The language of Reports in general English language testing:

A corpus-based analysis.

Maria N. Melissourgou

University of the Aegean (Greece)

Abstract

Candidates of international English language exams are often asked to write a report. However, this kind of report is different from the frequently appearing research reports or reports in professional contexts. Information to guide learners specifically about this genre is scarce and not well documented. Therefore, the aim here, is to provide empirical data on the qualities competent writers use to achieve the communicative purpose of this genre. Having previously identified two basic variations of reports in this context, the Data Report and the Personal Observation Report, I describe the most prominent features of the language used in personal view reports based on quantitative evidence. Using a sub-corpus of the WriMA (Writing Model Answers) corpus (Melissourgou & Frantzi, 2015) I conduct a genre and corpus analysis. The WriMA corpus is a genre-based corpus especially created for this type of research. Examining frequent common words, keywords and patterns, we explain how forms relate to functions at semantic, organizational and stylistic level. Finally, I provide statistical data on sizes (text, paragraph, sentence, word-length) classified according to four CEFR levels (Common European Framework for Languages, Council of Europe, 2001).

1. Introduction

This study refers to all interested parties involved in English language testing (candidates, teachers, raters, educational material writers, testing institutions) as it investigates reports, one of genres appearing in the writing tasks of these exams. Throughout the analysis there is a genre-perspective so this study may also be of interest to anyone involved in genre studies (applied linguists,

apprentice writers or people working on automatic genre detection). The main interest however is in supporting writing education.

Using corpus linguistics tools, representative genre-specific language data is processed accurately and objectively, describing what is frequent and how it is used. Evidence is then discussed and interpreted according to genre theory principles.

By placing this sub-genre into its 'context of situation' (Malinowski, 1923; Firth, 1950) I observe the way language functions in specific environments that are characterised by specific participants, roles and purposes. Following a Systemic Functional Linguistics (SFL) perspective, the aim is to help students understand "*what* situational factors determine *what* linguistic features" (Halliday, 1978:32). Another aim is to encourage the teaching of grammar not as a discreet component but as a "network of interrelated meaningful choices" (Haliday & Matthiessen, 2004) used to achieve communication.

Offering insight regarding the particular genre arms teachers, assessors and material writers with specific evidence in order to assist candidates' preparation and evaluate their performance.

2. Genre, register and the Systemic Functional Linguistics framework

According to Hymes (1972:277), "a normal child acquires knowledge of sentences, not only as grammatical, but also as appropriate. He or she acquires competence as to when to speak, when not, and as to what to talk about with whom, when, where, in what matter." Speakers and listeners during speech production and comprehension strategically make use of typified discourses, which are based on specific social purposes, cultural beliefs, roles and relationships. These discourses are typified in the sense that they are "carried out repeatedly through the same practices" (Tardy, 2009:12), thus becoming recognisable and associated with specific social contexts. They may then be referred to as genres.

The context of situation is mainly defined by the field, the tenor and the mode of discourse (Halliday, 1978). In other words, before going into the text one needs to know what is being said and with what purpose, who the participants are, what is the social relationship between them and what

channel of communication is adopted. The variety of language that matches the specific 'context of situation', the lexicogrammatical choices used to communicate within a specific context is called 'register' (Halliday & Hasan, 1985:38,39).

In the Systemic Functional Linguistics (SFL) perspective the "context plays a part in determining what we say; and what we say plays a part in determining the context. As we learn how to mean, we learn to predict each from the other." (Halliday, 1978:3). By taking a functional view of language the interest is in society rather than the individual, it is the social role of language that is explored and the functions that this language serves.

Supporters of the genre-based-writing-instruction and SFL are known as the Sydney School, as this approach developed in the context of the Australian school system (Christie, 1984; Kress & Knapp, 1992; Martin, 1985, 2009). They argue that students should be taught the primary traits of particular genres through explicit teaching and the modelling of prototype texts. Genre awareness is a prerequisite for teachers in order to help students gain control over texts, boost their confidence and prepare them for active participation in community discourses.

3. Methodology

Corpus linguistics methods process the language data quantitatively. The analysis of corpora offers research on language such accuracy, completeness and speed that it would be impossible to achieve using traditional methods. For this analysis I am using a 'pedagogic corpus', which according to its latest definition by Meunier & Gouverneur, (2009: 179-201), is "a large enough and representative sample of the language, spoken and written, a learner has been or is likely to be exposed to via teaching material, either in the classroom or during self-study activities."

