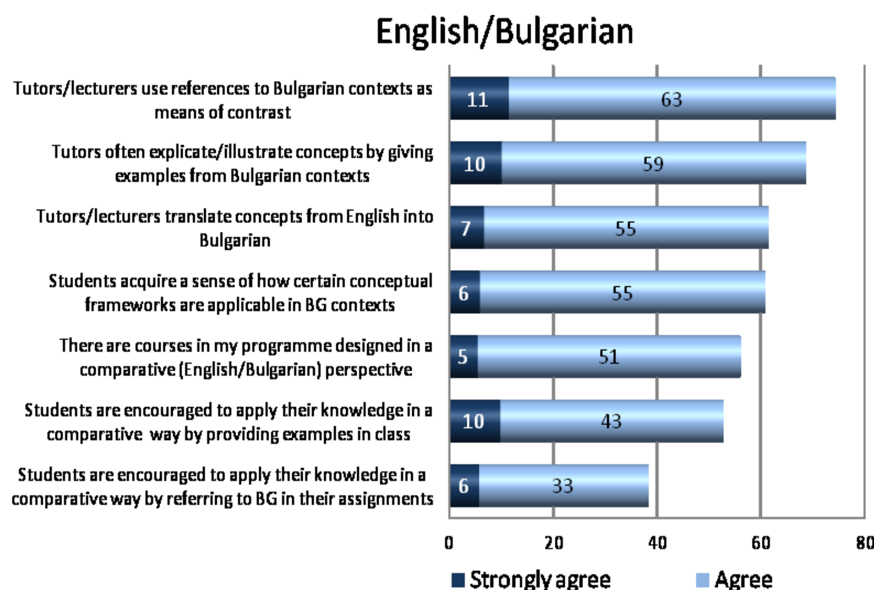


Figure 5

That the Bulgarian context is mostly used as a means of contrast and marking difference is a view which students across the three universities share. Consequently, when asked to phrase the impact of the programme in their own words, students often summarise the effect in terms of difference and contrast, as follows:

> When I first came in the university, I thought that English and my second language (French) will be useful and very important for my future. But now I see that they are not just this. I understand that the knowledge of a foreign language helps me understand a whole new culture which is far different from the Bulgarian. (BA ES+French VTU, year 1)

And

> The programme has definitely contributed to a better understanding of cultural differences between Bulgaria and English-speaking countries. (BA Applied Linguistics VTU, year 3)

The practice of translating concepts in the classroom space is slightly more widely spread (13.4% difference) among students of 'ES and other'. These are also the students who are more likely (with the same percentage difference) to state that they 'acquire a sense of how certain conceptual frameworks are applicable/function in relation to Bulgarian contexts.' Both the practice of translation and exemplification are markedly less present, in students' views, in SU. At the level of course design both ES major and 'ES and other' students have concordant views. The difference in views exceeds 10% when one compares the institutional contexts of SU and of VTU and PU. Students disagree most with the suggestion that they are able to bring the 'home' context into classroom discussion through examples and through application of their knowledge in a comparative perspective. With regard to the latter, disagreement is strongest among ES majors. In other words, while the application of 'home' context can be seen in tutor-led learning, students are not encouraged to apply the home context themselves. Further, the 'home' context is also markedly less pronounced within the institutionally documented mode of knowledge assessment—in assignments, exam papers, etc. which are the formal markers of student progress. Thus, there is a tension between student and tutor roles in actively engaging the 'home' context in academic discussion whereby tutors lead and control it. Also, there appears to be a clear demarcation

that knowledge and progress in pursuing ES pertain mostly to knowledge of and progress in the target context(s). A further discussion of some relevant points in view of this section can be found in Gupta (2010, 328-343).

6. Career goals and prospects

For over half of the ES students at the three universities, the programme they are in was their first choice. 57% have stated that it was true for them and 61% disagreed with the statement 'I wanted to study something else but didn't get admitted to the other programme.' Further, positive students' responses are evenly distributed between 'I plan to apply for a further degree in the same subject area' vs. 'a different subject area' which suggests that the ES degree offers both potential for continuous interest and motivation and opens scope for a further MA in a different area. Some ES BA graduates pursue MAs in management, business administration, economics or media studies. These choices are linked to students' views on how an ES degree is perceived in terms of professional development and employment opportunities.

The majority of students agreed with the statement, 'My degree in English definitely improves my employment prospects'. In the first year of study this conviction is expressed by 63%, in the second it drops by 20%, and by the fourth it goes up to 60%. That view is more firmly maintained by ES MA students—82%.

Institutional documents naturally firmly assert such a conviction, usually in similar terms, and students may well be influenced by them. SU's *English Philology Programme Profile* in 2008, for example, states that 'upon completion BA ES students can become:

> teachers in comprehensive and specialized secondary schools; teachers in colleges and universities; specialists in various departments of government administration, dealing with Bulgaria-UK and Bulgaria-USA relations; translators / interpreters from and into English, as well as of specialized issues and literature with regard to the UK and USA from other languages; journalists working with issues related to the UK and USA; consultants in the field of literature, language, culture, religions, historical and contemporary developments in the UK and the USA; experts and consultants in our [Bulgarian] and foreign companies, in the private or state tourist sector, in publishing, in libraries, etc.

The career paths that students themselves envisage are mapped in Figures 6 and 7 in terms of goals and desired prospective employers, both at the point of entry into university and at the moment of responding to the questionnaire.

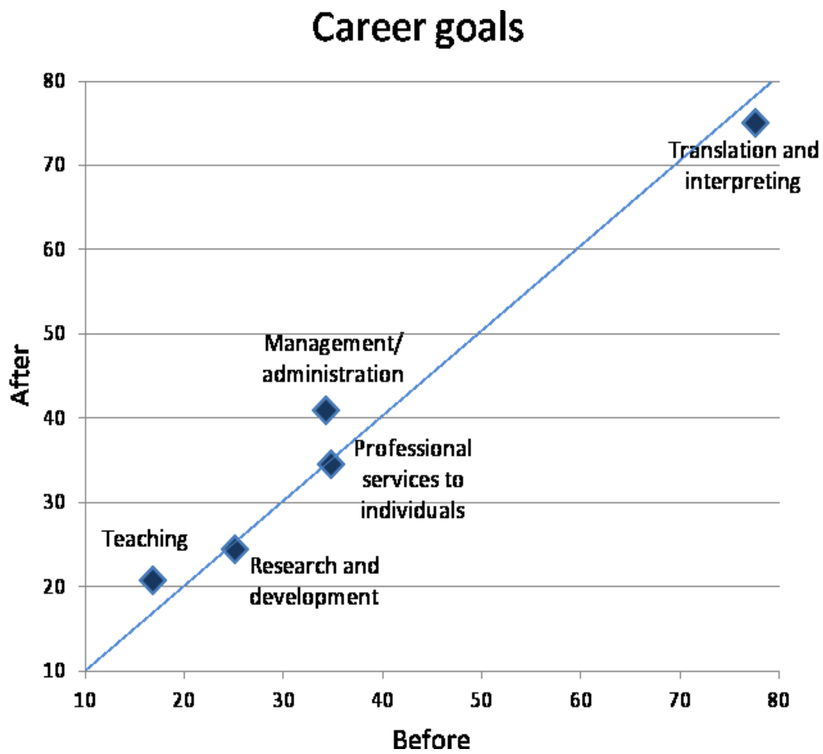


Figure 6

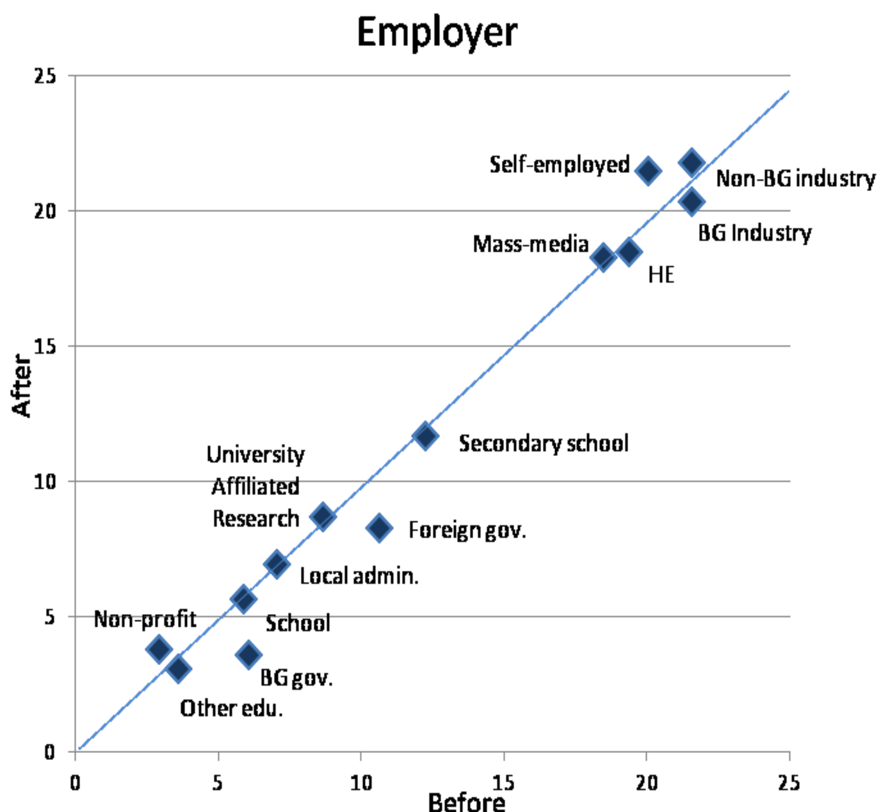


Figure 7

In terms of goals, professions of translating and interpreting score highest, though there is a slight lowering of rating with progress through the programme. Translation is historically one of the ‘strengths’ of ES degrees in terms of curriculum orientation, academic content and professional realization. It has been embedded in ES degrees as a major means of studying the target language since the introduction of English at the university level and is one of the key means of assessing students’ linguistic competence through English practice exams. The BA degree curriculum currently allocates about 100 contact academic hours to translation practice per year and there are a number of courses addressing translation within the core subject area (Translation Theory, Error

Analysis in Translation, etc). More recently, the ES major and 'ES and other' BA degrees have explicitly institutionalized translation as a professional qualification area and all universities are currently offering Translation-related programmes at the MA level. SU and VTU have introduced two delineated curricular strands for professional specialization 'Teaching' and 'Translation'. Until the 1990s only the former existed in an institutionalized form. Currently, at the MA level the translation-oriented programmes are 'Translation and Linguistics' and 'Conference Interpreting' for the three universities and 'Translation and Intercultural Communication' for PU. These developments are related to the socio-economic opening of the country since 1989 and have been most notably spurred since 2002 by Bulgaria's prospects for accession to the EU in 2007. For a number of students professional aspirations to be translators and interpreters guide their sense of the impact of ES on them. Echoing many others, one student says:

> Interpreting is an informative and enriching experience, one needs to have knowledge about a wide variety of topics. Every interpreting task has something to teach us. For those who want to work as interpreters learning never ends and I find this rather good for my personal growth. I also believe that facilitating communication between two parties is a noble mission, which is yet another aspect giving me professional satisfaction. (MA Conference Interpreting VTU)

The stable middle ground in students' views is occupied by 'professional services to individuals' and 'research and development'. The least desired aspiration among the respondents is 'teaching', which is actually one of the employment sectors for ES graduates. However, the prospect of teaching seems to gain in attractiveness as students progress further with their studies. ES programmes thus seem to compensate for the prevailing view in Bulgaria that teaching is not financially lucrative.[10] Indeed, ES students themselves note this occasionally:

[10] The average salary in the sphere of education for the first trimester of 2008 in the country was 488BGN (NSI data, 2009) for both private and state sectors. For 2007 the average annual salary in education is 414BGN in the state sector and 581BGN in the private (NSI data, 2008). State schools teacher salary at the beginning of one's career, the point where ES graduates will be, is currently fixed at 450BGN.

