

Using a comparable corpus to investigate lexical patterning in English abstracts written by non-native speakers

Carmen Dayrell¹ and Sandra Aluísio²

University of São Paulo (USP) - Brazil

¹Department of English

Rua Prof. Luciano Gualberto, 403/sala 14 - Cidade Universitária CEP: 05508-900 - São Paulo

²Centre of Computational Linguistics (NILC)/ Department of Computer Sciences

Caixa Postal: 668 - CEP: 13560-970 - São Carlos

E-mails: dayrellc@gmail.com, sandra@icmc.usp.br

Abstract

This study is set in a broad context of development of courses and computational tools to aid Brazilian graduate students in writing scientific papers in English. The main focus is on experimental research papers from the disciplines of physics, pharmaceutical and computer sciences. One of our primary objectives is to give students feedback and raise their awareness of the most typical lexical patterns used by their academic discourse community while at the same time draw their attention to the various available alternatives. Errors related to lexical use are by far the most frequent errors made by Brazilian graduate students when writing academic English. The aim of this paper is to carry-out a corpus-based study to investigate potential differences in the collocational patterns of *work* in abstracts written by Brazilian graduate students as opposed to abstracts collected from published papers of the same discipline. Relevant differences were found between the two subcorpora. The results were validated by examining the identified lexical patterns in a reference corpus of English abstracts. We also identified various items other than *work* which may be used to refer to the study described in the abstract as well as other lexical variations within the lexical patterns analysed.

1. Introduction

Scientific writing poses considerable challenges for non-native speakers of English. In addition to complying with the conventions and norms adopted by their academic discourse community, they also have to deal with the various difficulties involved in the complex process of writing in a foreign language. These problems are even more acute if the writer is an inexperienced researcher and he/she does not have full command of English grammar and usage at the sentence level.

Within the specific context of English Language Teaching (ELT), much effort has been spent on producing material to aid non-native speakers in overcoming the various problems which they may face when writing research papers. Swales (1990, 2004), Swales & Feak (2000), Weissberg & Buker (1990) are good examples of pedagogically useful studies which focus on the description of phrases and lexical patterns which are frequently used in academic discourse. It is also worth mentioning that several websites¹ are now available to provide users with practical guidelines when producing academic English.

Another important contribution is offered by studies based on learner corpora of academic English (among others, Thompson, 2001, 2006; Lee and Swales, 2006; Hyland, 2008a, 2008b). These studies opened up new perspectives and provided useful insights which enhanced

our understanding of underlying regularities in the language produced by students.

Benefits were also gained with computer-aided writing tools such as the *English Grammar and Spelling Software - Advanced Solutions for Your Writing*² which, in addition to a grammar and spell checker, also includes a dictionary, a thesaurus and a list of the most relevant collocates. Further achievements came from computational tools which take a step further and provide users with extracts from authentic research papers retrieved from a reference corpus of the discipline in question. This is the case of two writing tools developed at the University of São Paulo, namely AMADEUS and *Scipo-Farmácia*³. The former focuses on the disciplines of physics and computer science (Aluísio & Oliveira, 1995; Aluísio & Gantenbein, 1997; Aluísio *et al.*, 2001) while the latter deals with pharmaceutical sciences (Aluísio *et al.*, 2005; Genoves Jr. *et al.*, 2007). Similarly, Narita *et al.* (2003) and Anthony (2006) focus on English texts by Japanese speakers and developed computational tools to help user structure the text and produce adequate sentences in English.

Needless to say, various courses on English for Academic Purposes are offered worldwide. Most relevant to this study are the courses on academic writing offered annually by the University of São Paulo to graduate students. As we shall see in the next section, these courses have provided the data which is analysed in this paper. For the time being, what is important to mention is that the present study is set in a broad context, which includes a joined effort by various departments at the University of São Paulo to develop courses and computational tools to

¹ To mention but a few: *A Guide to Grammar and Writing*: <http://grammar.ccc.commnet.edu/grammar/>, *A Guide for Writing Research Papers Based on Modern Language Association (MLA) Documentation*: <http://www.ccc.commnet.edu/mla/index.shtml>, *Common Errors in English*: <http://www.wsu.edu/~brians/errors/>