The WriMA (Writing Model Answers) corpus was especially made for this type of study as there were no corpora available designed according to the strict criteria the researchers had in mind. It includes a large sample of model answers appearing in textbooks and websites addressing various well-known international English language exams. It is divided into sub-corpora according to genres.

Texts are additionally categorised according to the CEFR level they address (Council of Europe, 2001).

The Personal Observation Report sub-corpus, consists of 76 texts (17,702 word tokens and 2,888 word types). Texts relate to CEFR levels B1 up to C2. The sub-corpus has been POS (Part of Speech) tagged with TagAnt 1.1.2 (Anthony, 2014), and is analysed using WordSmith Tools v. 6 (Scott, 2015). The metadata stored in Excel sheets contain the prompt, the source, the CEFR level and the word limit, if there is one.

Firstly, the specific sub-genre is explored using the metadata of the corpus to define the context of situation, and then it is contrasted to other genre family members referring to appropriate literature.

The following text analysis is based on high frequent common words and keywords using the whole WriMA corpus (1151 texts, 253,025 tokens), as reference to extract keywords. Keywords are those whose frequency (or infrequency) in a text or corpus is statistically significant, when compared to the standards set by a reference corpus (Baker 2004; Bondi, 2010; Scott 1997; Scott & Tribble 2006).

Looking closer at these words, I explore the communicative roles that language plays looking for functional units and patterns affected by the context in situation. I observe the writers' 'manipulation' of certain words to convey meanings and adhere to the conventions of the specific genre. Functional interpretation is further supported by seeing words in their immediate context using the concordance tool (WordSmith Tools) and noticing their dispersion in texts (dispersion plot tool) when there is a need to check where a specific word is mostly used.

Statistical data such as text, sentence and paragraph length, words per sentence and word-length are also provided as means for each CEFR level. These measures have been widely used in genre analysis (e.g., Nesi & Gardner, 2012) and automatic genre detection (e.g., Stamatatos et al., 2001).

4. Placing the 'Personal View Report' into the 'Report' genre family

Research on the report genre as a whole has mostly focused on 'Research Reports' (Nesi & Gardner, 2012; Paltridge, 1997) and 'Workplace (or Professional) Reports' (e.g. Bondi & Danni, 2015; Flowerdew & Wan, 2010) or both (Devit et al., 2004). This is understandable as the first type is necessary for every student or scholar in higher education and the second is required in companies, organisations or other professional contexts. Research has also involved school-based reports, naming them 'Informational Reports' as this is the type of report widely used in the Australian primary school context (Board of studies NSW, 1998; Cope & Kalantzis, 1993; Derewianka, 1990; Knapp & Watkins, 1994; Martin, 1985).

However, the type of report investigated here resembles neither 'Research' nor 'Informational Reports'. The 'Personal Observation Report' (POR) is not similar to Research Reports, as reporting is not based on raw data, facts and figures. The themes are not generalisable as in the case of school-based Information Reports (e.g. Frogs, the life cycle of a grasshopper, our solar system). The subject is often very specific and the writer is asked to assess and draw personal conclusions based on subjective views and the proximity to a place or personal experience, which the recipients do not share. This is obvious by looking at the prompt:

e.g. *You work for a local magazine. A new take-away restaurant has opened in your area. The editor has asked you to visit it and write a report saying whether you recommend it or not.*

Most of these tasks seem to share more similarities with Workplace Reports since they are more product-oriented, that is, the ability to conduct research is not a prerequisite as in the case of academic Research Reports. In the previous prompt, the use of pronouns (e.g. whether *you* recommend it or not) shows that personal involvement is not only allowed but also required.

In PORs specific recommendations are often asked for instead of general conclusions. Here there is real-like need for immediate sharing of personal knowledge or views based on personal experience for practical reasons, as is the forthcoming assessment of an employee or a group of people visiting the writer's town soon. There is a practical need. As Nesi and Gardner (2012) explain referring to Workplace Reports, "the methodology employed by the professional report writer may

also be of little concern to the reader; clients are unlikely to care about the replicability of results, but do want the right solution to a possibly unique problem".

We could say then, that the POR is a workplace-like report. Since candidates for these exams are often quite young, these tasks are often adjusted to their interests and experience. The writer for instance, works in a children's camp and needs to assess an employee or works for a teenage magazine or needs to recommend places to visit for a visiting group of students.

