

UNPACKING L2 WRITING RESPONSES: A CORPUS-BASED STUDY ON TEACHER FEEDBACK TO STUDENT WRITING

by Tamás Merckle

University of Pécs, Pécs, Hungary

tamas[dot]merckle[at]hotmail[dot]com

Abstract

Over the past decades, corpus linguistics has become widespread in studies dealing with applied linguistics. Teachers have become acquainted with corpus linguistic methods and are using computer technology in their professional practice. A well-known example is the learner corpora with which researchers managed to attain invaluable results concerning various aspects of learner language. This article, however, presents a new field within corpus linguistics: the teacher corpus. A corpus on teacher language (instead of learner language) has a lot to offer in terms of methodology and pedagogy.

Introduction

Computer corpora have gained a foothold in the field of language teaching and other language related fields so much so that they have permeated all walks of language educational life. This means that, for instance, good dictionaries are inconceivable without relying on a renowned collection of spoken and written texts (e.g. Hornby, 1997; Sinclair, 1987a; 1987b). Corpus-based interlanguage L2 studies have resulted in effective pedagogical approaches (Kaszubski, 1998; Ringbom, 1998). The world of corpora can promote effective self-directed learning (Szirmai [2001], for example, presents an account on the corpus-based software *Contexts*). On the other hand, speakers of a foreign language can analyse their own language both in and outside classroom in order to get authentic data about what they or others are doing well and what needs to be improved. These are but a few possibilities that corpus linguistics has to offer.

Some teachers of ESL writing began to use corpus linguistics, not surprisingly, as they recognized the potential benefits of storing and using language for pedagogical purposes. The field of Second Language Writing (SLW) is drenched with corpus-related methodology touching upon every field such as English for Academic Purposes (e.g. Coxhead & Byrd, 2007), feedback-giving (e.g. Horváth, 2002), data-driven learning (e.g. Granger & Tribble, 1998) and several others. The present paper deals with the teacher corpus: a collection of

language data produced by teachers of EFL. The corpus is a collection of teacher comments that can be found at the bottom of the student papers, usually in red.

The paper is organised as follows: the background section will shed some light on the rationale of teaching second language writing: the implications for classroom use and the aims teachers and students may have. Evaluating student papers and genre-pedagogy will be in our focus as these have influential role from our corpus's point of view. Then, after the brief consideration of the background issues on which this corpus is based, the presentation and practical application of the PAC corpus will follow. Finally, the discussion of the results and the conclusion will end the paper.

Background: objectives and implications of second language writing

Teachers often have the responsibility of having to prepare their students for externally validated exams such as the school leaving examination. This kind of examination has its own requirements concerning writing – at both advanced and ordinary levels. The requirements are explicitly stated and are no secret, therefore the recipe seems to be easy: practice, practice and practice. This might be true, but one might encounter the potential pitfall of plunging into excessive writing practice without proper preparation on the teacher's part. Ferris (2007) argues that there are some options for teachers wishing to apply effective responses. The first is that they “consult the course rubric or grading criteria if such instruments exist” (p. 170). The school leaving exam, obviously, has this grading scheme, and it is available for teachers and students alike (see [Appendix A](#) for the writing assessment scheme of the Hungarian school leaving exam evaluation sheet).

It would be reasonable to say that if teachers see the aspects they will evaluate (the correctness of grammar, vocabulary, style, etc.) and the weight each aspect carries compared to other aspects, they might organise their work accordingly. If the school leaving exam requires students to write a letter to their friends in which they turn down an invitation, or a formal letter which enquires about summer courses in Britain, one might anticipate that the essential points such as formality and informality, paragraphing, vocabulary or grammar will be covered as well. However, it seems that the focus of attention has shifted to the importance of grammar (e.g. Zamel, 1985). A contrastive rhetorical study between L1 Hungarian and L2 English student writers showed that the observed Hungarian teachers of English, when it comes to evaluating writing, often resort to grammatical correctness as their main focus (Kiszely, 2003). This way, surface level correction can carry more weight in the teacher

feedback compared to other aspects such as style, and the writing evaluation practice can become lopsided.