> The English Philology course has definitely contributed positively to my development. I now have the confidence of speaking in English freely and fluently there is also a deeper understanding of the teachers' professions as well as students' responsibilities. (BA ES PU, year 4)

And

> I have become more responsible. I also discovered that I have a talent for teaching. (BA ES VTU, year 2)

Progress through an ES programme seems to raise career aspirations (with the highest difference) in the area of 'management / administration'. Further, among envisaged prospective employers students place higher premium on the non-Bulgarian industry sector, with Bulgarian-industry and self-employment ranking close to it. Also high on the expectation horizon is employment in the mass media and HE. The government sector, however, suffers most in terms of students' employment aspirations as the students progress through ES programmes.

7. Concluding remarks

In conclusion, it is clear that the pursuit of an ES degree attracts an increasing number and varied body of students, often involving in-country or international mobility from the region, a multi-linguistic background, and a wide-range of educational backgrounds in HS. The dominant motivation factor for pursuing an ES degree in Bulgaria appears to be the promise of professional realization. Students see in ES programmes an opportunity for professional development in a European job market, especially in translation and interpreting.

Interestingly, students also often see ES as a discipline vested with normative values along the lines of cultural diversity, respect of difference, etc. Comments in this direction include:

> My programme contributed to gaining experience communicating with a great many of different people. I enjoy cultural diversity. (BA ES PU, year 3)

> The programme […] has influenced my interest in different cultures, as in my own culture too. This programme has also helped me to learn that it's important to be able to respect the difference between mine and other cultures… (BA Bulgarian and English PU, year 3)

> My programme has broadened my horizon and has made me more aware of cultural differences. This gives me a better sense of identity, contributes to my communication skills, helps me define my interest and encourages my tolerance. (BA ES+French VTU, year 1)

Yet, more often than not, the terms in which diversity and difference are understood and expressed in students' responses reveal a sharp contrast between the context in which ES is conducted and the, so to speak, disciplinary 'text'—which seems often to assert homogenized and discrete totalities like the Bulgarian and the British nations. But that's outside the scope of this paper.

Acknowledgements

Thanks are due to the British Academy and to the Leverhulme Trust for funding the projects on English Studies in Non-Anglophone Contexts: East Europe, in the context of which this paper was written. I am very grateful to Maya Pencheva and Madeleine Danova of SU and Ludmilla Kostova of VTU for organising the logistics of the student survey in their universities and to Suman Gupta (The Open University) for his valuable comments to the initial drafts of this paper.

References

Engler, Balz, and Renate Haas, eds. 2000. *The European History of English Studies: Contributions towards the History of a Discipline.* Leicester: The English Association, for ESSE.

Georgieva, Patricia. 2002. *Higher Education in Bulgaria*. Monographs in Higher Education series, ed. L. Barrows. UNESCO.

Gupta, Suman. 2010. "Critical Practice in the English Studies Classroom: Observations in Bulgaria". *English Studies Journal*, Vol. 91, Issue 3: 328-343.

Holliday, Adrian. 2008. "Standards of English and Politics of Inclusion". *Language Teaching*, 41: 119-130.

Kayman, Martin, and Filomena Mesquita. 2005. *Report on a Survey of English Studies in Europe at the Turn of the Century*. www.essenglish.org/MAKreport.pdf.

O'Reily, Laurie M. 1998. "English Language Cultures in Bulgaria: A Linguistic Sibling Rivalry?" *World Englishes*, 17: 71-84.

Shurbanov, Alexander, and Christo Stamenov. 2000. "English Studies in Bulgaria". In *The European History of English Studies: Contributions towards the History of a Discipline* ed. Balz Engler and Renate Haas, 267-292. Leicester: The English Association, for ESSE.

Thomas, Desmond, and Maria Georgieva, eds. 2002. *Smaller Languages in the Big World; Sociolinguistics and ELT*. Plovdiv: British Council Bulgaria & Lettera.

Vesselinov, Dimitar. 2008. *Letopisna kniga na Fakulteta po klasicheski i novi filologi. (1888-1965)* [Annals of the Faculty of Classical and New Philologies]. Sofia: Sofia University Press.

Williams, Sadie. 2002. *Admission Trends in Undergraduate English: Statistics and Attitudes*. English Subject Centre Report Series; No. 1, April 2002.

Critical Discourse Analysis, An overview

Encarnacion Hidalgo Tenorio, University of Granada

Abstract
The aim of this paper is to see what Critical Discourse Analysis is. This implies scrutinising its origins, what it has meant to the academic world as a whole, how it encapsulates various trends with different theoretical backgrounds and methodological approaches, what are its limitations and its new developments. A simple practical example will show its potential.

Keywords: Critical Discourse Analysis; Foucault; Habermas; Systemic Functional Grammar; Linguistic Criticism; cognition; corpus linguistics.

1. Introduction

In this paper, I describe the heterogeneity of Critical Discourse Analysis (CDA), its power to attract and annoy, and its most exciting traits and weaknesses, which have caused debate and disagreement. More than two decades have passed from the analysis of excerpts to the study of large corpora, from allegedly interested selection to random collection of data. Its social implications encouraged its development. Leaving mere intuition aside and exploring the trace of ideology in texts other than literary ones contributed to its scientificity and helped broaden its scope.

I will pay attention to CDA as problem-oriented social research, founded in social history, semiotics and linguistics; to scholarly approaches that are also considered *critical*; to the objections raised against CDA; and to new trends trying to tackle its limitations. The question of what should be understood by *critical* is also addressed, with the aim of resolving misconceptions associated with this label. It is equally important to clarify commonly used terms, including *text*, *discourse* and *context* as well as others that have a central role in CDA itself, in particular, *ideology*, *power*, *dominance*, *prejudice* and *representation*. Further, because CDA has its origins in textual and linguistic analysis, I will address the question of why one particular theory of language, Halliday's (1985) Systemic Functional Grammar (SFG), has been widely adopted by CDA researchers. SFG is not the only linguistic theory used by CDA practitioners and I shall comment on the other methods of linguistic analysis that have been applied. However,

linguistics is not the only, or even the most important influence on the development of CDA, so I shall also give attention to others from sociology, social theory and philosophy.

Among the key authors in the discipline, I will highlight Wodak, Fairclough, Kress, van Leeuwen, van Dijk and Chilton, who represent the major approaches I refer to here. These scholars share interest areas such as inequality, control, literacy and advertising. While I cannot analyse one specific text from each perspective, I provide clues as to how it would be implemented and point to tools CDA has put forward to attack its allegedly fatal malady, overinterpretation.

This area of applied linguistics, which has variously been taken to be a paradigm, a method and an analytical technique, was originally known as *Critical Language Studies* (Billig 2003). It goes by various similar names. For instance, van Dijk (2009) prefers the term *Critical Discourse Studies*, suggesting that this may help see it as a combination of theory, application and analysis. The interest of this *cross-discipline* (van Dijk 1997) lies in attending to all types of semiotic artefacts, linguistic and non-linguistic. A central aim in all the various approaches is that critical analysis raises awareness concerning the strategies used in establishing, maintaining and reproducing (a)symmetrical relations of power as enacted by means of discourse. CD analysts focus on those features contributing to the fabric of discourse in which dominant ideologies are adopted or challenged, and in which competing and contradictory ideologies coexist.

2. What is discourse?

The first obstacle faced by newcomers to the field is the various definitions of the concept of *discourse*. In a modified version of a taxonomy by Bloor and Bloor (2007: 6-7), it is possible to make the following kinds of distinction:

- *discourse-1* is the highest unit of linguistic description; phonemes, morphemes, words, phrases, clauses, sentences and texts are below;
- *discourse-2* is a sample of language usage, generally written to be spoken, that is, a speech;

- *discourse-3* refers to the communication expected in one situation context, alongside one field and register, such as the discourse of law or medicine;
- *discourse-4* is human interaction through any means, verbal and non-verbal;
- *discourse-5* is spoken interaction only;
- *discourse-6* stands for the whole communicative event.

Wodak and Meyer (2009) associate this diversity with three different trends: The German and Central European tradition, in which the term discourse draws on text linguistics; the Anglo-American tradition, in which discourse refers to written and oral texts; and the Foucauldian tradition, in which discourse is an abstract form of knowledge, understood as cognition and emotions (Jäger and Maier 2009).

Gee's (1999) pair *small-d-discourse* and *big-D-discourse* encapsulates these senses above cogently: The former refers to actual language, that is, talk and text. The latter, to the knowledge being produced and circulating in talk; to the general ways of viewing, and behaving in, the world; to the systems of thoughts, assumptions and talk patterns that dominate a particular area; and to the beliefs and actions that make up social practices. Chilton's (2004) $language_L$, $language_l$, $discourse_D$ and $discourse_d$ are very much in the same line. Cameron (2001) does not use these labels but her meaning is comparable when she distinguishes between the *linguists' discourse* (i.e. language above the sentence and language in use) and the *social theorists' discourses(s)* (i.e. practice(s) constituting objects). In a similar vein, van Dijk (1997) proposes linguistic, cognitive and socio-cultural definitions. He first argues that discourse is described at the syntactic, semantic, stylistic and rhetorical levels. Secondly, he adds that it needs to be understood in terms of the interlocutors' processes of production, reception and understanding. And, thirdly, he points to the social dimension of discourse, which he understands as a sequence of contextualised, controlled and purposeful acts accomplished in society, namely, a form of social action taking place in a context (i.e. physical setting, temporal space plus participants). Since context is mostly cognition, that is, it has to do with our knowledge of social situations and institutions, and of how to use language in them, van Dijk claims that each context controls a

specific type of discourse and each discourse depends on a specific type of context.

From Widdowson's perspective (2004), texts can be written or spoken, and must be described in linguistic terms and in terms of their intended meaning. Discourse, on the other hand, as text in context, is defined by its effect. In his words, discourse "is the pragmatic process of meaning negotiation", and text, its product (2004: 8). Co-textual relations are concerned with text and contextual relations with discourse; that is, text cohesion depends on discourse coherence.

CD analysts Fairclough and Wodak (1997: 276) refer to the following senses: Language use in speech and writing, meaning-making in the social process, and a form of social action that is "socially constitutive" and "socially shaped". The concept Fairclough finally opts for is *semiosis*, in order to include not only linguistic communication but also, for example, visual communication, as well as to generalise across the different meanings of the term discourse. Semiosis plays a part in representing the world, acting, interacting and constructing identity, and can be identified with different "perspectives of different groups of social actors" (Fairclough 2009: 164). Discourses can be appropriated or colonised, and put into practice by enacting, inculcating or materialising them. In contrast, texts are "the semiotic dimension of events" (*ibid*.), where we can find the traces of differing discourses and ideologies (Weiss and Wodak 2003).

The origin of the latter ideas can be tracked back to philosopher Michel Foucault (2002: 54), for whom discourses are "practices that systematically form the objects of which they speak". In their interpretation, Fairclough and Wodak (1997: 261) add that discourses "are partly realized in ways of using language, but partly in other ways", for example visual semiosis. Texts are the only evidence for the existence of discourses, one kind of concrete realisation of abstract forms of knowledge; at the same time, they are interactive and influenced by sociolinguistic factors. In the process of constructing themselves in society, individuals internalise discourses that comprise the core of a community of practice, in the sense that such discourses control and organise what can be talked about, how it can be talked about and by whom. Social practices are meaningful and coherent in that they conform to discourse principles. As manifestations of ideologies, discourses form individual and collective consciousness, and consciousness influences

people's actions; that is, through the repetition of ideas and statements, discourse solidifies knowledge (Jäger and Maier 2009), and reflects, shapes and enables social reality. Furthermore, it can be defined by the activities participants engage in, and the power enacted and reproduced through them; thus, we can speak about feminist or nationalist discourse, doctor-patient or classroom discourse, the discourse of pity, whiteness or science, or hegemonic and resistant discourses. To Foucault's definition, van Leeuwen (2009: 144) adds that discourse involves social cognitions "that serve the interests of particular historical and/or social contexts", represent social practices in text, and transform or recontextualise them. As will be seen later, van Dijk places considerable emphasis on this notion.