Workplace simulation is understandable, as the candidates of these exams may need to use the language they are being tested on in workplace environments. However, seeing this genre in high-stakes language exams, one cannot help thinking that the certificates obtained may also be used for university entrance and this kind of tasks does not prepare them for the Research Reports they are going to use in that case. In fact, the need for reading and interpreting graphs in a scientific way, as in the case of reports in academic IELTS (International English Language Testing Services) exams, is not evident here. Few rubrics mention a previously conducted survey and in rare cases the results are included in the rubric. These may be presented in the form of diary-notes or main points, hypothetically written by the writers themselves. Mainly, the writers are asked to report based on their own views.

e.g. *You went to <place> for five days to attend your sister's wedding. Look at the extracts from the diary you kept during your trip there. Then write your report on the trip.*

Between these two reports in language testing, this has seemed to me the most prevalent difference and has been the reason for coining the names 'Personal Observation Report' and 'Data Report' respectively. Although the POR is distinctly different from other family members and it is a common requirement in exams involving millions of candidates every year, it has not been researched enough.

5. Textual analysis

5.1. Structure based on thematic sub-units

The presence of headings and subheadings is a characteristic feature of this genre. Sometimes there is a main heading and other times there is some introductory data before the text is divided into smaller parts. In table 1, we can see that in 22,3% of the total texts there is use of section headings only. A main heading is used in 28,9% of the reports and it is combined with section headings in 26,3 % of these texts. Nearly half of these reports (44,6%) include introductory data in a steady pattern of *To-From-Date-Subject.* This is combined with section headings in 38,1% of the texts. Finally, there is no use of main headings, section headings or data in only 4,2 % of the reports where the layout resembles any type of essay that is, the text is simply divided by paragraphs. It is worth mentioning here that no preference was observed in using any of this across CEFR levels.

Table 1. The use of headings

Section heading only	Main Heading		Introductory data		Nothing
22,3%	28,9%		44,6%		
	M.H. + S. H	only M.H	I.D. + S. H	only I.D.	4,2%
	26,3%	2,6%	38,1%	6,5%	

- *M.H.= Main heading, S.H.= Section heading, I.D.= Introductory data*

In cases where there are sub-headings their role is to divide the texts in thematic sections and classify interesting parts of the object/place/ person described. Sometimes these sections are defined by the question itself, for example:

You are a secretary for a primary school. The headmistress has asked you to assess a newly established local wildlife park to see if it is suitable to use for field trips. Write your report, describing what there is to see and do there and what the facilities and prices are like.

However, most of the times the writers have to decide on their own and prioritise on the qualities they are going to focus on.

This is an important step as it will define the outline of the description and facts taken from this description will perhaps be used later on to justify recommendations. Since there is a word limit in these tests careful planning at this stage seems necessary.

Regarding headings, three words are used really often: *Introduction, Conclusion, and Recommendation(s)*. The word *report* is used both in main headings *[A report on...]* but most commonly in the first sentence stating the writer's aim *[The purpose of this report is to / The aim of this report is to]*.

Then there is a wide range of different words or two to there-word phrases used as section headings. Thematically speaking though there are two categories, which appear often though using varied combinations of words. One of them has to do with 'cost' and the other with 'strengths and weaknesses' as shown in table 2.

Table 2. Common thematic categories in section-headings

'Strength/weaknesses' headings	'Cost' headings
Weaknesses	Prices
Strengths	Cost
Positive features	The food and the prices
Positive/ Negative points	Projected costs
Positive/ Negative sides	Prices and service
Positive/ Negative aspects	Availability and cost
Problems	Room prices

Problem areas	Quality and price
	Running costs
	Cost and service
	Facilities and prices
	Prices young people can afford