It is easier to correct deviations of hard-and-fast rules than correct elusive aspects such as style. Evidently, every teacher would immediately draw a red pen when seeing the sentence *She didn't give no kisses to him*, because they know that formal rules do not allow the speakers of English to use two negatives in one sentence. However, there are no exact rules for style, even, for 'the impression on reader' (which is one of the school leaving exam categories). Therefore, teachers rarely put out from the safe shores of grammar.

According to Ferris (2007), teachers have another tool to rely on besides grade rubrics: the specific requirements of a given task. Obviously, letter writing requires other strategies than writing an argumentative essay. In this respect, Hungarian teachers are doing what they are expected to do: make the students practise the genre they are expected to write in the school leaving exam. This is what Second Language Acquisition calls washback effect (Ellis, 1997). The rule of it is simple: if *a* is required, but *b* is not, then as a teacher I will make my students practice *a*, because I want them to know that, and no one will enquire about whether they know *b*. For instance, argumentative essay is virtually unknown in the Hungarian school leaving exam as only letter writing is required. This is not to say that no information is disseminated about argumentative essays in classes, but dauntingly less compared to letter writing.

Thirdly, "the teacher might also provide feedback tailored to the needs and progress of individual students" (Ferris, 2007, p. 170). This is what we call 'differentiated pedagogy' in Hungary, which is a panacea if one has time for it, as not only needs-based annotation is required, but needs-based writing lessons are also necessary.

The introduction of genre-based pedagogies would respond to some of the concerns of those teachers that see beyond the sway of letter writing and experiment with teaching various genres. Identifying these genres that may be beneficial for our students empowers teachers to teach them effectively. Knowing what to teach has consciousness-raising function, as the teacher's responsibility here is to "illuminate the genres that matter to students" (Hyland, 2007, p. 160), as well as to take into consideration the following methodology-related principles:

- *Writing as a social activity*: writing is a purposeful activity where intentions, purposes and audience are equally important.
- *Writing is needs-oriented*: teachers should keep in mind what kind of writing is needed in the students' future situations.

- *Explicit outcomes and expectations are needed*: teachers should be explicit about what is expected from students and what the intended way of achieving that during the course.
- *Learning to write is a social activity*: teaching is a sequence of “scaffolded developmental steps in which teachers and peers play a major role” (Hyland, 2007, p. 153, the principles are based on Hyland, [2007, p. 152]).

Teaching writing with the help of models is a common practice. The course books are filled with authentic writing material from which students can infer the characteristics of a genre. With appropriate teacher instruction, students can render these features into their own writing practice. Thus, it would be logical to think that L1 materials as models are the main facilitators of effective writing instruction. That is, using these should enable students to create texts that are similar to the model texts in terms of characteristics and quality. The situation is not as simple as that, however. Gilquin *et al.* (2007), when talking about the relationship between L1 corpora and English for Specific Purposes (ESP) as well as English for General Purposes material design, enunciated that there are several problems with it, namely that (a) they are hardly corpus-informed, (b) the corpora that are used in practice are native corpora, which are different in terms of errors: native novice writers experience different difficulties than novice non-native writers. A corpus can provide students with lots of authentic data, but differentiating between the needs of L1-L2 as well as novice and expert student writers is a key when using corpora as models.

The division between learner and native corpora is not a new idea. The International Corpus of Learner English (started by Granger, 1993; 1994; 1996) is probably the best known example of a collection of learner language. Teachers and students alike that are part of a corpus research project can familiarize themselves with the writing processes, the features of various genres and contribute to the enhancement of the texts stored through the synthesis of their previous work and the knowledge they inferred from corpus analyses. Horváth (2002, p. 192, emphasis added) suggested that not only students can learn from corpus projects but teachers as well, because “[teachers] mark up text by students, who, in turn, attempt to understand, critique and apply the comments. Yet this part of the pedagogical process is often *lost to research and pedagogy* when the teacher shares written comments with the student writer.” The PAC corpus, my own collection of secondary EFL teacher comments, aims to show one way we can prevent the loss of valuable research and pedagogical information.