3. What makes DA critical?

CDA is naturally embedded within Critical Theory, a paradigm developed in the last three decades whose critical impetus originates in the Frankfurt School, especially Habermas. As Wodak and Meyer (2009: 6) recall, in 1937 Horkheimer urged social theory to critique and change society, which meant to improve its understanding by integrating social sciences, to show how social phenomena are interconnected, to produce knowledge that helps social actors emancipate themselves from domination through self-reflection, and to describe, explain and eradicate delusion, by revealing structures of power and ideologies behind discourse, that is, by making visible causes that are hidden. The scope of CDA is not only language-based. Its critical perspective attracts scholars from various disciplines, as well as activists. Their concern lies with unveiling patterned mechanisms of the reproduction of power asymmetries. Anthropology, linguistics, philosophy and communication studies, among others, may share this inclination.

From its inception, CDA was a discipline designed to question the status quo, by detecting, analysing, and also resisting and counteracting enactments of power abuse as transmitted in private and public discourses. For some, to be critical might imply to be judgemental. However, this is not the case here, because, as Jäger and Maier (2009: 36) state, this kind of critique "does not make claims to absolute truth". CDA is understood to be critical in a number of different ways: Its explicit and unapologetic attitude as far as values and criteria are

concerned (van Leeuwen 2006); its commitment to the analysis of social wrongs such as prejudice, or unequal access to power, privileges, and material and symbolic resources (Fairclough 2009); its interest in discerning which prevailing hegemonic social practices have caused such social wrongs, and in developing methods that can be applied to their study (Bloor and Bloor 2007). All this makes CDA an example of research aiming for social intervention. Fairclough and Wodak (1997) add that a critical reading goes beyond hermeneutics. In their view, CDA aims at demystifying texts shaped ideologically by relations of power; it focuses on the opaque relationship between discourse and societal structure; and it does so through open interpretation and explanation, by relying on systematic scientific procedures, that is, by achieving distance from the data and setting them in context. Self-reflection concerning the research process is a must. In sum, CDA seeks to expose the manipulative nature of discursive practices, and improve communication and well-being by removing the barriers of assumed beliefs legitimised through discourse.

4. The origins of CDA

The philosophical and linguistic bases on which CDA is grounded are certain branches of social theory and earlier discourse analysis, text linguistics and interactional sociolinguistics. Certain proponents of CDA are influenced by Marx's critique of the capitalist exploitation of the working class, his historical dialectical method, his definition of ideology as the superstructure of civilisation (Marx and Engels 1845/2001), and his notion of language as "product, producer, and reproducer of social consciousness" (Fairclough and Graham 2002: 201). Some also draw on Althusser's (1969/1971) conception of *interpellation*, which describes the way an individual can be aware of themselves as a constructed subject within discourse on their becoming part of someone's utterances. Likewise, Gramscian *hegemony* (1971) influences a number of CDA scholars. It formulates the idea that power can be exercised and domination achieved not only through repressive coercion, oppression and exploitation, but also through the persuasive potential of discourse, which leads to consensus and complicity.

Habermas (1981) is frequently cited by CDA writers. His key contribution in the theory of communicative action is the notion of

validity claims, which, according to him, are universally presupposed in all discourse. He further maintains that language can be used either strategically or in a manner oriented to understanding. In the latter, validity claims can be challenged and defended in a communication situation that is free from coercion, is only based on rational argument, and permits access to all who are affected by the discourse. These characteristics are absent from the strategic use of language; it is to challenging the strategic use of language that CDA pays attention. Foucault (1972), in contrast to Marx and Habermas, thinks that consciousness determines the social production process. Despite contesting the existence of an autonomous subject, he believes in the individual's involvement in the practical realisation of power relations. Discourses are produced by all individuals, then, especially those who have the right to use all resources (Jäger and Maier 2009).

In the late 1970's, the University of East Anglia nursed a new trend of analysis, as linguists and literary theorists were interested in linguistic choice in literature (see Fowler 1986). Later on, they would focus on other texts of relevance in the public sphere, especially the mass media. This did not mean only a terminological change (i.e. from *linguistic criticism* to *critical linguistics*). The new label, which is sometimes taken as synonymous with CDA, implied a new attitude in academe: The scholar's commitment against social injustice. The East Anglia School proposed Hallidayian linguistics for the analysis of news texts (Hodge and Kress 1993). Language as social semiotic, the three metafunctions, and transitivity and modality became staples in this new discipline. Chomsky's grammar (1957) was also appropriated since one of its main concerns is describing the implications of syntactic transformations: Passivisation and nominalisation have been the focal point of many a CDA work.

5. Examination of approaches to CDA

Notwithstanding obvious similarities, especially as regards agenda and scope, proponents of schools of CDA differ according to theoretical foundations or methodology. Some tend towards deduction and others proceed inductively. The former base their explanations on a few examples; the latter scrutinise a larger collection of data; without doubt, this can be more time-consuming but absolutely reliable and unbiased.

All in all, they all generally attend to categories such as tense, deixis, metaphor, attributes or argumentative topoi.

Fairclough's Dialectical-Relational Approach to CDA is an essentially Marxist framework, anchored in his (1989, 1995) research on language, ideology and power, where we find a very influential terminology, including *dominance*, *resistance*, *hybridisation of discursive practices*, *technologisation of discourse* and *conversationalisation of public discourse*. As Wodak and Meyer (2009) explain, Fairclough highlights the semiotic reflection of social conflict in discourses, which translates into his interest in social processes (i.e. social structures, practices and events). A pragmatic side of this approach is his support for *critical language awareness*, which to him is essential in language education (Fairclough 2007). Fairclough describes the following procedure: The scholar looks at one social problem with a potential semiotic dimension. This dimension is analysed by identifying its styles (or semiotic ways of being), genres (or semiotic ways of acting and interacting) and discourses (or semiotic ways of construing the world). Later, the differences between styles, genres and discourses are identified. Next, the researcher studies the processes by means of which the colonisation of dominant styles, genres and discourses is resisted. The focus then shifts to the structural analysis of the context, and the analysis of agents, tense, transitivity, modality, visual images or body language. Eventually, interdiscursivity is dealt with. Regardless of the apparent neatness of this methodology, Fairclough (2009) denies there is one single way of analysing any problem. Interestingly, he believes that, after selecting one research topic, the scholar constructs their object of research by theorising it. Its transdisciplinarity is one of the outstanding strengths of one approach where researchers may prefer (detailed but not always too rigorous) analysis of few data, selected, sometimes, by using somewhat unclear methodology and, to some extent, opaque style encouraging less critical thinking than one might expect.

Van Dijk's Socio-Cognitive Discourse Analysis is an approach characterised by the interaction between cognition, discourse and society. It began in formal text linguistics and subsequently incorporated elements of the standard psychological model of memory, together with the idea of *frame* taken from cognitive science. A large part of van Dijk's practical investigation deals with stereotypes, the reproduction of ethnic prejudice, and power abuse by elites and resistance by dominated groups.

Van Dijk also emphasises the control of discourse dimensions as a means to gain access to power. A further element in his account of discourse production and comprehension is the *K-device*, which is shorthand for personal, interpersonal, group, institutional, national and cultural knowledge (van Dijk 2005). Cognition, realised in collective mental models as a result of consensus, is the interface between societal and discourse structures (van Dijk 2009). While societal structures influence discursive interaction, in the latter the former are said to be "enacted, instituted, legitimated, confirmed or challenged by text and talk" (Fairclough and Wodak 1997: 266). Van Dijk (2009) believes CDA needs a model of context such as Moscovici's (2000) social representation theory: One individual's cognition is informed by dynamic constructs known as *social representations*, that is, the concepts, values, norms and images shared in a social group, and activated and maintained in discourse. He advocates the analysis of semantic macrostructures, local meanings, formal structures, global and local discourse forms, specific linguistic realisations and context. The aspects he focuses on are coherence, lexical and topic selection, rhetorical figures, speech acts, propositional structures, implications, hesitation and turn-taking control. Despite its power, in this approach, it is my belief that intersubjective agreement between scholars is not fully guaranteed by a slightly deficient explanation of how to apply some of the rules identified by van Dijk in discourse practice; thus, method and conclusions are open to multiple interpretation.

The Discourse-Historical Approach (DHA) (Wodak and colleagues) attempts, *inter alia*, to describe those cases where language and other semiotic practices are used by those in power to maintain domination (Reisigl and Wodak 2009). Initially, DHA was concerned with prejudiced utterances in anti-Semitic discourse. Recent developments include the discursive construction of national sameness and the social exclusion of out-groups through the discourses of difference, and the reconstruction of the past through sanitised narratives. The general approach reflects sociolinguistics and ethnography; it also gives an important place to Habermas's notion of the public sphere and to strategic communicative action as opposed to ideal communication oriented to understanding. Its central tenet is the importance of bringing together the textual and contextual levels of analysis. The model of context used in DHA invokes historical knowledge understood in terms

of four layers: (a) the linguistic co-text, (b) the intertextual and interdiscursive level, (c) the extralinguistic level, and (d) the socio-political and historical level (Wodak and Meyer 2009). The interconnection between various texts and discourses leads directly to the notions of *de-contextualisation* and *recontextualisation*, processes in which elements typical of a particular context can be taken out of it and inserted into a new context with which it has not been conventionally associated. DHA has further produced a series of analytical and descriptive tools, drawing on linguistic models and argumentation theory. In particular, DHA lists six strategies for identifying ideological positioning (i.e. *nomination*, *predication*, *argumentation*, *perspectivisation*, *intensification* and *mitigation*) which are analysed as part of a larger process that includes also the characterisation of the contents of a discourse, linguistic means of expression and context-dependent linguistic realisations of stereotypes. One of the strengths of DHA is the emphasis on the combination of observation, theory and method, and the continuum between application and theoretical models. Its historical, political and sociological analyses are also, in my view, an important part of its methodology, especially in relation to systems of genres, although the lack of a fully systematic procedure in this regard is one of its weaknesses.

The Duisburg School is heavily influenced by Foucault's work. A particularly strong underlying conviction is that it is discourse that makes subjects (Jäger and Meier 2009). In other words, an individual's sense of who they are arises from their imbrication in systems of historically contingent meanings communicated by institutionalised patterns of behaving, thinking and speaking. This kind of framework, sometimes referred to as Dispositive Analysis, also draws on social constructivism (Laclau 1980) and activity theory (Leont'ev 1978), and claims that social selves are constituted in a semiotic network that includes not only linguistic mediation of various kinds but also architectural arrangements, legal practices, customs, rituals, modes of moral thought, social institutions and so forth. Their notion of discourse is built upon "an institutionalized way of talking that regulates and reinforces action and thereby exerts power" (Link 1983: 60). While Foucault's approach is relatively vague as regards discourse in its linguistic manifestation, the Duisburg approach pays attention to metaphors, references, style, implied meaning, argumentation strategies, the sources of knowledge,

and agentive structures and symbols. Like DHA, the Duisburg approach advocates an analytic procedure. After selection of a particular subject matter, analysis is focussed on one domain, such as the media. This is followed by the structural analysis of one discourse strand ("what is said and sayable at a particular point in time", Jäger and Maier 2009: 46) and of typical discourse fragments, that is, the different topics each text deals with. Linguistic-discursive practices are explored through the analysis of texts; non-linguistic discursive practices, through ethnographical methods; and materialisations, through multimodal analysis and artefact analysis. The interest of an approach like this, concerned, *inter alia*, with everyday racism, patriarchy in immigration discourse or the discourse of the right, may be diluted, however, behind the imbalance between its complex theoretical apparatus and what may seem to look as only content-based analysis.