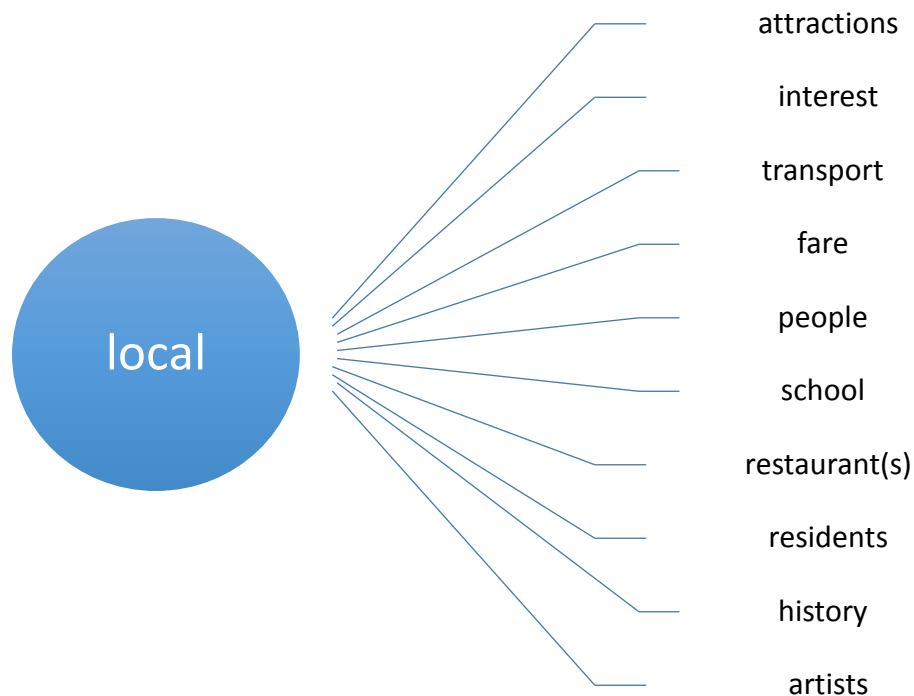**5.2. Factual language- Specific information of local interest**

Going back to the prompts we notice that the writer may be working for a magazine or may be a student representing the college students, a company employee, a pupil privileged to be asked by teacher, a part-time employee in a place where young people spend time or a member of a committee. In other words, the writer has been chosen to write a report because of particular abilities or roles in the community. The addressee is usually of higher status; it can be the Editor of a magazine, the Principal, an international research group, the company manager, the coordinator, the chairman of the local board of school governors, the visiting group-leader, the teacher or head teacher.

The writer is trusted to be critical and objective based on hypothetical previous evidence related to his/her performance. In some cases, the ability of the writer to combine studies with a part-time job may also be used as a trick to add extra competence on his/her profile and a real-life view gained from the working sector. This is why, specific, factual language is needed. However, most of the times the interest remains local and the writer knows the hypothetical reader of the report.

As we saw on table 2, mentioning prices is common in these reports. This is done by using specific numbers and percentages and referring to specific currencies. Numbers are also used for dates and ages. However, quantification may also involve grouping sizes in a more general sense [e.g. a wide range of/ a number of / many/ several/ few].

Humans are identified by name and surname; specific locations are also named (restaurants, schools, sports clubs, streets, shopping centres). When humans are not named they are grouped together in distant and general categories (students, people, teenagers, children, teachers, visitors, staff, group) and there is no presence of relatives as it often happens in stories for example. Whether it is for people, places or facilities the adjective *local* comes second in the adjectives used, which shows the proximity as a common cause for the person chosen to report. Figure 1 shows the most frequent collocates of *local* on the right.

<u>Figure 1. Collocates of *local*</u>

5.3. Comparison and evaluation

Comparing features is basic in these reports (see table 2) in order to describe and evaluate the specific entity contrasted to other available options.

Comparison is also common showing the degree of the quality (*very, quite, rather, too* plus adjective) or adjectives in the comparative form. There is extensive use of *more* and *more than* and

comparative adjectives used are more than double the number of those in superlative form. Looking at the dispersion plot, *more* is used throughout the reports. This is because comparison is also used at the justification stage, something is suggested because *it is more + adjective.*

Well + past participle is frequently used to evaluate *[e.g. a well-stocked bookshop/ the complex is well-managed/ a large well-equipped gym].* Conjunctions and adverbials such as *however, although, despite the,* help to differentiate between the qualities or the features described and go from strengths to weaknesses *[e.g. Everything is arranged in an organised manner, which makes shopping easy. However, the supermarket is not close to the town centre and finding a parking space can be difficult. / Although the local bus service works well during the week with regular, punctual buses, the weekend is a different matter.]*

5.4. Recommendation and justification

Recommendation is a key stage in these reports and often specifically expressed as a section-heading. Then writers may choose to justify their recommendation more clearly. Depending on the type of the question, justification may appear close to the end of the text if the purpose has been to present/describe and then recommend and justify or it may be present throughout the text when the task asks for recommendations only.