Corpus design principles

A corpus cannot be set up without having well-defined factors such as “purpose, language community, text types, representativeness, encoding and storage facilities” (Horváth, 2001, p. 101). Currently, the PAC corpus has the following characteristics:

- the first Hungarian corpus of teacher annotations, containing three schools’ practice in connection with EFL writing. The scripts of the three schools are transferred into three subcorpora, named *Budapest1*, *Budapest2* and *Budapest3* (these subcorpora make the PAC corpus);
- the language community is made up of EFL teachers, currently no other criteria (age, experience, number of years spent with teaching, etc.) are identified;
- the corpus is entirely anonymous, does not reveal the name of the teachers and schools;
- incorporate teacher comments for student writing, as teachers use marginal comments and end-notes, it must be mentioned that only end-notes are scrutinized here;
- contains 5414 words;
- allows a variety of generalisations;
- the annotations presented are voluntarily submitted.

My aim is to extend the corpus so that it becomes representative to the teacher community of a school, or teachers within a school. The current size of the corpus does not achieve the intended minimum, however, I assume this smaller size is sufficient to come up with relevant conclusions and to show the way it works in practice. Several studies show (e.g. Granger, 1993, Mark, 1998) that smaller corpora are enough to help to draw firm conclusions.

Clearly, permission was needed to start gathering the corpus. In most of the cases, the teachers were willing to submit their annotations together with the students’ texts. Then I typed the handwritten annotations into the appropriate subcorpora and saved the Word file. The subsequent steps were simple: the typed text allowed me to copy-paste the large amount of text into the CL programs found at [Compleat Lexical Tutor](#), or [LexTutor](#) (Cobb, 2007). This website offers students and teachers a range of programs that help analyse texts from various viewpoints.

As for the proportion of the words that make up the corpus currently, Table 1 shows the distribution of the type-token ratio. Tokens stand for the actual number of words that can be found in a corpus. Type is a term used for different words in a corpus (Leech, 1997; Kennedy, 1998). Ratio refers to the proportion calculated from type and token. Although the numbers are relatively low, it can be stated that the texts are rather homogeneous.

Table 1. Type-token ration of the three subcorpora within the PAC corpus

Subcorpus	Tokens	Types	Ratio
<i>Budapest1</i>	1691	373	4.533
<i>Budapest2</i>	1828	312	5.858
<i>Budapest3</i>	1895	371	5.107

Having presented an overall picture of the PAC corpus, I now intend to show how this corpus can be used for research purposes. It should be pointed out that although the full corpus contains three subcorpora (as can be seen in Table 1), the following analysis will present two of these (*Budapest1* and *Budapest2*).

Design of the study

1. Research question

As attested previously, the teacher corpus can respond to several questions teachers may have, among them to the following one, which is in the focus of the present research project.

- *Do the comments teachers make about student writing resonate with student expectations?*

In order for the teachers to be effective, it is crucial to know what the individual learner needs, with what she struggles the most, and with which points she needs a helping hand. To acquire data on this subject, questionnaires will be used to collect quantitative information about their needs. The question will be answered by means of the questionnaires as well as CL methods. The comparison of these will highlight the (non-)existing disparity between expectations.

2. Research methods

To answer the above question, we need to have two sets of data. One of these is the practice of the teachers, whereas the other is the expectations of the students. These will be analysed as follows: (1) the two schools have their own subcorpus containing the teacher comments. A [Word-frequency analysis](#) will yield a rank containing all the words that appear in the corpus starting from the most frequent and listing all the words rank ordered according to frequency.

(2) Then, these words are put into the five categories that a school-leaving exam offers (grammar, vocabulary, style/tone, organisation, content). For instance, the word *grammar* will belong to the *grammar* group, *paragraphs* will go into *organisation*. Thus, the five categories will have words with frequency numbers. This will allow us to infer the main focus of the teachers at a given school. (3) Expectations are enquired by means of a questionnaire (see [Appendix B](#)). In this short survey, students are asked to rank order these categories based on their needs – the most problematic will come first and they go backwards. This will result in a list that can be used to compare against the actual practice of the teachers.