² <http://www.whitesmoke.com/>

³ *Scipo-Farmácia* can be accessed at <http://www.nilc.icmc.usp.br/scipo-farmacia/>

aid Brazilian graduate students in writing scientific papers in English. The main focus is on experimental research papers from the disciplines of physics, pharmaceutical and computer sciences. Our long-term objective is two-fold: to improve course materials and resources for academic English and to provide computer-aided writing tools with linguistic knowledge so as to enable the automatic identification and correction of errors at the lexical, syntactical and rhetorical levels.

This paper is part of a larger project which aims to investigate errors made by Brazilian graduate students in academic writing. Our primary aim is to carry out a corpus-based study on the collocational behaviour of lexical items which frequently pose a challenge for Brazilian writers. The focus is on errors related to lexical use which is by far the most frequent error made by Brazilian students when writing scientific papers in English (Genoves Jr. *et al.*, 2007). These refer to misuse of a word to express a particular meaning. They may refer to direct translations of a Portuguese item into a false cognate in English (*pretend for intend*) or errors made in a common idiom (*as for such as*) or common collocation (*do contributions for make contributions*). There are other cases which are related to naturalness, that is, the writer's lexical choice is not most frequently used in that particular context, although its semantic meaning is fairly appropriate.

This paper focuses on the lexical patterning of the item *work*. A pilot study is carried out to investigate potential differences in the collocational patterns of *work* in abstracts written by Brazilian graduate students in comparison with abstracts collected from published papers of the same discipline. This idea relies on Sinclair's (1991:6,108 and 2003:3) suggestion that words do not occur randomly in a text but are instead closely associated with their surrounding context. According to Sinclair's (1991:6), the use of a given lexical item is related to specific lexical and grammatical patterns. Thus, by identifying differences in the lexico-grammatical patterning of abstracts written by students and published abstracts, we hope to be able to raise students' awareness of their most frequent errors as well as to draw their attention to the use of chunks which are regularly used within their academic discourse community. The results are validated by examining the identified lexical patterns in a reference corpus of English abstracts which were collected from papers published in reference journals of the disciplines under analysis. All procedures described below are carried out by means of the software package *WordSmith Tools*, version 4.0 (Scott, 2004).

2. The Comparable Corpus of English Abstracts (CCEA)

The data analysed in this paper is drawn from a monolingual comparable corpus of English abstracts which consists of two separate subcorpora: one made up of abstracts written by Brazilian graduate students and the other consisting of abstracts from published papers. The subcorpus of English abstracts written by students

(hereafter EA-STS) contains 84 abstracts which were collected in four courses on academic writing offered to graduate students from the disciplines of pharmaceutical sciences (20 abstracts), biology/genetics (11), physics (27) and computer science (26) at two universities in Brazil between 2004 and 2006. Here, we examine the first version of the abstracts, that is to say, abstracts handed in before the course started.

The subcorpus of English abstracts extracted from published papers (hereafter EA-PUB) was designed to match the specifications of the EA-STS subcorpus so that the two collections could be made comparable. Thus, EA-PUB also includes 84 abstracts from the same four fields of research, paying special attention to the number of abstracts in each. The abstracts were randomly selected from of various academic journals, using the WebBootCat tool⁴ as a starting point to select websites to be consulted. WebBootCat is a tool designed to help users to quickly produce corpora from any domain or subject (Baroni *et al.*, 2006). In an attempt to diversify the selection of journals as much as possible, no more than 3 abstracts were selected from each journal. In terms of number of words (tokens), the EA-STS and the EA-PUB subcorpora contain 18,004 and 21,061 words respectively.

An important methodological point to make here is that by published abstracts we do not mean that they are all written by native speakers of English. What is assumed is that they are of acceptable quality because they have been published by recognised bodies of a given discipline. Thus, published abstracts are presumably more likely to comply with the pre-established conventions adopted by the discourse community in question. Another difference between the two subcorpora is that most abstracts included in the EA-PUB come from papers by more than one author. This may mean that they have been more carefully revised and edited and hence less likely to contain deviations.