There is a prominent strand of CDA that advocates the use of Halliday's SFG. This is the framework of linguistic description used by Fairclough, as it was also by Fowler et al. (1979) and Hodge and Kress (1988). In those studies that make use of SFG, different linguistic descriptions of the same piece of reality are claimed to stand for different constructions of that reality. Thus one and the same historically occurring event can be described as a *riot*, a *demonstration* or a *protest*; and social actors can be presented as agents or victims by selection of grammatical coding. More generally in this approach, text types represent social practices (i.e. regulated ways of doing things), which involve participants, actions, performance modes, presentation styles, times and locations, resources and eligibility conditions. Theo van Leeuwen (2009) has developed SFG's formal framework for the classification of the semiotic system of social actor types and for the classification of the different ways in which social actors can be linguistically represented. According to this author, *deletion*, *substitution*, *rearrangement* and *addition* are the transformations that elements of a social practice undergo through discourse. *Recontextualisations* add the *what for* and the *why* (or *why not*) of a social practice. In discourse, van Leeuwen hypothesises, social actors can be included or excluded; actions can be represented dynamically or not, as if there were no human agents or the opposite; we can generalise them, or make them stand for specific references, abstractions, symbols; as for practices or their elements, these can be set in a context, or

reallocated. By making explicit the ranges of ways in which texts represent social actors, their actions and purposes, van Leeuwen seeks to analyse how specific discourses legitimise some of these actors and practices and intentions rather than others. His concern with an overall comprehensive analysis of complex semiotic phenomena (the language of images included), which is not yet complete, by way of linking various disciplines may be both one of its strengths and one of its drawbacks at the same time.

Finally, I will mention Mediated Discourse Analysis (Scollon and Scollon 2005), which highlights the role of ethnography and semiotics. As with the DHA, this emphasises the diachronic dimension, and texts are viewed as situated discourse (Scollon 2003). Further, Scollon and Scollon revisit Bourdieu's notion of *habitus* (a system of permanent, identical characteristics, which, by integrating past experiences, mediates perceptions, judgments and actions), and develop it by appropriating Nishida's (1958) concepts of *action-intuition* and the *historical body* (a combination of the social and the psychological). Scollon and Scollon (2005) claim that discourses are always present at any moment. They depict the individual as an actor "embodied in the society of various social groups" (Scollon 2003: 172). Subsequently, one of their goals is to find the link between individual action and public discourse, so as to achieve an understanding of how we internalise the social world and how, through action-intuition, the historical body of a social actor is externalised. I must admit that their optimistic standpoint as to the possibility of researchers acting in the world in order to make a difference in the actual world and their concern with problems in intercultural communication is more than exciting; however, some relative lack of very detailed concrete methodological guidelines may deter scholars.

6. Critiques of CDA

The merits and demerits of CDA research have been the object of a certain amount of critique. The problems that have been picked up concern context, cognition, partiality and the linguistic model employed. Most critics do not call into question the existence or epistemological relevance of CDA, perhaps with the exception of Widdowson and Chilton, but are aware of its shortcomings: Its theoretical foundations are

quite tangled in many cases, and the use of concepts and categories may seem to be inconsistent, which does not encourage the production of a systematic theory. Eclecticism, if lacking in justification, can be a source of contradiction.

Although Widdowson (2004) does not oppose CDA's cause, he casts doubts on its modes of analysis. He cannot agree with the way CDA uses SFG, where meaning is understood as a condition of texts, taken from them, not put into them. He points out that there is a gap between addresser meaning and addressee interpretation of this meaning, on the grounds that the perlocutionary effect is not a feature of texts but a function of discourse, in which the addressee's assumptions are shaped by their knowledge and beliefs. Hallidayan grammar offers interesting devices for the description of semantic meaning (or *signification*); however, to Widdowson's eye, this is defective because it fixes on isolated sentences instead of utterances. He adds that, in this framework, the concept of context is as essential as it is indeterminate. If meanings are understood as properties of the interaction between words and contexts, interpretation is an imprecise process. In a nutshell, Widdowson regards some CDA approaches as examples of the *functional fallacy*, by which he means the idea that pragmatic meaning (or *significance*) may be produced directly by signification. He maintains that abstracting sentences from their contexts and choosing examples relevant to the ongoing research does not lead CD analysts to produce analysis in the strict sense of the term: Pretexts influence how to approach texts and the type of discourses derived from them. To him, CDA is a biased, unprincipled, conventional, decontextualised cherry-picking of linguistic features, closer to impressionistic commentary, which supports interpretation and yields simplistic findings. Widdowson argues that CDA is critical in the sense that it has moral appeal, socio-political justification and liberal ideological positioning. And he accepts that the CD analyst observes issues that are relevant in areas other than the scholarly world and addresses how control is exercised though language. However, he strongly urges that CDA should adopt a critical attitude towards its own purposes, methods and practices, be explicit in methodological procedures, which must be replicable, and apply consistent principles and systematic linguistic theory. In all, CDA should comprise systematic analysis of entire texts, co-texts and contextual relations.

Although Chilton has contributed many papers on discourse that have a social-critical intention, his (2005) paper is critical of CDA, maintaining that what CDA lacks is a cognitive theory of language that could show how discourse affects social cognition and vice versa. Cognitive frame theory, conceptual metaphor theory and blending theory can explain better than traditional CDA approaches (including SFG) why phenomena such as racism and prejudiced thought can occur. In addition, Chilton hints that the work by critical analysts is based on no particular scientific programme and may simply reflect a universal ability in non-expert cheating-detection, going so far as to suggest that, taking into account its audience and scope, CDA may be of limited social import.

Billig (2003) thinks that CDA has the crucial characteristics of a critical approach: The claim to be critical of the current social order and of the approaches which do not critique the current social order's domination patterns. He also recognizes the importance of CDA's claim that non-critical approaches prevail in the academic world, resulting in keeping existing power relations unchallenged; and he supports CDA's insistence that an interdisciplinary approach is needed. Nevertheless, he recognises that, through naturalisation and institutionalisation, a critical approach may itself become a dominant discourse and, consequently, a dominant discipline, with the shortcomings of the approaches it criticises.

Other critical voices include those of Martin and Blommaert. Martin's approach (1992) critiques CDA because of its inability to put into practice its social-based ambitions, so that in the end it observes social phenomena we mostly dislike, producing very persuasive materials on why they are offensive, but failing to suggest practical action. Instead he proposes Positive Discourse Analysis (Martin 2000). He argues that, for the purpose of social change, scholars should analyse not only texts they find objectionable but also texts they find admirable and motivating. As for Blommaert (2005), he deals with the discursive production of inequality and the need for self-critique in CDA. Drawing on linguistic ethnography, he refers to a particular kind of parochialism. He considers that one of the shortcomings of this field is that generally most work produced to date pays attention to texts of relevance in the West since, as a rule, CDA is not applied to societies other than the First World. He further criticises CDA on the grounds that it has demarcated boundaries around itself as a field: Much discourse analysis can be

critical, says Blommaert, without subscribing to the underlying assumptions and methods of CDA.

7. New directions in CDA

New formats and materials involved in communication have encouraged new avenues of CDA research, analysis of the multimodal properties of texts being one significant innovation. Other trends include developing the connection between CDA and cognitive linguistics, analysing gender semiosis, and bringing corpus linguistics into CDA.

In recent decades, attention has been drawn to just one communicative mode, verbal language. However, music and pictures are the basis for the meaning-making process in the audio and visual modes; the size, colour and frame of a news report are important to guide the addressee's engagement with the text; the distribution of images and the timing of news are significant in TV and the press; body posture, gestures and the use of space help construct our text and talk (Hodge and Kress 1988; Kress and van Leeuwen 2006).

Having underlined CD analysts' tendency not to go beyond the verbal detail and the sentence level, so that it is hard to explain what happens in the mind when reception takes place, Chilton (2005) stresses the necessity to look at how we construct knowledge, that is, which procedures are necessary for individuals to share views and build mental models. This implies, for instance, his concern with the role of cognitive frames in facilitating human processing of information, and with the cognitive strategies deployed in order to infer other people's intentions. In the same vein, O'Halloran (2003) had already concluded that there must be a link between sentential structures and mental representations, all of which seem to be controlled by discourse rules, an idea that needs to be reassessed in the light of connectionism (McClelland and Kawamoto 1986). In his model of reading for gist, O'Halloran discusses the extent to which lay readers attend to absences from a news text (especially as far as causal relations are concerned), and considers how to avoid mystification in the interpretation stage (Hart 2010).

Feminist Critical Discourse Analysis aims to analyse the relationship between gender and language, which mostly means examining enactments of power by men and women in the public domain (see papers in Lazar 2005). Sexism, victimisation, emancipation and the

construction of identity are key issues in question. Feminist research has been prolific in investigating the role of gender in politics, the media, the workplace and the classroom context. The evolution from the *deficit*, *difference* and *dominance* approaches to a shift to discourse (see Litosseliti 2006) has been crucial. In current research, attention is not directed towards whether men and women speak differently, or whether the language of females is a deviation from the male norm (traditionally, the excuse to explain male superiority at some levels), but to comprehend gender as a dynamic construct (Cameron and Kulick 2003; Sunderland 2004).

Finally, it has been argued that the use of corpus linguistics techniques in CDA may help to avoid or reduce researcher bias. Quantitative computer-aided corpus approaches can address large data sets, and the focus of analysis can take into account collocations, keyness, semantic preference and semantic prosody. What appears to be obvious and is taken-for-granted are checked against the data at the same time that unexpected findings arise in the procedures of description and analysis (Mautner 2009). Patterns of preferred and dispreferred lexis and structures (Baker 2006) facilitate detecting of the ideologies of hegemonic discourses associated with particular texts. The limitations of corpus approaches to CDA are that they work with very little context, may hinder close reading and can help us learn only about the verbal domain. Corpus-informed CDA can give the impression that it is a mechanistic or positivist approach. O'Halloran and Coffin's (2004) discussion of over- and under-interpretation counters this view, however.

8. An example of CDA

Analysis in CDA can be top-down, where analysts begin with their understanding of the content; or bottom-up, where the starting point is the linguistic detail. In practice, however, some combination of both is in play. The analyst looks for what is encoded in sentences (i.e. *signification*) and its interaction with context (i.e. *significance*). In this respect, the analyst is merely doing what an ordinary reader would normally do, but with more conscious attention to processes of comprehension, their possible effects, and their relationship to a wider background knowledge than the ordinary reader may assume to be relevant. Depending on the approach, various linguistic devices are paid

attention to with a view to understanding their contribution to some potential strategically intended meaning that may be linked with ideology. The clues found are interpreted, and some explanation is expected about them and their implications. For example, many CDA writers use Hallidayan linguistics, focusing on the ideational and interpersonal metafunctions of language, specific deployment of which may, it is claimed, reveal author-to-audience power relationship or be connected (by further interpretation) with the author's positioning. Here I show CDA at work by studying the British Press discursive construction of the killing of Saddam Hussein's sons.

The analysis of how people speak or write about crime is interesting because, by describing the verbal construction of perpetrators, we can decipher the discourse of wrongdoing, and gain an insight into the conception of the world held by the speaker or writer. Thus, the language of the journalists who presented Saddam Hussein's sons as agents of evil may also tell us much about the journalists themselves and the newspapers they work for. To illustrate my point, I have studied all the news articles published in the UK on 15 December 2003, both on paper and on the Internet (when available), one day after Saddam Hussein was captured by US troops. The newspaper subcorpus consisted of 56 items, including all sorts of articles and other materials (see Table 1). The web newspaper subcorpus was a 17,492 word collection of 16 articles taken from the *Daily Mail*, 7 from *The Guardian* and 3 from *The Independent*.