3. Participants

Two subcorpora (*Budapest1* and *Budapest2*) will be analysed here. The number of the teachers that contributed to the two corpora is eleven (six and five, respectively). As for the students, ten students were chosen from each of the two schools (altogether twenty), and were asked to fill in the questionnaire. All students belong to classes in which the EFL teachers taught that contributed to the corpus. That makes it possible to compare the expectations and the practice of teachers.

The research proper

1. Corpus analyses

In terms of the actual analysis applied here, I have made use of the [Word-frequency Indexer](#), as suggested above. Its simple procedure allows us to gather quantitative data on the teacher feedback. Thus, the present analysis relies only on statistical data and to this end it can result in only statistical results. However, as far the educational benefits are concerned, a lot can be learnt from a statistical comparison as well, regardless of their seemingly impersonal nature. Without any predilection, we can have exact statistical data on teacher practice. As this comparison can be done quickly, useful and applicable information is available for teachers to learn from.

A limitation of the procedure applied is that it does not check the context in which the word appears. This, however, causes no problems. When we have, for example, the word *topic*, one would say, it belongs to a content-related annotation. But if that word in a particular context does not represent content problem, it may have a fallacious effect as it increases the number of content annotations in our list. Consider the following sentence:

*You should include a **topic** sentence in the second paragraph.*

The word-frequency indexer will only tell us that the word ‘topic’ appears in the text, but does not tell us what it refers to. The word *topic* on its own refers only to a content annotation, however, the context (*topic sentence*) tells us that it is actually an organisation comment. If we had not checked the context of the word, we would have put it into the ‘content annotation’ category, which would be a mistake.

However, the ensuing example on *Budapest1* corpus will show a way to overcome this problem. The list presented is the word-frequency result. The percentages show the amount that a word covers in the whole corpus. *Letter*, for instance, constitutes 0.92 percent of the whole corpus.

0.92%	LETTER	0.25%	WORDS
0.67%	TOPIC	0.17%	INFORMAL
0.59%	MISTAKES	0.17%	INTRODUCTION
0.59%	PARAGRAPH	0.17%	LOGICAL
0.50%	GRAMMAR	0.17%	PARAGRAPHING
0.42%	FORMAL	0.17%	STYLE
0.34%	SPELLING	0.17%	TENSES
0.34%	STRUCTURE	0.17%	VOCABULARY
0.25%	GRAMMATICAL		

These words are put into the five categories (grammar, vocabulary, organisation, content, style) as described above.

Clearly, what we see is that significant importance is placed on topic, in other words, content-related problems. In order to make sure topic is not a part of the phrase ‘topic sentence’, thus it is not organisational hint, Lextutor’s online concordancer (<http://www.lex tutor.ca/concordancers/>) was used. This program shows a chosen word in the context in which it was uttered by creating concordancing lines of the text. Consider the following example with *topic*, taken now from the *Budapest1* corpus and shown as the concordancer program presents them.

968. Let them tell us one **TOPIC** *each* otherwise the
 969. good analogies are included nice **TOPIC** sentences. *I* would rephrase
 970. Tell your reader why your **TOPIC** *is* interesting what you know
 971. tintrouction It introduces the **TOPIC** *so* when writing an
 972. paragraph elaborated the **TOPIC** *well* I would add more details
 973. paragraph they have a mixture of **TOPIC** *which* are related to the
 974. The **TOPIC** *you* presented is clear

Line 969 refers to an organisational piece of advice, however, the other ones are clearly topic-related. The same procedure was applied to ‘mistakes’, as this word can refer to a number of different mistakes. Having checked the sentences with the word ‘mistake(s)’, it is revealed that these are mainly grammar- and spelling-related hints (with one being vocabulary mistake):

536. writing to apply see your **MISTAKE**. *Otherwise* it is good. Just
 537. Keep it up Minor spelling **MISTAKES** *and* some grammatical
 538. with some grammatical **MISTAKES**. *Be careful* does *and*
 539. and some grammatical **MISTAKES** *here* and *there* but on
 540. you have lots of spelling **MISTAKES**. *If* you use word-processor
 541. Apart from some grammar **MISTAKES** *see* them this is a wonderful
 542. attention to the vocabulary **MISTAKES** *strong* wind
 543. Apart from the grammatical **MISTAKES** *you* presented a very good work

The next aspect in the frequency line is organisation [with the word *structure*], followed by vocabulary [words] and style [style] is the last in the list.