3. Data Analysis

This paper focuses on the lexical item *work*. The main rationale behind this choice is the fact that *work* is one of the most frequent lexical items in the EA-STS (9th position), with 38 instances, and it occurs only nine times in the EA-PUB. This difference in the number of instances may be interpreted as an indication that the item behaves differently in the two subcorpora.

Work is used as a noun in the vast majority of instances of both subcorpora (89%): 34 occurrences in the EA-STS and eight occurrences in the EA-PUB. These are therefore the focus of the study and all instances of *work* as a verb or an adjective are discarded.

In both collections, *work* tends to be part of recurring lexico-grammatical patterns and refer to the study described in the abstract. Two instances in the EA-STS and three instances in the EA-PUB are exceptions. In these cases, *work* either refers to someone else's work or it

⁴ Further details of this web service can be found at <http://www.sketchengine.co.uk/auth/wbc/mycorp.cgi>.

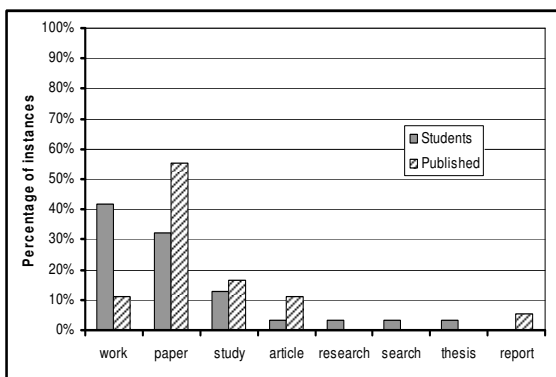
is related to the effort required to do a given task. Table 1 summarises these findings. VERB refers to any verb which appeared in the first position on the right of the work such as *aims*, *presents*, *shows*, etc. In pattern (iii), NOUN is used to indicate the various nouns which appeared in that particular position such as *aim* and *objective*. Optional items are indicated between brackets.

	Patterns	Students' Abstracts	Published Abstracts
i.	in this/my/ the present work	13	2
ii.	This/Our work VERB	13	2
iii.	the (main) NOUN of this/ the present work	6	1
iv.	<i>work</i> does not refer to the study described in the abstract	2	3
TOTAL		34	8

Table 1: Lexical patterns identified in the CCEA

Taking into consideration the low number of times *work* occurs in the EA-PUB subcorpus, our next step is to identify word(s) other than *work* which occur within these specific recurring lexical patterns in published abstracts. Here, we are interested in items which may also be used to refer to the study in question, even though they may not be exactly synonymous. Once these items are identified, we go back to the EA-STTS subcorpus and examine whether these same items are used by students in these specific contexts.

For instance, pattern (i) refers to the sequences *in this/the present/my* ***. Eight different lexical items are used: *work*, *paper*, *study*, *article*, *report*, *thesis*, *research* and *search*. *Search* occurs once in the EA-STTS subcorpus; it is most probably to be a mistranslation and meant to be *research*. Also, the pattern *in my work* occurs once in the EA-STTS subcorpus. Pattern (i) occurs 31 times in the EA-STTS and 18 times in the EA-PUB.

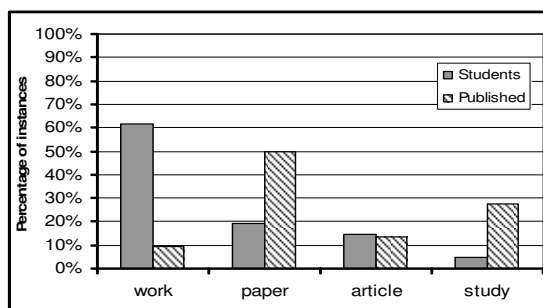


Graph 1: Pattern (i) - *in this/the present/my* ***

Work is the most frequent item in the EA-STTS (42%) and *paper* is the most frequent in the EA-PUB (56%).