Table 1. Newspaper corpus

	Daily Mirror	***Daily Mail***	***The Independent***	***The Guardian***	***TOTAL***
Authored articles	10	8	13	8	**39**
Un-authored articles		7			**7**
Editorials	1	1	1	1	**4**
Comments				2	**2**
Speeches			2		**2**
Extracts of speeches				2	**2**
TOTAL	**11**	**16**	**16**	**13**	**56**

A full analysis might focus on metaphor, modality, transitivity and lexical selection, among other features. Given that I want to examine the

construction of opinion, I look at transitivity, which in SFG is treated as part of the ideational metafunction (see Halliday 1985), and provides a powerful tool for the analysis of how meaning is embodied in the clause. In particular, the category of transitivity deals with the linguistic means we have for expressing, through language, our experience of the world around us. Its components are *participants*, the *processes* these are involved in and the *circumstances* in which processes take place. There is some tendency for clause constituents to be expected to have a specific semantic role: Subjects are generally agents; direct objects, patients; indirect objects can be recipient or beneficiary; subject complements, attributes, and so on. As for the types of processes, Halliday proposes six: Material (action, event), behavioural, mental (perception, affect, cognition), verbal, relational (attribution, identification) and existential. In spite of this looking like a precise taxonomy, it can still be difficult to identify sometimes which is which when there seems to be overlapping (see Table 2).

Table 2. Processes in the corpus

Process	**Positive %**		**Negative %**		**Neutral %**	
	They	***He***	***They***	***He***	***They***	***He***
Behavioural	0.00	0.18	0.52	0.35	0.17	0.53
Existential	0.17	0.00	0.00	0.00	0.17	0.18
Material Action	6.22	5.27	17.79	30.40	21.59	19.33
Material Event	1.38	0.18	1.90	2.99	2.59	1.58
Mental affection	0.69	0.53	2.59	1.93	2.25	0.35
Mental cognition	0.00	0.00	0.35	0.00	2.94	0.88
Mental perception	0.00	0.00	0.00	0.00	0.35	0.35
Relational attribution	3.28	4.22	6.56	16.17	1.73	5.27

Relational identification	0.00	0.70	0.00	1.05	0.00	1.93
Relational possession	0.00	0.18	0.00	0.53	0.17	0.35
Verbal saying	2.07	0.53	3.28	1.41	21.24	2.64

In my corpus, the *West* (the so-called Coalition) is in the main involved in material (action) and verbal (saying) processes (see Table 3 below). Since they are identified with many of the values and beliefs supported by the privileged voice behind these newspapers, they are *sayers* (or, simply, speakers) whose voice is heard profusely in several ways. They are also represented as agents, or doers, who carry out generally neutral or more positive actions than their "High Value Target". The *others* of the West, the Iraqis, are scared, dead or victimised sufferers; very few are said to be actors or heard articulating their thoughts. That is, they are those affected by the processes and as such are often encoded as grammatical objects. As for Saddam, who is described as perverse, is accordingly demonised, dehumanised and objectified—this representation of him arising in large part also as a result of the journalist's selection of verbs expressing the processes in which he is a participant. So far, none of this is surprising. The position of the newspapers is reflected in their negative depiction of the dictator and the marginalisation of all voices other than the non-problematic West. The ideology of the media is equally transparent in the next set of examples, which concern the dictator's sons.

Table 3. Semantic roles in the corpus

	Participant	Positive %		Negative %		Neutral %	
		They	***He***	***They***	***He***	***They***	***He***
Active voice	Actor	5.28	3.16	12.85	20.91	21.13	8.08
	Behaver	0.00	0.18	0.35	0.18	0.18	0.53
	Beneficiary	0.35	0.00	1.06	0.70	2.82	0.53
	Carrier	3.17	3.87	6.51	15.82	1.76	4.39
	Cause	0.00	0.00	0.53	0.18	0.00	0.18
	Existent	0.00	0.00	0.00	0.00	0.35	0.18
	Goal/phenomenon	0.00	0.35	4.58	5.98	0.00	6.50
	Identified	0.00	1.05	0.00	0.70	0.00	1.76
	Inducer	0.00	0.00	0.00	0.18	0.00	0.18
	Possessor	0.00	0.00	0.00	0.00	0.18	0.00
	Recipient	0.00	0.00	0.00	0.18	0.88	0.18
	Sayer	2.46	0.53	2.46	1.05	20.95	2.46
	Senser	0.53	0.53	0.88	0.35	5.46	1.41
Passive Voice	Actor	0.35	0.00	1.94	1.93	0.70	0.53
	Carrier	0.00	0.35	0.18	0.88	0.00	0.88
	Cause	0.00	0.00	0.18	0.35	0.00	0.00
	Goal/phenomenon	0.00	1.93	0.35	4.57	0.70	5.45
	Identified	0.00	0.00	0.00	0.00	0.00	0.18
	Inducer	0.00	0.00	0.00	0.53	0.00	0.00
	Recipient	0.00	0.00	0.00	0.00	0.18	0.00
	Sayer	0.00	0.00	0.18	0.18	0.00	0.00
	Senser	0.00	0.00	0.18	0.00	0.35	0.00

Nominalisation is a powerful means of abstraction that may, in certain contexts, disguise, or make less salient, agents, responsibility and circumstances. For example, the *Guardian* text speaks about "the death of Uday and Qusay" (the end of their lives without indicating the cause) and "the killing in July of Saddam's two sons" (an unmentioned someone or something caused them to stop living). In its editorial, we also read: "Uday and Qusay … perish[ed] in a murderous blizzard of bullets". Here the instrumental cause of their death is made explicit (i.e. the large number of something as annoying and unpleasant as bullets), but there is no direct reference to any trace of the animate agent pulling the trigger.

The *Daily Mirror* reports Saddam's sons as being "killed in a battle", and adds that "Uday and Qusay … died in July guns blazing, after waging a four-hour-battle with American forces". Readers see that there has been a change of state in the patient argument, something natural and unavoidable. Although we get information about how this took place, the agent remains unknown. This is explained in terms of the circumstances of this event (expressed, for instance, in the clause "after waging …"). The human agency is thus not stated explicitly and the instrumental causation is expressed indirectly. In some way, such indirectness may appear to justify what happened: There was a long fight between opposing groups, one of which is vaguely present; it was in July; bullets were fired quickly and continuously; and both perished. In the last scenario, there is something else: Uday and Qusay were dead because they started this battle and continued it over a period of time. The American forces only fulfil the role of *goal*, the Hallidayan term for the participant affected by (not the agent of) some material process.

The *Daily Mail* tells that they were "killed after opening fire on an overwhelming US force which surrounded them in July", "in a fire-fight against US forces", and that they "died in a shoot-out with US troops". In the first case, the journalist shows the episode as if Saddam's sons had to be affected necessarily by an agentless action, which eventually caused their death because it was them who started shooting at the Americans. In the second case, we only know that someone or something caused their death in a battle that involved the use of guns rather than bombs or any other weapon. US forces are merely a circumstance in the whole event. As for the third, although the supposed agents of their demise are mentioned indirectly as minor participants, the fact that the journalist

prefers the lexical item *shoot-out* helps imagine a fight in which several people shot at each other. In other words, the idea of reciprocity is presented here, in contrast to the previous examples. Furthermore, the noun *troops* indicates that there were other people involved in this event, that the US force was not an abstraction, and that there was some human agency behind it.

As shown above, the point I have been considering has to do with avoidance of agency. *To die* is a material (event) process. It happens to animate beings, and it happens accidentally or not. If we use this lexeme, we focus on the affected participant, not on the cause (e.g. an illness, misadventure), the agent (e.g. living thing acting deliberately or not), the instrument (e.g. gun, knife, rope, poison) or the circumstances (e.g. in water, in action) concurring to provoke it. If we prefer the verb *to kill*, we refer to a material (action) process that involves an affected goal and an actor that causes the object's death. As for *to perish*, it means that the patient dies as a result of an accident or very harsh conditions; these very hard conditions were the bullets shot by the US troops, against which they fought to the death, an idea that is celebrated by the newspaper if compared with Saddam Hussein's apparent passivity.

The images resulting from the newspapers' construction of the scene are the following: We know what happened to both (they died), when (July), where and how (in a battle, in a fire-fight, guns were blazing), and why (they waged a battle with American forces, they opened fire on a US force). Sometimes the actor or cause of their deaths is referred to indirectly as another circumstantial element (waging a four-hour-battle with American forces, in a fire-fight with US forces) or as the affected object of their actions (opening fire on a US force). Thus, these journalists depict Uday and Qusay as responsible for their own deaths, since they started the attack. Linguistically speaking, the US force is a contingent participant or suffers Saddam's sons' actions, but it never acts. This may have been an unconscious choice. Nevertheless, in one case the writer adds clarifying contextual information: The US force was overwhelming. The journalist emphasises that an amount or quantity is much greater than other amounts or quantities. The situation is perceived as one in which one side is at a disadvantage. Therefore, the affected goal changes as well as the notion of who is the patient.

This practical exercise develops some of the main tenets of CDA: Aim, type of data and analytical approach. Methodologically speaking, I

have studied one aspect such as transitivity which is very relevant to understanding people's positioning. This is one way to get to know which your identity is, and how you see the world and how you perceive others. To me, the application of this perspective is useful because it deals with ideologically loaded material systematically. It is made even more robust by the use of the complementary tool of corpus linguistics, which has allowed me to scrutinise frequency patterns and made it easier to discover the traces of what is (and is not) essential in the text, in order to comprehend better speakers' and writers' intentions. If the latter avoid some structures, this may be an indication of their fears and their taboo areas. If, on the contrary, they prefer some others, this may show they are worried about or even obsessed with those problems around which their discourse revolves. Despite the systematicity of this type of approach, I must agree, however, this is not "a mechanical procedure which automatically yields 'objective' interpretation" (Fowler 1986: 68). Furthermore, Chilton (2004: 111) is right when he says that "labelling stretches of language as serving strategic functions is an interpretative act". All in all, I believe that certain devices lead any researcher to analogous conclusions, and that certain linguistic patterns have certain implications other scholars can also examine when replicating similar experiments. The microanalysis of a text helps to support this point; its macro-analysis can be used to avoid misjudgement. Everything is meaningful in language. The selection of one item implies at the same time the exclusion of some others (Fairclough 1995: 210). In Fowler's words, "[d]ifferences in expression carry ideological distinctions (and thus differences in representation)" (1986: 4).

Categories such as ideology and power are present in the analysis of discourse practice (Fairclough 1995: 11). That is something I find revealing, because representation has to do with power. It is the powerful that have the chance to represent others (and themselves) in a light they may find more or less convenient. Metaphorically speaking, it is those who are powerful that give (or do not give) voice to those represented in their discourse practices. Their control of the media can allow them to arrange the ordering of events, and then to obscure or give more prominence to some participants instead of others.

9. Conclusion

CDA is an infant discipline gradually maturing. Curiously, several of its strengths can be taken simultaneously as the source of its weaknesses. Some of the exponents of the critical paradigm may themselves be lacking in a self-critical attitude since CDA has become an established discipline (Billig 2003). However, its general critical outlook has encouraged the development of new approaches, in an attempt to answer new research questions, and allay doubts about its method and theoretical grounds.

Its inter- and transdisciplinary nature still needs to be carried forward before it yields fruit. The ambition that CDA can help raise awareness about the unequal social conditions of minorities makes it a worthy enterprise. Nevertheless, both proponents and audiences are often familiar with this asymmetry and usually hold similar views: CDA is mostly consumed by CDA scholars not by the average woman or man in the street. Furthermore, despite CDA practitioners' activist orientation, their recent achievements only range from adjustments in the perception of a particular unjust state of affairs to cosmetic changes in advertising, news reports or political speeches.