Budapest2 subcorpus has the following features:

0.84%	STRUCTURE	0.36%	GRAMMAR
0.60%	CONCLUSION	0.24%	AUXILIARIES
0.60%	FORMAL	0.24%	BODY
0.60%	PARAGRAPHS	0.24%	CONTENT
0.60%	TOPIC	0.24%	INFORMAL
0.48%	INTRODUCTION	0.24%	STYLE
0.48%	VOCABULARY	0.24%	WORDS

It is clear that organisation is of crucial importance in this corpus [structure and conclusion], tone is analysed [formal], and content as well [topic]. It should be pointed out that the word *topic* was analysed by a concordancer to ensure it is a content-related word. Lexis is also touched upon in the comments [vocabulary], immediately followed by grammar [grammar and auxiliaries]. Style falls out of the center of attention but is still in the list [style].

Results of the questionnaires

We now have data about the frequency of the aspects appearing in the annotations. The question is whether the practice outlined above is in line with student expectations or not. The items in the questionnaire were designed in a way that the questions exclude each other from

being at the same rate, which means that a hierarchy can be drawn concerning problem points. When designing the questionnaire, it was also a principle to apply only six options on the Likert scale, so that students are forced to take a stand. The items were inverse-questions of each other, which is expected to provide us with an easily analysable order of aspects. These orders will be compared to the CL lists mentioned above, so that the discrepancy between needs and practice can be spelled out. Table 2 and 3 show the results of the questionnaire phase, the former deals with the *Budapest1*, whereas the latter refers to *Budapest2*.

Table 2. Answers of the Budapest1 students

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Overall
Item1 (grammar)	5	6	4	5	5	6	6	6	5	6	5.4
Item2 (vocabulary)	5	3	3	3	4	3	4	2	4	5	3.6
Item3 (organisation)	3	5	6	6	6	6	5	5	4	5	5.1
Item4 (style/tone)	1	2	1	1	1	2	1	1	1	1	1.2
Item5 (content)	4	5	5	4	5	5	6	5	6	6	5.1

The overall value indicates the practice as students view it: according to them, grammar is a hot issue and is desired to handle (item1, value 5.4), whilst organisation and content-oriented hints compete for the second place (item3 and item5, value 5.1). Vocabulary is slightly above average (item2, value 3.6), and tone fell out of the scope the students questioned (item4, value 1.2).

The other questionnaire shows us the following result:

Table 3. Answers of the Budapest2 students

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Overall
Item1 (grammar)	6	6	6	4	6	6	6	6	5	6	5.7
Item2 (vocabulary)	4	5	4	5	5	4	5	4	4	3	4.3
Item3 (organisation)	5	4	5	3	4	4	4	3	6	6	4.4
Item4 (style/tone)	2	2	3	2	2	2	2	1	3	3	2.2
Item5 (content)	6	6	5	6	5	3	6	5	5	5	5.2

The following results can be perceived by looking at Table 3: students of *Budapest2* see grammar feedback as a constant part of the comments (item1, value 5.7), and content is of similar importance (item5, value 5.2). Vocabulary and organisation are also required to be touched upon by the teachers (item2 and item3, value 4.3 and 4.4, respectively). As opposed to the above aspects, tone is somewhat neglected (item4, value 2.2).

Discussion

The previous sections have presented the language analysis of two teacher subcorpora of the PAC corpus as well as the expectations of the students. We are now in a position to compare the results so that we can come up with edifying upshots.