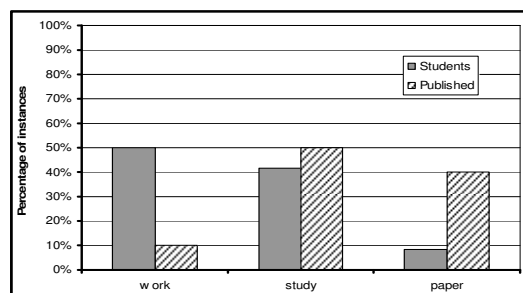
However, it is interesting to notice that *paper* appears in as much as 32% of instances in the EA-STTS whereas *work* occurs only twice (11%) in the EA-PUB. *Study* is the only item which is used in similar percentage in the two subcorpora. *Article* appears in a higher proportion in the EA-STTS than in the EA-PUB (11% compared to 3%). *Thesis*, *research* and *search* occur once in the EA-STTS and *report* appears once in the EA-PUB.

Pattern (ii) refers to the sequences *this/the present/the/our* *** VERB in the beginning a clause. In addition to *work*, three other items are used in this context: *paper*, *article* and *study*, yielding chunks such as *this paper examines* and *this study presents*. The pattern occurs 21 times in the EA-STTS and 22 times in the EA-PUB. In the EA-STTS, *work* accounts for the vast majority of instances (62%) and *paper* is the second most frequent item with 19% of occurrences. By contrast, *paper* occurs in 50% of the cases in the EA-PUB whereas *work* is only used twice (9%). *Study* is also more frequent in the EA-PUB (27%) in comparison with the EA-STTS (5%). *Article* is the only item which is used in similar percentage in both subcorpora (14%).



Graph 2: Pattern (ii) - *this/the present/our* *** VERB

By examining the lexical variations within pattern (iii) - *the (ADJ) NOUN of this/the present* *** -, we find that, in addition to *work*, *paper* and *study* are also used.



Graph 3: Pattern (iii): *the (ADJ) NOUN of this/the present* ***

As for the position of NOUN, for the overwhelming majority of instances in both subcorpora (83% in the EA-STTS and 70% in the EA-PUB), it is related to 'aim' (*aim*, *objective(s)* and *purpose*). We also find *result* and *contributions(s)* in the EA-PUB and *expectancy* in the

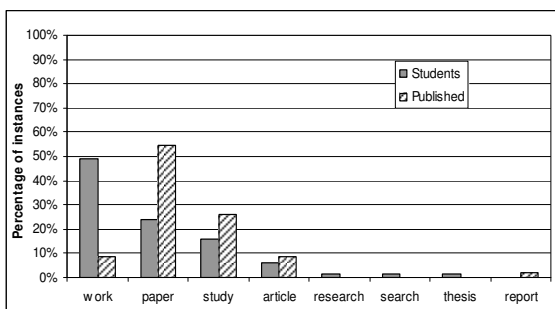
EA-STS, which is most probably a mistranslation of the Portuguese item *expectativa* (expectation). One instance in the EA-STS shows *object of this work*. Adjectives (ADJ) such as *main*, *primary* and *key* may occur before the noun. *Study* shows a high frequency of occurrence in both subcorpora: 42% of instances in the EA-STS and 50% of instances in the EA-PUB. However, *work* accounts for 50% of instances in the EA-STS and *paper* represents 40% of instances in the EA-PUB.

Taking into consideration the overall number of recurring patterns in the corpus, no striking difference is found between the percentages of each across the two subcorpora (Table 2). Patterns (i) and (ii) are used in fairly similar proportion in both collections; pattern (ii) is slightly more frequent in the EA-STS.

	Patterns	Students' Abstracts	Published Abstracts
i	<i>in this/the present/my</i> ***	31 (48%)	18 (36%)
ii	<i>this/the present/the/our</i> *** VERB	21 (33%)	22 (44%)
iii	<i>the</i> (ADJ) NOUN of <i>this/the present</i> ***	12 (19%)	10 (20%)
	TOTAL	64 (100%)	50 (100%)

Table 2: Lexical patterns identified in the CCEA

By examining the overall number of times each lexical item is used in each subcorpus, we find that *work* occurs in 51% of the instances in the EA-STS compared to 11% in the EA-PUB subcorpus. *Paper* is the most frequent item in the EA-PUB, representing 54% of all occurrences. It is the second most frequent item in the EA-STS (24%). *Study* shows a higher percentage of instances in the EA-PUB component (30%) in comparison with the EA-STS subcorpus (16%).