Adjectives are essential during the recommendation stage. They are used to describe existing positive qualities *[e.g. the restaurant offers excellent food]* and then recommend the place/person because of them, or to suggest ways of improvement so that the place /person will acquire these qualities *[e.g. New computers with Internet connections would be an excellent resource for students].* Positive connotation adjectives outnumber those with negative connotations *[e.g. more, good, young, new, excellent, popular, available, great, best, easy, helpful, modern, positive, reasonable, friendly, pleasant, suitable, useful].*

Modals are also used to suggest improvements, mainly *could* and *should [e.g. although improvements could certainly be made, students should be encouraged to.../ more exercise time could be introduced].]*

Conjunctions such as *because*, *as, so,* are extensively used to justify the recommendations made *[e.g. There are two attractions that may be taken into consideration, as they seem suitable/ students would have benefitted from a specialised language course as some communication problems were experienced / My suggestion is to visit the (name of place) because it is special].* Cause and effect relationships are also used to make suggestions and justify them using 'by + gerund' *[e.g. By featuring these activities, people would have the chance to, the problem could be resolved by adding a few chairs]* and *if- conditionals [e.g. if there is a prize to motivate students, I think this will encourage them].*

5.5. Subjectivity versus Objectivity

Although the writers are allowed to express personal views and in some cases are specifically asked to do so it seems that this is done mostly by describing and presenting features rather than imposing conclusions right from the start. Using description and comparison stages the writers lead the reader to the conclusion that what is recommended is in fact the best option available. While there is some subjectivity in these reports, this is controlled and wrapped in an objective package, more discrete than the subjectivity one can find in an essay where the writer is asked to express his/her opinion for example.

Complex phrases may add formality and objectivity avoiding to address the recipients directly *[e.g. It would be a good idea to, ... may be taken into consideration, taking into account the suggestions made above, it is hoped that, it is recommended that...]* using passive instead of active voice. However, active present tense is more frequent.

In the following table (table 3), one can see how phrases including the verb *recommend* are mostly in active rather than passive voice and the extent of personal involvement.

Table 3. Recommending

Active

(The purpose/aim of this report is to … and) to recommend (9)

I would recommend (4)

I strongly recommend (4)

I recommend (3)

I am going to recommend (2)

We recommend (2)

I would not hesitate to recommend (1)

I would certainly recommend (1)

I would definitely recommend (3)

(I am writing this report) to recommend (1)

I highly recommend (1)
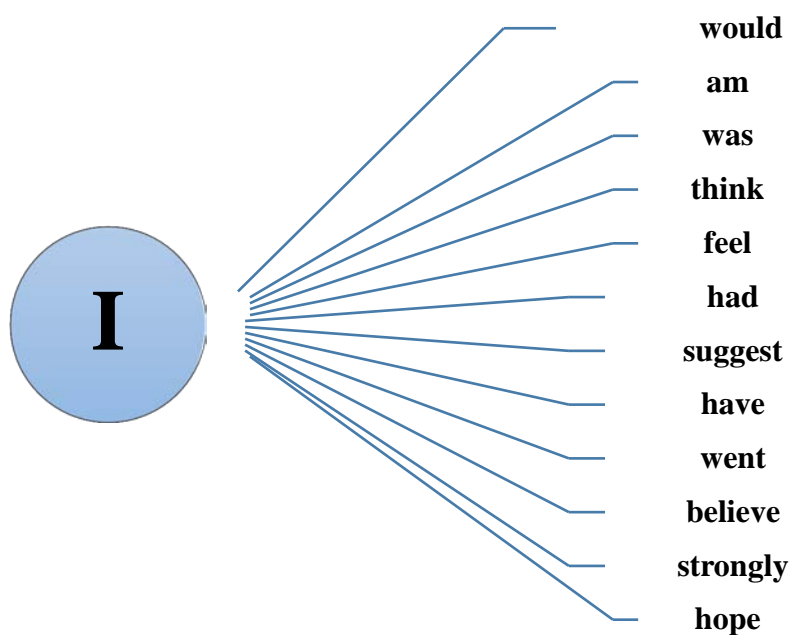
I am recommending (1)

Passive

It is recommended that (4)

It is hoped that the measures recommended above (will be adopted) (2)

<place> comes highly recommended for… (1)

_____

Personal involvement is also evident in the high frequency of the personal pronoun *I* and *we*. Out f the texts used in this sub-corpus 64% included the pronoun *I* and 33% the pronoun *we*. In figure 2, the most frequent collocates of *I,* show some participation in visiting the places evaluated as well as clear expression of personal beliefs, opinions and suggestions.