At the first glance it seems that the *Budapest1* subcorpus focuses mainly on grammar and content. These are the top priority for the teachers at that particular school. Organisation follows the first two aspects, and tone and style are hardly considered. The expectations, however, are somewhat similar to practice. Students require grammar, content and organisation responses – the ones that are considered in the responses. For some reason, tone scored low among students – owing perhaps partly to the fact that other factors are considered

more important, or maybe because they were not really aware of what is meant by tone. All in all, what we see is that in the school of *Budapest1* corpus expectations and practice meet.

One can find an opposite trend in *Budapest2* corpus. Although similar percentages could be seen concerning aspects, it seems that the following order can be delineated: teachers go for organisation first, then tone and content. After them, they provide comments on vocabulary, and grammar is, interestingly, fifth in the list. This opposes student expectations, namely, that they would appreciate grammar instructions first, then content, vocabulary, organisation and style, respectively. This, of course, does not mean that the teachers at that particular school provide ineffective comments, but it may be an instructive lesson for teachers wanting to apply needs-based pedagogy. It should also be pointed out that it is early to come up with general conclusions, these results are the first steps towards a large-scale conclusion.

Conclusion and implication for future studies

The research project presented here has attempted to investigate the potential discrepancy between teachers' practice and students' needs. The first step of a needs-based pedagogy is that the teachers know what is needed and provide remedy and help for that particular problem point. As has been shown, this is the case concerning *Budapest1* subcorpus. As opposed to *Budapest1*, *Budapest2* fails to provide needs-based pedagogy. However, it should be mentioned that the results and conclusions presented here are limited to this study. Repeating it at a later time might result in different outcomes.

In the future we need to extend the scope of the corpus to other schools and aim at attaining a higher number of scripts and token numbers. This corpus is a small one to infer general conclusions, but, despite its size, it can already be used for practical purposes already. The present study may serve as a new signpost for future directions and utilisation of corpora.

References

- Bernstein (1990). *Class, Codes and Control. The Structuring of Pedagogic Discourse* (vol. 4.). London: Routledge and Kegan Paul.
- Cobb, T. (2007). *Compleat Lexical Tutor*. Available at <http://www.lextutor.ca>.
- Coxhead, A., & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing*, 16, 129-147.
- Ellis, R. (1997). *Second Language Acquisition*. Oxford: Oxford University Press.
- Ferris, D. (2003). *Response to Student Writing*. New York: Lawrence Erlbaum Associates.