Graph 4: Overall number of times each lexical item is used in the EA-STS and EA-PUB subcorpora

4. Discussion

The figures above indicate a clear tendency of students to use the item *work* when referring to the study described in the abstract whereas published abstracts show a marked preference for the word *paper*. However, one cannot afford to ignore that the EA-PUB subcorpus is very

limited in size and hence it does not allow the researchers to make generalizations on the collocational behaviour of these two items in scientific abstracts.

An important point to make here is that our findings are very much consistent with the results revealed by Orasan (2001) on the use of the word *paper* in scientific abstracts. Orasan (ibid.) examines 917 abstracts (146,489 words) from the disciplines of artificial intelligence, computer science, biology, linguistics, chemistry and anthropology. *Paper* is frequently used as the subject of verbs such as *present* (62 times), *describe* (50), *be* (45), *introduce* (15) and tends to yield patterns like *this paper presents* (44) or *this paper describes* (39). Another clear pattern is the sequence *in this paper*, which occurs 143 times in the corpus. Taking into consideration the items used to refer to the study described in the abstract, Orasan (ibid.) shows that *paper* is the most frequent option (53% of the instances) in relation to other items such as *study*, *research* and *work* (Table 3).

Item	Number of instances	% of instances
paper	499	53%
study	170	18%
research	154	17%
work	111	12%
TOTAL	934	100%

Table 3: Number and percentage of instances of the items used to refer to the study in question (Orasan, 2001)

Similar to our study, Orasan (2001) also concludes that the high frequency of these patterns in abstracts is not by chance but instead that it is a strong indication that they are frequently used in this specific context of abstracts. However, it is important to bear in mind that Orasan (ibid.) uses a corpus which includes abstracts from various disciplines and does not focus on the specific recurring lexical patterns that we are interested here. Thus, in order to validate our findings and be able to obtain a clearer picture of how *work* and *paper* are used by the academic discourse communities in question, we necessarily need access to a reference corpus of abstracts which matches the specifications of the texts included in the CCEA. Here, we use a corpus consisting of 723 scientific abstracts from the disciplines of physics (369) and pharmaceutical sciences (354) (Genoves et al., 2007). All abstracts were collected from reference journals of these two disciplines such as *Physical Review Letters (A-D)*, *Science*, *Nature* and *Biotechnology Progress*. The overall size of the corpus is 115,913 words (tokens).

We first focus on the three recurring patterns discussed above and look at number of instances in which *work* and *paper* are used. The analysis is extended to include other items which may also be used to refer to the study described in the abstract.

Unlike the results discussed above, we find that, in the reference corpus, *study* is by far the most frequent item, with 51% of the instances (Table 4). *Paper* is the second

most frequent item, accounting for 21% of the occurrences. *Work* is used in 17% of the instances. The reference corpus also shows that the item *review* can also be used in these specific contexts; however, it does not occur in the CCEA.

	Item	Number of instances	% of instances
1.	study	76	51%
2.	paper	31	21%
3.	work	26	17%
4.	review	8	5%
5.	article	4	3%
6.	report	3	2%
7.	research	1	1%
	TOTAL	149	100%

Table 4: Number and percentage of instances for the lexical items in the reference corpus

The reference corpus also reveals several variations of the patterns under analysis. For instance, pattern (i) is the most frequent pattern in the reference corpus with 101 instances. In addition to *in this* and *in the present*, which account for the vast majority of instances – 85 and 10 respectively, we also find *in our* (3), *in the current* (2) and *in the* (1). Seven lexical items are used to refer to the study in question: *study*, *paper*, *work*, *article*, *review*, *report* and *research* (Table 5). *Study* is the most frequent item, accounting for 53% of the instances. *Paper* (22%) is slightly more frequent than *work* (16%).