Figure 2. Most frequent collocates (on the right) of the pronoun _I_.

I → would, am, was, think, feel, had, suggest, have, went, believe, strongly, hope

Quantification, as previously shown, is also used both objectively (numbers, percentages) and subjectively in rough evaluations (a wide range of, many, a number of).

Nominalisation is a common technique used by writers to 'pack meaning' and add a sense of objectivity. Processes as well as properties can be turned into things allowing for more information in the same sentence. Nominalisation is evident in section-headings _[e.g. Introduction, Conclusion,_

Recommendation] and in introductory data after the Subject category. *[e.g. Subject: Suggestions on...].* Some examples of nominalisation instances found in these reports can be seen in table 4.

Table 4. Nominalisation

Noun	Nominalisation instances
suitability	The purpose of this report is to assess the suitability of…
choice	The city museum is a good choice for…
increase	The slight increase in the cost of…
improvements	… although improvements could certainly be made
suggestions	… if the zoo follows these suggestions…
information	… would have appreciated more detailed information…
organisation	… involving the organisation of new entertainment

5.6. Corpus statistics

Measuring sizes and the general organization of the texts (table 5), we notice that there is an upward trend for the number of words used in each CEFR level with the largest increase from B2 to C1 (66 words). The same happens for the number of paragraphs with an exception between B2 and C1 levels which share the same value. Sentences used across texts also follow a steady increase from B2 to C2 but we can see almost no change between B1 and B2. Words used on average consist of 4 to 5 letters. All levels on average use a 5-paragraph model and texts open and close with slightly smaller paragraphs than the ones in the middle.

Table 5. Size and organization patterns

Average/ CEFR	B1	B2	C1	C2
Word length per text	167	200	266	291
Paragraphs per text	4,6	4,9	4,9	5,1
Sentences per text	11	11,5	13,5	16,1
Words per sentence	15,2	17,4	19,7	18,1
Mean word length	4,6	4,6	4,9	4,9
Sentences per paragraph 1	1,2	1,5	1,4	1,8
Sentences per paragraph 2	2,8	3	4,1	3,7
Sentences per paragraph 3	3,2	2,7	4,1	3,2
Sentences per paragraph 4	2,5	2,5	2,3	3
Sentences per paragraph 5	2	2,3	1,8	2,1

6. Implications for teaching and assessment

Writing in testing contexts can be challenging. Time constraints and word-limits put extra pressure on candidates who need to be well prepared and confident. Inferences drawn about test-takers' language abilities often result in high-stakes decisions such as university admission, graduation, citizenship and immigration issues or access to jobs and promotion in professional contexts. Anxiety can affect response. Therefore, teachers need to be well informed and use teaching time effectively. They should first try to cover the distance between second-language students and authentic contexts by looking at the genre properties beyond the text. It has been shown that largely important aspects such as the writer-recipient relationship and roles as well as the communicative purposes of the genre can often be understood by reading the prompt carefully.

This analysis has also shown that PORs share some stereotypical properties that could be explored and discussed while modeling and deconstructing the genre. Following a Genre-based-Writing-Instruction approach, the teachers are able to bring these specific uses of language to conscious awareness. If teachers are themselves aware of the basic common characteristics of the specific register they can shortcut the whole procedure by guiding students during the deconstruction stage. It should be noted however, that the findings of this study are not to be given as a rigid template but carefully brought to the attention of the students as common generic tendencies.

Additionally, a range of activities could be built on these language features for further practice preferably incorporated in similar texts.

Finally, these findings can aid both everyday classroom assessment and exam raters' work in raising awareness of the specific genre by providing particular language features and organisational properties used frequently by competent writers. Quantitative evidence from a range of representative genre-specific texts can offer more solid answers than intuition as to what should be taught and what should be sought for during assessment.

References

Anthony, L. (2014). *TagAnt (Version 1.1.2)* [Computer Software]. Tokyo: Waseda University.

Baker, P. (2004). 'Querying keywords. Questions of difference, frequency and sense in keyword analysis.' *Journal of English Linguistics*, 32(4), 346-359.

Board of studies. (1998). *K-6 English syllabus: Modules*. Sydney: Board of Studies.