Drawbacks notwithstanding, the adventure of CDA is to look into how discourses construct participants in communication as individuals with allegiances to the collective, and to embark on the analysis of the discursive means by which the world comes into existence. If this finally may bring increased understanding of social processes and structures, and ultimately perhaps, increased understanding of effects on social actors' views and actions, CDA must have a role in the social sciences.

Acknowledgements

My warmest thanks to Leanne Bartley, Andrew Blake, Paul Chilton Graeme Porte, J.C. Pascual and Salvador Valera.

References

Althusser, Louis. 1969/1971. "Ideology and Ideological State Apparatuses: Notes towards an Investigation." *Lenin and Philosophy and other Essays*. Ben Brewster, trans. N.Y., Monthly Review Press: 127-86.

Baker, Paul. 2006. *Using Corpora in Discourse Analysis*. London: Continuum.

Billig, Michael. 2003. "Critical Discourse Analysis and the Rhetoric of Critique." Gilbert Weiss and Ruth Wodak, eds. *Critical Discourse Analysis*. Houndmills: Palgrave Macmillan. 35-46.

Blommaert, Jan. 2005. *Discourse Analysis: A Critical Introduction*. Cambridge: CUP.

Bloor, Meriel and Thomas Bloor. 2007 *The Practice of Critical Discourse Analysis. An Introduction*. London: Hodder Arnold.

Cameron, Deborah. 2001. *Working with Spoken Discourse*. London: Sage.

———. and Don Kulick. 2003. *Language and Sexuality*. Cambridge: CUP.

Chilton, Paul. 2004. *Analysing Political Discourse. Theory and Practice*. Routledge: London and N.Y.

———. 2005. "Missing Links in Mainstream CDA: Modules, Blends and the Critical Instinct." Ruth Wodak and Paul Chilton, eds. *A New Agenda in (Critical) Discourse Analysis*. Amsterdam and Philadelphia: John Benjamins. 19-51.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Fairclough, Norman. 1989. *Language and Power*. London: Longman.

———. 1995. *Critical Discourse Analysis*. London: Longman.

———. 2000. *New Labour, New Language?* London and N.Y.: Routledge.

———. 2005. "Critical Discourse Analysis in Transdisciplinary Research." Ruth Wodak and Paul Chilton, eds. *A New Agenda in (Critical) Discourse Analysis*. Amsterdam and Philadelphia: John Benjamins. 53-70.

———. 2009. "A Dialectical-Relational Approach to Critical Discourse Analysis." Ruth Wodak and Michael Meyer, eds. *Methods of Critical Discourse Analysis*. London: Sage. 162-86.

Fairclough, Norman and Phil Graham. 2002. "Marx as a Critical Discourse Analyst: The Genesis of a Critical Method and its Relevance to the Critique of Global Capital." *Estudios de Sociolingüística* 3.1: 185-229.

Fairclough, Norman and Ruth Wodak. 1997. "Critical Discourse Analysis." Teun A. van Dijk, ed. *Discourse as Social Interaction, Vol 2*. London: Sage. 258-84.

Foucault, Michel. 2002. *The Archaeology of Knowledge*. London: Routledge.

Fowler, Roger. 1986. *Linguistic Criticism*. Oxford: OUP.

Fowler, Roger, Bob Hodge, Gunther Kress and Tony Trew. 1979. *Language and Control*. London: Routledge.

Gee, James P. 1999. *An Introduction to Discourse Analysis*. London: Routledge.

Gramsci, Antonio. 1971. *Selections from the Prison Notebooks of Antonio Gramsci*. Quintin Hoare and Geoffrey Nowell Smith, ed. and trans. London: Lawrence & Wishart.

Halliday, Michael A.K. 1985. *An Introduction to Functional Grammar*. London: Edward Arnold.

Hart, C. 2010. *Critical Discourse Analysis and Cognitive Science: New Perspectives on Immigration Discourse*. Basingstoke: Palgrave.

Hodge, Robert and Gunther Kress. 1988. *Social Semiotics*. London: Polity Press.

———. 1993. *Language as Ideology*. London and N.Y.: Routledge & Kegan Paul.

Jäger, Siegfried and Florentine Maier. 2009. "Theoretical and Methodological Aspects of Foucaldian Critical Discourse Analysis and Dispositive Analysis." Ruth Wodak and Michael Meyer, eds. *Methods of Critical Discourse Analysis*. London: Sage. 34-61.

Kress, Gunther and van Teun Leeuwen. 2006. *Reading Images. The Grammar of Visual Design*. London: Routledge.

Laclau, Ernesto. 1980. "Populist Rupture and Discourse." *Screen Education* 34: 87-93.

Lazar, Michelle M., ed. 2005. *Feminist Critical Discourse Analysis. Gender, Power and Ideology in Discourse*. Houndmills: Palgrave Macmillan.

Leont'ev, Alexi N. 1978. *Activity, Consciousness, and Personality*. Englewood Cliffs: Prentice Hall.

Link, Jürgen. 1983. "Was ist un was bringt Diskurstaktik." *kulturRRevolution* 1: 6-21.

Litosseliti, Lia. 2006. *Gender and Language: Theory and Practice*. London: Hodder Arnold.

Martin, James R. 1992 *English Text: System and Structure*. Amsterdam: John Benjamins.

———. 2000. "Close Reading: Functional Linguistics as a Tool for Critical Discourse Analysis." Len Unsworth, ed. *Researching Language in Schools and Communities: Functional Linguistic Perspectives*. London: Cassell. 275-302.

Marx, Karl and Friedrich Engels. 1845/2001. *The German Ideology Part One, with Selections from Parts Two and Three, together with Marx's 'Introduction to a Critique of Political Economy'*. N.Y.: International Publishers.

Mautner, G. 2009. "Checks and Balances: How Corpus Linguistics can Contribute to CDA." Ruth Wodak and Michael Meyer, eds. *Methods of Critical Discourse Analysis*. London: Sage. 122-43.

McClelland, James L. and Alan H Kawamoto. 1986. "Mechanisms of Sentence Processing: Assigning Roles to Constituents." James L. McClelland and David E. Rumelhart, eds. *Parallel Distributed Processing: Explorations in the Microstructures of Cognition, Vol. 2*. Cambridge, MA: MIT Press, 272-325.

Moscovici, Serge. 2000. *Social Representations*. Cambridge: Polity Press.

Nishida, Kitaro. 1958. *Intelligibility and the Philosophy of Nothingness.* Tokyo: Maruzen.

O'Halloran, Kieran. 2003. *Critical Discourse Analysis and Language Cognition*. Edinburgh: Edinburgh UP.

O'Halloran, Kieran and Caroline Coffin. 2004. "Checking Overinterpretation and Underinterpretation: Help from Corpora in Critical Linguistics." Caroline Coffin, Ann Hewings and Kieran O'Halloran, eds. *Applying English Grammar. Functional and Corpus Approaches*. London: Arnold. 275-97.

Reisigl, Martin and Ruth Wodak. 2009. "The Discourse-Historical Approach." Ruth Wodak and Michael Meyer, eds. *Methods of Critical Discourse Analysis*. London: Sage. 87-121.

Scollon, Suzanne. 2003. "Political and Somatic Alignment: Habitus, Ideology and Social Practice." Gilbert Weiss and Ruth Wodak, eds. *Critical Discourse Analysis. Theory and Interdisciplinarity*. London and N.Y.: Palgrave Macmillan. 167-98.

Scollon, Ron and Suzie Wong Scollon. 2005. "Lighting the Stove: Why Habitus Isn't Enough for Critical Discourse Analysis." Ruth Wodak and Paul Chilton, eds. *A New Agenda in (Critical) Discourse Analysis*. Amsterdam and Philadelphia: John Benjamins. 101-17.

Sunderland, Jane. 2004. *Gendered Discourses*. Basingstoke & N.Y.: Palgrave Macmillan.

Van Dijk, Teun A. 1997. “Discourse as Interaction in Society.” *Discourse as Social Interaction, Vol 2.* Ed. Teun A. van Dijk. London: Sage. 1-37.

———. 2003. “The Discourse-Knowledge Interface.” Gilbert Weiss and Ruth Wodak, eds. *Critical Discourse Analysis. Theory and Interdisciplinarity*. London and N.Y.: Palgrave Macmillan. 85-109.

———. 2005. “Contextual Knowledge Management in Discourse Production: A CDA Perspective.” Ruth Wodak and Paul Chilton, eds. *A New Agenda in (Critical) Discourse Analysis*. Amsterdam and Philadelphia: John Benjamins. 71-100.

———. 2009. “Critical Discourse Studies: A Sociocognitive Approach.” Ruth Wodak and Michael Meyer, eds. *Methods of Critical Discourse Analysis*. London: Sage. 62-86.

Van Leeuwen, Teun. 2006. “Critical Discourse Analysis.” Keith Brown, ed. *Encyclopedia of Language and Linguistics, Vol. 3*. Oxford: Elsevier. 290-94.

———. 2008. *Discourse and Practice. New Tools for Critical Discourse Analysis*. Oxford: OUP.

———. 2009. “Discourse as the Recontextualization of Social Practice: A Guide.” Ruth Wodak and Michael Meyer, eds. *Methods of Critical Discourse Analysis*. London: Sage. 144-61.

Weiss, Gilbert and Ruth Wodak. 2003. “Introduction: Theory, Interdisciplinarity and Critical Discourse Analysis.” Gilbert Weiss and Ruth Wodak, eds. *Critical Discourse Analysis. Theory and Interdisciplinarity*. London and N.Y.: Palgrave Macmillan. 1-32.

Widdowson, Henry G. 2004. *Text, Context, Pretext. Critical Issues in Discourse Analysis*. Oxford: Blackwell.

Wodak, Ruth and Michael Meyer. 2009. “Critical Discourse Analysis: History, Agenda, Theory and Methodology.” Ruth Wodak and Michael Meyer, eds. *Methods of Critical Discourse Analysis*. London: Sage. 1- 33.

Wodak, Ruth and Gilbert Weiss. 2005. “Analyzing European Union Discourses: Theories and Applications.” Ruth Wodak and Paul Chilton, eds. *A New Agenda in (Critical) Discourse Analysis*. Amsterdam and Philadelphia: John Benjamins. 121-35.

Reviews

Nørgaard, Nina, Beatrix Busse and Rocio Montoro. 2010. *Key Terms in Stylistics*. London and New York: Continuum.

Using a toolbox as a metaphor for the interdisciplinary nature of stylistics, the authors of this book state that their aim is to "provide the reader with an overview of the stylistics toolbox, the tools available in it, the different linguistic paradigms and branches of stylistics which have produced and/or employ the tools as well as key thinkers in the field." This is, indeed, a laudable aim and the book is a very welcome contribution to the field of stylistics.

The introduction defines stylistics as the analysis of "the ways in which meaning is created though language in literature as well as in other types of texts" (p. 2). This is a very broad definition which extends the scope of stylistics well beyond the language used in literature to nonfictional forms of language, such as advertising, and news reports, in fact, to all kinds of language use, aligning stylistics with functional approaches to language in general, which are also concerned with the creation of meaning in texts. The introduction also provides a brief historical background, which focuses mainly on developments from the 1960s to the present day and a short section on the growing significance of stylistics in education and academia, where it bridges the traditional gap between literature and linguistics.

The rest of the "stylistics toolbox" is divided into four "key" parts: branches, terms, thinkers and texts. *Key Branches* contains 15 subsections each dealing with one branch. These include well established branches of stylistics, such as historical stylistics, narratology, and feminist stylistics, and other branches which have emerged more recently. Some of these are primarily concerned with the texts themselves, such as corpus linguistics, for instance, which examines large amounts of data in order to identify norms and deviations in language use, and multimodal stylistics, which examines how meaning is created not only by words but by other means such as pictures, layout, visual images etc. Other recent branches in stylistics are concerned with the producers and receivers of texts, emphasizing, as the authors point out, "that the production of meaning needs to be accounted for as a double exercise encompassing as much text-informed inferences as the

mental processes that allow text comprehension" (p.1). These include entries on pragmatic stylistics, cognitive stylistics/cognitive poetics, reader response criticism and Emotion: stylistic approaches. Each section describes the most important features of this branch, its main practitioners and, in most cases, also gives examples how the approach has been used, as well as a critique of the approach, discussing various problems and points of criticism.