	Item	Number of instances	% of instances
1.	study	54	53%
2.	paper	22	22%
3.	work	16	16%
4.	article	4	4%
5.	review	3	3%
6.	report	1	1%
7.	research	1	1%
	TOTAL	101	100%

Table 5: Items within pattern (i) in the reference corpus

Pattern (ii) – *This *** VERB* – occurs 38 times in the corpus (Table 6). In addition to *this*, which appears in 23 instances (61%), the following appears before our search item, mentioned in order of frequency of occurrence: *the present* (5), *our* (5), *the current* (3) and *the performed* (1). One instance shows *the study presented here* VERB. All these occurrences have been considered as variants of pattern (ii). Five different lexical items are used to refer to the study in question, in order of frequency: *study*, *work*, *paper*, *review* and *report*. *Study* is again the most frequent item in pattern (ii), accounting for 40% of the instances.

	Item	Number of instances	% of instances
1.	study	15	40%
2.	work	9	24%
3.	paper	7	18%
4.	review	5	13%
5.	report	2	5%
	TOTAL	38	100%

Table 6: Items within pattern (ii) in the reference corpus

Pattern (iii) occurs 10 times in the corpus (Table 7). Only three items appear within this pattern. Here again, *study* is the most frequent item, representing 70% of the instances (Table 6).

	Item	Number of instances	% of instances
1.	study	7	70%
2.	paper	2	20%
3.	work	1	10%
	TOTAL	10	100%

Table 7: Items within pattern (iii) in the reference corpus

As can be seen, in the reference corpus, *study* is the most frequent item within the three patterns. *Paper* is the second most frequent item in patterns (i) and (iii) and *work* comes third in the frequency ranking. For pattern (ii), it is interesting to notice that *work* is more frequent than *paper* (24% compared to 18%). By contrast, in the EA-STS, *work* is the most frequent item in the three patterns whereas *study* shows a high percentage of instances for pattern (iii) only. For patterns (i) and (ii), *paper* is the second most frequent item.

In terms of percentage of instances for each pattern, we notice that the reference corpus shows a strong preference for pattern (i) (68%, Table 8). This same tendency is seen in EA-STS abstracts, although not as marked (Table 2).

	Patterns	Reference Corpus
i	<i>in this/the present/current/our/the ***</i>	101 (68%)
ii	<i>this/the present/current/performed/our *** VERB</i>	38 (25%)
iii	<i>the (ADJ) NOUN of this/the present ***</i>	7 (10%)
	TOTAL	149 (100%)

Table 8: Number and percentage of instances for each pattern in the reference corpus

5. Final Remarks

This paper has examined the collocational behaviour of item *work* in abstracts written by Brazilian graduate students as opposed to abstracts collected from published papers of the same discipline. Relevant differences were found between the two subcorpora. Taking into

consideration the same specific contexts, the former displayed a strong preference for the item *work* whereas the latter showed a clear tendency to use *paper*. The results were validated by examining the identified lexical patterns in a reference corpus of English abstracts. *Study* was by far the most frequent item in the reference corpus. *Paper* came second, showing a slightly higher proportion than *work*.

Given that our long-term objective is to provide support to the development of course materials and computer-aided writing tools to aid Brazilian graduate students in writing scientific papers in English, this study took a step further and searched for items other than *work*, *paper* and *study* which may also be used to refer to the study described in the abstract. We also looked at instances which could be regarded as variants of the identified lexical patterns.

Thus, in addition to contrasting collocational patterns of *work* in abstracts written by students and published abstracts, this study has identified various lexical items used in specific lexical patterns as well as described their usage according to frequency. These findings can be incorporated into course materials and computational resources. This would enable us to raise students' awareness of the most typical lexical patterns used by their academic discourse community while, at the same time, it allows us to draw students' attention to the various other alternatives available to them when writing academic English.

6. Acknowledgements

We would like to thank our Brazilian sponsors FAPESP and CNPq (Brazil) for providing the financial support needed in this research.