- Ferris, D. (2007). Preparing teachers to respond to student writing. *Journal of Second Language Writing*, 16, 165-193.
- Frequency Home*. Retrieved July 15, 2008, from <http://www.lex tutor.ca/freq>.
- Gilquin, G., Granger, S., & Paquot, M. (2007). Learner corpora: The missing link in EAP pedagogy. *Journal of English for Academic Purposes*, 6, 319-335.
- Granger, S. (1993). International Corpus of Learner English. In J. Aarts, P. de Haan & N. Oostdijk (Eds.), *English Language Corpora: Design, Analysis and Exploitation*. (pp. 57-71). Amsterdam: Rodopi.
- Granger, S. (1994). The learner corpus: A revolution in applied linguistics. *English Today*, 10 (3), 25-29.
- Granger, S. (1996). Learner English around the world. In S. Greenbaum (Ed.), *Comparing English World-wide* (pp. 13-24). Oxford: Clarendon Press.
- Granger, S., & Tribble, C. (1998). Learner corpus data in the foreign language classroom: form-focussed instruction and data-driven learning. In S. Granger (Ed.), *Learner English on Computer* (pp. 199-209). London: Longman.
- Hornby, A. S. (1987). *Oxford Advanced Learner's Dictionary of Current English*. Oxford: Oxford University Press.
- Horváth, J. (2001). *Advanced Writing in English as a Foreign Language: A Corpus-based Study of Processes and Products*. Pécs: Lingua Franca Csoport.
- Horváth, J. (2002). Pedagogical annotation of learner corpora. In B. Hollósy & J. Kiss-Gulyás (Eds.), *Studies in Linguistics – A Supplement to the Hungarian Journal of English and American Studies* (pp. 192-210). Debrecen: University of Debrecen.
- Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing*, 16, 148-164.
- Ringbom, H. (1998). Vocabulary frequencies in advanced learner English: A cross-linguistic approach. In S. Granger (Ed.), *Learner English on Computer* (pp. 41-52). London: Longman.
- Kennedy, G. (1998). *An Introduction to Corpus Linguistics*. Rowley, Mass.: Newbury House Publishers.
- Kiszely, Z. (2003). *Students' Writings in L1 Hungarian and L2 English: Rhetorical Patterns, Writing Processes and Literacy Backgrounds*. Unpublished PhD dissertation. University of Pécs.
- Kaszubski, P. (1998). Enhancing a writing textbook: A national perspective. In S. Granger (Ed.), *Learner English on Computer* (pp. 172-185). London: Longman.
- Keh, C. L. (1996). Feedback in the writing process: A model and methods for implementation. In T. Hedge & N. Whitney (Eds.), *Power, Pedagogy & Practice* (pp. 294-306). Oxford: Oxford University Press.
- Leech, G. (1997). Introducing corpus annotation. In R. Garside, G. Leech & T. McEnery (Eds.), *Corpus Annotation: Linguistic Information from Computer Text Corpora* (pp. 1-18). London & New York: Longman.
- Leki, I. (1990). Coaching from the margins: Issues in written response. In B. Kroll (Ed.), *Second Language Writing: Research Insights for the Classroom* (pp. 57-68). Cambridge: Cambridge University Press.
- Mark, K. (1998). Applying learner corpus insights to the design of instructional manuals. In P. Lewis (Ed.), *Exploring Relationships in CALL* (pp. 119-125). Tokyo: JALT CALL N-SIG.
- Online Concordancers*. Retrieved July 15, 2008, from <http://www.lex tutor.ca/concordancers>.
- Sinclair, J. (Ed.). (1987a). *Collins COBUILD English Language Dictionary*. London: Collins.

- Sinclair, J. (Ed). (1987b). *Looking up: An Account of the COBUILD Project in Lexical Computing and the Development of the Collins COBUILD English Language Dictionary*. London and Glasgow: Collins ELT.
- Szirmai, M. (2001). *The Theory and Practice of Corpus Linguistics*. Debrecen: Debreceni Egyetem Kossuth Egyetemi Kiadója.
- Zamel, V. (1985). Responding to student writing. *TESOL Quarterly*, 19 (1), 79-101.

Appendix A

Source: www.om.hu [Hungarian Ministry of Education, English language matura]

A feladat teljesítése, a megadott szempontok követése [fulfilment of the task, following the instructions]	5	
Hangnem, az olvasóban keltett benyomás [impression on the reader]	2	
Szövegalkotás [text creation]	4	
Szókincs, kifejezés mód [vocabulary, self-expression]	5	
Nyelvhelyesség, helyesírás [language accuracy, grammar]	5	
Íráskép [layout]	1	
Összesen [overall]	15	

Appendix B

Dear student,

I am conducting a research project at the University of Pécs, and in order for me to have as accurate a picture about my topic as possible, I need your help. I would like to ask you to consider the following questions and circle the answer you feel is correct. There are no good or wrong answers, with these questions I would like to receive information about you as a student.

This questionnaire is about your English writing experiences either in class or at home. Think of your habits, the way you write, the way your teacher helps you, and any problem points as well.

Thank you for answering these questions. (1 – absolutely not true, 6- absolutely true)

1. To what extent do you think you need more grammar feedback from the teacher?

1 2 3 4 5 6

2. To what extent do you think you need more vocabulary feedback from the teacher?

1 2 3 4 5 6

3. To what extent do you think you need more organisation feedback from the teacher?

1 2 3 4 5 6

4. To what extent do you think you need more style feedback from the teacher?

1 2 3 4 5 6

5. To what extent do you think you need more content feedback from the teacher?

1 2 3 4 5 6

This is the end of the questionnaire. Thank you for filling it out.