Key terms is an alphabetical lists of terms which are central to the various branches of stylistic analysis, including terms for phonological features, e.g. alliteration, metre and rhyme; pragmatic features, e.g. deixis, entailment, politeness, presupposition, and speech acts; lexical features, e.g. collocation, metonymy, semantic features, e.g. modality, functional features, e.g. ideational, interpersonal and textual meaning, grammatical metaphor, and discoursal features, e.g. lexical cohesion, and coherence. Most of the entries provide definitions and discussions, and there are many cross references to other entries. Some of the entries are not key terms as such but accounts of theoretical frameworks, e.g. Blending theory, Possible Worlds Theory and Mental Spaces Theory, and others are descriptions of stylistics associations and journals. There is also, under the heading *Critique of stylistics,* an account of earlier controversies over the value of stylistics as a discipline. All this is extremely valuable information, but, as they are not what one expects to find under the heading "key terms", they may not be easy for readers to find. It would perhaps have been better to present the theoretical frameworks in the *Key Branches* part of the book and the rest in the introduction.

Key thinkers describes the chief practitioners of stylistics, e.g. Roger Fowler and Mick Short, as well as other scholars who have had an impact on the field. These include the chief proponents of the main linguistic approaches, such as functionalism, e.g. Michael K. Halliday; formalism, e.g. Noam Chomsky, cognitive linguistics, e.g. Charles Forceville; corpus linguistics, eg John Sinclair and multimodality, Theo van Leeuwen, as well as "the father of linguistics", Ferdinand de Saussure**.** Each section includes biographical details and descriptions of their main contributions to the different branches of stylistics. Sometimes there is, however, an overly evaluative tone in the description, which is somewhat out of place in a reference guidebook, with expressions such as "the amiability and clarity with which his academic work is

delivered”(p.199); “moved on to prestigious positions in the academic world” (p.207), making the entry read more like a Curriculm Vitae than an informative account.

Key texts, the final part of the book, is not a collection of texts as such, but a select bibliography. This is a useful, comprehensive list of the chief contributors to the field of stylistics.

As a reference guide to stylistics, this book provides a comprehensive and up-to-date coverage of stylistics, incorporating the latest developments in the field. It does not, however, always entirely live up to the goals of the key terms series “To offer undergraduate students clear, concise accessible introductions to core topics.” Sometimes this is because the authors assume too much prior knowledge by, for instance, using some terms without explanation, e.g. “paratactic construction” (p.162), a term used in Systemic Functional Grammar. In general, however, this is due to a lack of clarity both in the organisation of the content and in the language in which this is presented. The latter is, unfortunately, further marred by somewhat unnatural English expressions (e.g. … *has been enormously felt* p.113). One gets the impression that the book has been put together in haste. Despite these shortcomings, which can be remedied by careful editing and revision, the main strength of the book remains. This is the wealth of information it contains and the contribution it makes to the field of stylistics today.

Jennifer Herriman
University of Gothenburg, Sweden

Armstrong, Charles I. 2009. *Figures of Memory: Poetry, Space, and the Past*. Palgrave Macmillan. 244pp. ISBN 13: 978-0-230-22353-0 hardback. £50.

The first figure of memory to confront the reader of this book is René Magritte's *La mémoire* (1948) on the cover. The reddish splash on the sculpted head might be blood from a wound, but it is also a patch of vivid colour standing out against the dead pallor of marble or plaster. This serves to introduce some of the ambiguities of memory as pain and pleasure, both wound or trauma and vivid and vital component of the life of the imagination. This sophisticated, theoretically informed study of ten poets from Wordsworth and Tennyson to Jeremy Prynne and Alice Oswald, well-versed in recent criticism, dwells on ambivalence, on memory in poetry as less or more than mimesis, nostalgia or almost new creation, as celebration not forbidding mourning. Memory can be collective as well as individual, so it is appropriate that the study includes the Irish poets W.B. Yeats, Seamus Heaney and Eavan Boland who come from a memory-haunted country that may need to do a lot of forgetting. It is also appropriate that prominence is given to two of the most influential theorists and practitioners of literary remembering, the Wordsworth of the *Prelude* and the T.S. Eliot of 'Tradition and the Individual Talent'. New light challenging the established poetics of memory is reflected from the radically revisionist work of three contemporary writers: the book concludes with an assessment of feminist critique of male tradition in Eavan Boland, future-oriented spatial aesthetics in the late modernism of J.H. Prynne and presentist ecopoetics in Alice Oswald.

The approach through the theme of memory delivers gains and losses, winners and losers. Variousness triumphs over tidiness and Yeatsian dreams of order. Poetry, criticism and the imagination are championed and vindicated both severally and collectively. Much-maligned mimesis as an aspect of art recedes even further into the twilight zone. The imperfections of hindsight are identified but championed rather than reproached. Wordsworth is reaffirmed as not just a monument but a living and energising presence in contemporary poetry. Christina Rossetti's intense lyrical opacities appear in a different and more interesting light in the company of Wordsworth, Tennyson and

Stefan George. Prynne's notoriously challenging work is sympathetically discussed in a way that should encourage baffled readers to try again.

But the rather oddly-focused Eliot discussion, only rather tangentially engaged with *The Waste Land*, perhaps says more about memory than about Eliot. Boland and Heaney are left alone and palely loitering rather than integrated into the rest of the discussion. It seems odd to pay so little attention to *Station Island*, Heaney's most direct and ambitious engagement with earlier writers such as Carleton and Joyce, the undead presences which haunt the Irish writer. It was less helpful than the author may have hoped to discuss Heaney immediately after Eliot: the links with Wordsworth are much stronger and Wordsworth's 'Fair seedtime had my soul' from the Prelude provides an epigraph for the important autobiographical sequence 'Singing School' in *North*. But 'Singing School' is not discussed and possible connections with Heidegger receive more attention than the actual connections with Wordsworth. While Yeats's ambition to conflate individual recollection with 'the collective memory of the nation, rooted in its native soil' is accurately identified and described there is no real recognition of the pressure this remembered legacy exerted on Heaney's negotiations of collective memory or how he survived it.

Hardy's increasingly recognised influence on contemporary poetry earns him a place in this book, but the manner of his incorporation, based on the Poems of 1912-13, is perhaps more ingenious than helpful. He is grouped with the Yeats of 'Under Ben Bulben' and Tennyson in Arthurian mode in a chapter entitled 'Weird Wests: Victorian and Post-Victorian Displacements of Nostalgia.' There is indeed a common reference to the west, whether Victorian Cornwall or Yeats's Connaught or mythological Arthurian territory, but these are rather different places even if Hardy's memories are Tennysonian as well as personal. One is tempted to devise other groupings such as 'Noxious Norths: Ibsen and Heaney' or 'Savage Souths: Faulkner and Shelley'. Even so, the reader is introduced or reintroduced to rewarding poems and there are there stimulating and suggestive readings. Hardy's haunting, enigmatic phrase 'chasmal beauty' in 'Beeny Cliff' is persuasively linked with the paradox of continuity and disjunction over time. It is for moments likes these, or for Armstrong's reading of Prynne's reading of Wordsworth's 'The Solitary Reaper', that one will come back to this book.

There are a few miscellaneous oddities. Armstrong is appropriately alert to remembered European and classical influences pressing upon the treatment of memory in English poetry and quotes Mallarmé and Rilke to good effect. But why quote Mallarmé only in translation when Rilke is quoted in German, followed by a translation? Proust's extended explorations of memory in *A la recherche du temps perdu* had a profound influence on modernist and contemporary writing, yet there are only two very brief passing references to him. Scott is ignored altogether though his negotiations of collective memory in poems and historical fictions cast a long shadow across literary Europe. Collective memory invites consideration of its unacknowledged sibling the collective unconscious, but Armstrong ignores the invitation and Jung is never mentioned. The Greek term *anamnesis* ('calling to mind', 'recollection') is important in Platonic and Aristotelian doctrine, which is presumably why the Greek title of Aristotle's *On Memory and Reminiscence* is transliterated, as *Peri mnemes kai anamnesis*, but *anamnesis* is actually in the genitive case in the title: it should read *anamneseos*. While readings are usually careful and sensitive it is not clear how many readers would recognize *In Memoriam* cxix as 'a fetishistic roll call of bodily impressions' even if they read it out of context.

But these are small matters. This is a thoughtful and informed study which renews our respect for poetic transformations of what has gone before.

Norman Vance
University of Sussex

Clarkson, Carrol. 2009. *J.M. Coetzee: Countervoices*. Houndsmill: Palgrave Macmillan.

Writing is, according to J.M. Coetzee 'a matter of awakening the countervoices in oneself and embarking upon speech with them' (quoted in Clarkson 7-8). An investigation of the countervoices in Coetzee's writing and his interest in problems of language is what Carrol Clarkson presents in *J.M. Coetzee: Countervoices*. The concept is related, both in Coetzee as well as in Clarkson's reading of it, to Bakhtin's dialogism. However, Coetzee makes, Clarkson stresses, and important observation about a lack in Bakhtin's understanding of dialogism in Dostoevsky: dialogism in Dostoevsky is not reducible to ideological positioning or novelistic technique, but something that grows out of Dostoevsky's own moral character (Clarkson 9). For Clarkson, therefore, Coetzee's concept of countervoices is not intellectual spice added to his fiction in the form of staged dilemmas, but a matter of intellectual involvement (Clarkson 8).

Since Coetzee was awarded the Nobel Prize for Literature in 2003 there has been a gradual emergence of a field of Coetzee criticism. Coetzee is a writer who eludes easy labels, even if the label 'South-African writer' has been one of the most resilient ones. The great variety in the Coetzee criticism is a good illustration of the difficulties of assigning Coetzee to a tradition: from the first monograph to appear, *The Novels of J.M. Coetzee: Lacanian Allegories*, by Teresa Dovey (1988), via socio-historical analyses, such as David Attwell's *J.M. Coetzee: South Africa and the Politics of Writing* (1993) and readings informed by post-structuralist theory, such as Derek Attridge's *J.M. Coetzee and the Ethics of Reading: Literature in the Event* (2004) to a more recent trend where Coetzee is read in conjunction with questions in analytic philosophy, such as in Stephen Mulhall's highly original and interesting *The Wounded Animal: J.M. Coetzee and the Difficulty of Reality in Literature and Philosophy* (2009), Coetzee continues to interest critics from different schools and traditions. Clarkson is the first, however, to start from Coetzee's preoccupation with linguistics and follow the repercussions of the exigencies of linguistics in questions of an ethical and aesthetic nature. Clarkson's discussions are compelling and interesting and undoubtedly constitute a novel and important contribution to the growing field of Coetzee criticism.

Clarkson's position within the field is difficult to determine. On the one hand, her interrogation of the authority and ethical aspects of the speaking subject, the ethics of address and the authority to speak on behalf of someone, be it 'the one who is absent'—the implication of the third person (Clarkson 37)—or a 'we', link Clarkson to central concerns in the post-colonial tradition. However, Clarkson has none of the ideological perspectives or the socio-historical approach which often characterises that tradition. One the other hand, even if Clarkson argues—convincingly—against the critics who see Coetzee's works as allegories for post-structuralist ethics (primarily Levinas), Lacan, Derrida, Lacoue-Labarthe and other continental thinkers are important reference points in many of the discussions. However, capital-T 'Theory' at no point overpowers the text. What characterises *J.M. Coetzee: Countervoices* is strict argumentation, which reminds the reader more of the analytic tradition in philosophy than the continental one. Perhaps the best way to position this book is to say that, like Coetzee in his writing, Clarkson engages intellectually the countervoices that arise.