7. References

- Aluísio, S.M., Barcelos, I., Sampaio, J., Oliveira Jr., O. (2001). How to Learn the Many Unwritten 'Rules of the Game' of the Academic Discourse: A Hybrid Approach Based on Critiques and Cases. In *Proceedings of the IEEE International Conference on Advanced Learning Technologies*. Madison, Wisconsin. Los Alamitos, CA: *IEEE Computer Society*, 1, pp. 257--260.
- Aluísio, S.M., Gantenbein, R.E. (1997). Educational Tools for Writing Scientific Papers. In *VIII Simpósio Brasileiro de Informática na Educação*, ITA, pp. 239--253.
- Aluísio, S.M., Oliveira Jr., O. (1995). A Case-Based Approach for Developing Writing Tools Aimed at Non-native English Users. In *Lecture Notes in Artificial Intelligence* 1010, pp. 121--132.
- Aluísio, S.M., Schuster, E., Feltrim, V.D., Pessoa Jr., A., Oliveira Jr., O.N. (2005). Evaluating Scientific Abstracts with a Genre-specific Rubric. In *The 12th International Conference on Artificial Intelligence in Education - AIED*, 18-22 July, Amsterdam.
- Anthony, L. (2006). Developing a Freeware, Multiplatform Corpus Analysis Toolkit for the Technical Writing Classroom. In *IEEE Transactions on Professional Communication*, Vol. 49, No. 3, pp. 275--286.
- Baroni, M., Kilgarriff, A., Pomikálek, J., Rychlý, P. (2006). WebBootCaT: instant domain-specific corpora to support human translators. In *Proceedings of EAMT 2006*. Oslo, pp. 247--252.
- Genoves Jr., L., Lizotte, R., Schuster, E., Dayrell, C., Aluísio, S. (2007). A two-tiered approach to detecting English article usage: an application in scientific paper writing tools. In *Proceedings of the International Conference RANLP 2007*. Borovetz, Bulgaria, Sep 26, 2007, pp. 225--239.
- Hyland, K. (2008a). Academic Clusters: Text Patterning in Published and Postgraduate Writing. In *International Journal of Applied Linguistics*, 18(1), pp. 41--61.
- _____ (2008b). As Can Be Seen: Lexical Bundles and Disciplinary Variation. In *English for Specific Purposes*, 27, pp. 4--21.
- Lee, D. and Swales, J. (2006). A Corpus-Based EAP course for NNS doctoral students: Moving from Available Specialized Corpora to Self-compiled Corpora. In *English for Specific Purposes*, 25, pp. 56--75.
- Narita, M., Kurokawa K., Utsuro, T. (2003). Case Study on the Development of a Computer- Based Support Tool for Assisting Japanese Software Engineers with their English Writing Needs. In *IEEE Transactions on Professional Communication*, Vol. 49 (3), pp. 194--209.
- Orasan, C. (2001). Patterns in scientific abstracts. In *Proceedings of Corpus Linguistics 2001 Conference*. Lancaster University, Lancaster, UK, pp. 433--443
- Scott, M. (2004). *WordSmith Tools* version 4. Oxford: Oxford University Press.
- Sinclair, J. (1991). *Corpus Concordance and Collocation*. Oxford: Oxford University Press.
- _____ (2003) *Reading Concordances*, London: Pearson Education Ltd., Longman.
- Swales, J.M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- _____ (2004) *Research Genres. Explorations and Applications*. Cambridge: Cambridge University Press.
- Swales, J.M., Feak, C.B. (2000). *English in Today's Research World: A Writing Guide*. Michigan: The University of Michigan Press.
- Thompson, P. (2001). Looking at Citations: Using Corpora in English for Academic Purposes. In *Language Learning and Technology* 5(3), pp. 91--105.
- _____ (2006). Assessing the Contribution of Corpora to EAP Practice. In Z. Kantaridou, I. Papadopoulou & I. Mahili (Eds.), *Motivation in Learning Language for Specific and Academic Purposes*. Macedonia: University of Macedonia [CD-ROM].
- Weissberg, R., Buker, S. (1990). *Writing up Research: Experimental Research Report Writing for Students of English*. Englewood Cliffs (NJ): Prentice Hall Regents.