Bondi, M. (2010). Perspectives on keywords and keyness: An introduction. In Bondi, M. & Scott, M. (eds.), *Keyness in Texts*. Amsterdam-Philadelphia: John Benjamins, 1-18.

Bondi, M. & Danni, Y. (2015). Forward-looking statements in CSR reports: a comparative analysis of reports in English, Italian and Chinese. In Formato F. & Hardie A. (eds.), *Corpus Linguistics 2015 conference abstract book*. Lancaster: UCREL, 58-59.

Christie, F. (1984). Varieties of written discourse. In *Language studies: Course study guide*. Geelong, Victoria: Deakin University.

Cope, B. & Kalantzis, M. (1993). *The Powers of Literacy: A Genre Approach to teaching Writing.* Pittsburgh: University of Pittsburgh Press.

Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment.* Cambridge: Cambridge University Press.

Derewianka, B., (1990). *Exploring How Texts Work*. Newtown, Australia: Primary English Teaching Association.

Devitt, A. J., Reiff, M. J., Bawarshi, A. (2004). *Scenes of writing: strategies for composing with genres.* US: Pearson Education.

Firth, J. R. (1950). Personality and language in society. *The Sociological Review*, 42 (1), 37-52.

Halliday, M.A.K. (1978). *Language as social semiotic: the social interpretation of language and meaning*. Newcastle upon Tyne: Edward Arnold.

Halliday, M.A.K. & Hasan, R. (1985). *Language, context, and text: aspects of language in a social-semiotic perspective.* Oxford: Oxford University Press.

Halliday, M.A.K. and Matthiessen, C. (2004). *An Introduction to Functional Grammar*. London: Arnold.

Hymes, D. H. (1972). On communicative competence. In Pride, J.B.& Holmes, J. (eds.), *Sociolinguistics: selected readings*. Harmondsworth: Penguin, 269–293.

Knapp, P., & Watkins, M. (1994). *Context-Text-Grammar: Teaching the Genres and Grammar of School Writing in Infants and Primary Classrooms*. Sydney: Text productions.

Kress, G. and Knapp, P. (1992), Genre in a social theory of language. *English in Education*, 26: 4–15.

Malinowski, B. (1923). The problem of meaning in primitive languages. Supplement 1 to C. K. Ogden and I. A. Richards, *The meaning of meaning: a study of the influence of language upon thought and of the science of symbolism* (8th edition, 1946). NY: Harcourt Brace and World, 296-336.

Martin, J.R. (1985) *Factual writing: exploring and challenging social reality*. Geelong, Victoria: Deakin University.

Martin, J.R. (2009) Genre and language learning: A social semiotic perspective. *Linguistics and Education*, 20(1), 10–21.

Melissourgou, M. N. and Frantzi, K. T. (2015). *Investigating genre-specific qualities in the Writing part of popular EFL exams: The need for a specialised corpus*. Paper presented at LIF 2015, Language in Focus International Conference, Cappadocia, Turkey.

Meunier, F., & Gouverneur, C. (2009). New types of corpora for new educational challenges: Collecting, annotating and exploiting a corpus of textbook material. In Aijmer, K. (Ed.), *Corpora and language teaching*. Amsterdam and Philadelphia: John Benjamins, 179-201.

Nesi, H., and Gardner, S. (2012). *Genres across the Disciplines: Student writing in higher education*. Cambridge: Cambridge University Press.

Paltridge, B. (1997). *Genre, Frames and Writing in Research Settings*. Amsterdam/ Philadelphia: John Benjamins.

Flowerdew, J. and Wan, A. (2010). The linguistic and the contextual in applied genre analysis: The case of the company audit report. *English for Specific Purposes*, 29(2), 78-93.

Scott, M. (1997). PC analysis of key words – and key key words. *System,* 25(1), 1-13.

Scott, M. (2015). *WordSmith Tools v. 6*, Stroud: Lexical Analysis Software.

Scott, M. and Tribble, C. (2006). *Textual Patterns: Keywords and Corpus Analysis in Language Education.* Amsterdam: John Benjamins.

Stamatatos, E., Fakotakis, N., Kokkinakis, G. (2001) Automatic Text Categorization in Terms of Genre and Author. *Computational Linguistics*, 26(4), 471–495.

Tardy, M. C. (2009). *Building genre knowledge*. Indiana, US: Parlor Press.