In 1981 J.M. Coetzee published an article on Kafka's unfinished short story 'The Burrow' ('Der Bau') (reprinted in Coetzee, *Doubling the Point: Essays and Interviews* (1992), 210 ff.). Coetzee begins his article by quoting the first sentence of Kafka's story: 'Kafka's story "The Burrow" begins: "I have completed the construction of my burrow and it seems to be successful".' (*Doubling the Point*, 210). This article as well as Kafka's story may serve as illustrations of the aspects of Coetzee's writing that Clarkson investigates and the characteristics of Coetzee's writing which inspires Clarkson's investigations. The first sentence can be read as an allegory for the relation between the artist and his work or the writer and his writing, the possibility of an authorial intention and the elements of unpredictability and chance. Coetzee's approach to this story is not, however, an exploration of its allegorical dimensions even if the question of the relation between the artist and his work clearly is important to Coetzee in many of his texts, *Foe* and *Slow Man* being in their different ways the most obvious examples. Instead, Coetzee investigates—in linguistic and rather technical but nonetheless very interesting ways—'Time, tense and aspect in Kafka's "The Burrow"'. For Clarkson, perspectives such as this in Coetzee's non-fictional as well as his self-reflective pauses around linguistic inflections, etymologies and syntactic phenomena in his fictional writing serve as her impetus for

investigating 'in what ways [...] seemingly innocent linguistic choices on the part of the writer have ethical consequences for the position of the speaking or writing self in relation to those whom one addresses, or in relation to those on whose behalf one speaks, or in relation to a world one attempts to represent or create in writing' (Clarkson 1).

One example of how Clarkson establishes the relation between linguistic choices and ethics is her discussion of the speaking or writing 'I'. According to Clarkson, Coetzee's questioning of the speaking or writing 'I' does not start from an ideological position, but from a realisation of linguistic limitations: 'linguistic exigency dictates that writing cannot *but* imply an 'I' who writes', Clarkson argues. Having realised this limitation, which exists on the linguistic level, Coetzee therefore 'engages linguistic and literary strategies to question the authority of that 'I'' (Clarkson 21). The ethical dimension of choices at the linguistic level can be seen, therefore, in Coetzee's responses to what is linguistically and grammatically given. Discussions of the position of the writer and the nature of authority is not new in Coetzee criticism. What Clarkson brings to this discussion, however, and this is a feature of the entire book, is a sustained and meticulous investigation of Coetzee's own non-fictional writing, the relatively small set of interviews and Coetzee's own theoretical references as the framework for the discussion. Clarkson's research of what we perhaps can call Coetzee's formative years as a writer—the period from his time in London (his master thesis and the time portrayed in *Youth*), via his doctoral dissertation, to which Clarkson makes several references, during Coetzee's time in Texas and up to his time as an academic and professor at the University of Cape Town (the articles and interviews reprinted in *Doubling the Point* are central here), situates Coetzee's fiction (and non-fiction) within a much wider theoretical and philosophical context than has hitherto been done.

J.M. Coetzee: Countervoices does not offer a book-by-book analysis of Coetzee's writing as most extant monographs on Coetzee do. Instead the chapters are organised according to different topics which are linked to grammatical phenomena of central concern to Coetzee. Chapters 1 ('Not I') and 2 ('You') deal with 'Coetzee's careful exploration of the grammar of person in Roman Jakobson and Emile Benveniste carries through to the fields of aesthetics and ethics to become what Coetzee calls the "deep semantics of person"' (Clarkson 18). Chapter 3 ('Voice')

'constitutes a hinge between [the] discussion of the implications of Coetzee's experiments with structuralist conceptions of the "death of the author" in the first half of my book and a reinstating of some notion of authorial consciousness in the second half' (*ibid.*), whereas '[c]hapters 4,5 and 6, 'Voiceless', 'Names' and 'Etymologies', carry the idea of authorial consciousness through, but in relation to its situatedness within ethical, cultural and historical contingencies' (*ibid.*). The final chapter, 'Conclusion: We', draws together the discussions from all chapters, readdresses the grammar of person and speaks about what is at stake ethically when we say 'we'.

Clarkson's arguments are carefully organised and coherently presented. One chapter links to the next, and Clarkson frequently makes interesting and illuminating references to discussions in other places in the book. Combined with the thorough investigation of Coetzee's own theoretical references *J.M. Coetzee: Countervoices* emerges as a solid, interesting and apposite contribution to Coetzee criticism. The way Clarkson constantly refers back to Coetzee's own theoretical framework and her strategies of linking her own discussions throughout the book, however, lend this monograph a sense of completeness, which leaves at least this reader wondering to which degree he should read the discussions as an exposition of the theoretical underpinnings of Coetzee's writing and to what extent Clarkson takes the discussions beyond what was Coetzee's inspiration. This problem of determining the limits of the author's and the text's references, or even the impossibility of it, which is the problem every reading must admit, is an interesting illustration of one of the aesthetic problems which Coetzee explores. The final 'lesson' of *Elizabeth Costello*, 'At the Gate', is probably the best example of this problem, since that is perhaps the most elaborate exploration of the question of what remains of artistic creativity and artistic integrity within the structures which determine the meaning of the text. In 'At the Gate' Coetzee employs Kafka and the force from his short story 'Vor dem Gesetz' (which also appears in *The Trial*) as an example. In similar fashion we may ask, when reading Clarkson's book, at which point the framework which theoretical references constitute become determining for the path down which the discussion moves. I am reminded of the beast's cunning in Kafka's story mentioned above: the conspicuous hole in the ground is not the real entrance to the burrow. The real entrance is as well concealed as a thing in this world can ever

be. But still, anyone can stumble upon it. Thus Clarkson moves in the region in which the question about the text's determining structures is continuously be raised, and in that sense the discussions presented all return to this central aesthetic challenge for Coetzee, but also in profound ways to Clarkson's book title: countervoices.

J.M. Coetzee: Countervoices to my mind leaves the reader wanting more on two counts. First, even if frequent and very interesting references are made to Coetzee's fictional works, Clarkson uses these references to highlight and develop the theoretical discussions inspired by Coetzee's formative influences. In other words, no sustained interrogation of a single fictional work is presented. This aspect of *J.M. Coetzee: Countervoices* is clearly the result of a conscious choice on the author's part: this is not *that* book. The book cannot be faulted for that. Instead, readers will be looking forward to the critical discussions that undoubtedly will arise in its wake. Second, the final chapter 'Conclusion: We' reads both as a conclusion, but perhaps more as yet another interesting topic, a topic which deserves more attention that it gets in this chapter. In short, however, *J.M. Coetzee: Countervoices* presents a novel approach to Coetzee's writing and opens up new and important perspectives on one of the world's most critically challenging authors.

Kjetil Enstad
The Norwegian Military Academy

About the Contributors

EMMANUEL ADEDAYO ADEDUN is a Senior Lecturer in the Department of English, University of Lagos, Nigeria and a Visiting Scholar at the Centre for Language, Discourse and Communication, King's College, University of London. His research interests are Sociolinguistics and application of sociolinguistic insights to Advertising, Film and Pedagogy.

BEGOÑA CRESPO has published on the processes of lexicon enlargement in the History of English, genre/text-type taxonomy and the diachronic development of English scientific writing. She is currently supervising some PhD Dissertations on the language of science. She has co-organised several conferences on English Philology, Medieval Studies and Corpus Linguistics and summer courses on women studies. She is the main researcher of the project "Women and scientific discourse: social differences and language variation in English (1700-1930)". She has published in journals such as *Quaderni di Semantica, Studia Neophilologica, English Studies, Studia Anglica Posnaniensia and Literary and Linguistic Computing* among others as well as book chapters in D.S Brewer, Rodopi, Peter Lang, Lincom Europa or John Benjamins.

JENNIFER HERRIMAN does research within Systemic Functional Linguistics. Her main focus is on the organization of information structure in English and Swedish texts. She is professor of English linguistics at the Department of Languages and Literatures, University of Gothenburg.

ENCARNACIÓN HIDALGO TENORIO is a Senior Lecturer at the University of Granada, Spain, where she teaches a variety of undergraduate and postgraduate courses on English language and linguistics. Her research interests include gender studies, critical discourse analysis, corpus linguistics and Irish theatre.

MILENA KATSARSKA is Lecturer in American and Culture Studies at the University of Plovdiv, Bulgaria. Her publications are in the fields of culture studies, intercultural communication, disciplinarity and education. Among them are *Migration, Modern Nationalism and Nostalgia for the Homeland in the Age of Globalization* (Plovdiv University Press 2007, co-editor), *English Studies on This Side: Post-2007 Reckonings* (Plovdiv University Press 2010, co-editor), the chapter contributions "The Bulgarian Connection in *Harry Potter*" in Suman Gupta, *Re-Reading Harry Potter* (Palgrave Macmillan 2009, 2 ed.) and "English Studies in Non-Anglophone Contexts: Bulgaria" in Maria Georgieva and Allan James Eds. *Globalization in English Studies* (Cambridge Scholars Publishing, 2010).

DONKA KESKINOVA is Lecturer in Statistical Methods and Programming at the Department of Sociology at the University of Plovdiv, Bulgaria. She has participated in over 300 surveys and research projects funded by the WIPO (World Intellectual Property Organisation); EU Sixth Framework; EU Copernicus Program; The British Council; The Friedrich Ebert Foundation; Bulgarian Ministry of Culture; Bulgarian Ministry of Law; Bulgarian Helsinki Committee; BBSS Gallup International; Bulgarian Academy of Sciences – among others. Her current consultancy research includes network analysis and cluster analysis of the regional economy in the South East of England; Mapping of the regional health-technology cluster in the South East of England; and Mapping of the copyright sectors in the Bulgarian economy. Her main publications are in the field of cluster analysis, network analysis, and multiple correspondent analysis.

SIGNE OKSEFJELL EBELING holds a PhD in corpus-based contrastive linguistics and is associate professor of English language at the University of Oslo, Norway. Over the past 15-20 years she has worked on various corpus projects, including the English-Norwegian Parallel Corpus (University of Oslo) and the British Academic Written English corpus (Oxford Brookes University). Her research interests include contrastive analysis, phraseology and learner language.

ARNE OLOFSSON has an academic background in Latin and Greek, but he switched to English linguistics in the 1970s. From 1976 to 2010, when he retired, he worked at the English department of the University of Gothenburg, from 2001 as a professor. He has been involved in teaching (linguistics, proficiency and translation), academic leadership (head of department for 12 years) and research (four main fields: linguistic variation, including grammaticalization, word-formation, including etymology, lexicography, mainly bilingual, and English proficiency testing, including construction and evaluation of tests. His dissertation, *Relative Junctions in Written American English*, based on the Brown Corpus, was published in Gothenburg Studies in English in 1981. Its main theme, the choice of relativizers, reappeared in half a dozen articles 2001–2009 based on other corpora. Other books deal with the influence of Latin and Greek word-formation: *De internationella orden* and *Klassiska ord*, both in collaboration with Alvar Ellegård.

JOANA SALAZAR NOGUERA is an associate professor at the University of the Balearic Islands in Palma de Mallorca, Spain. She has participated in several European research groups and has created the educative website http://www.spanishkid.org to promote multicultural understanding. She has published on the adequacy of EFL methodologies in the Spanish educative context. She coordinates the Research Development Centre on Secondary Teacher Education of the ATEE (Association of Teacher Education in Europe).

MOJISOLA SHODIPE is a lecturer in the Department of English, University of Lagos. Her research interests are Variation Studies and Sociophonology.

JUSTIN WAGEMAN is an associate professor in the School of Education at North Dakota State University. He directs a curriculum organization that has provided professional development to American educators for over a decade. His research and teaching interests include language methodology, program evaluation, professional development, and assessment.