TEACHER STANCE AS REFLECTED IN FEEDBACK ON STUDENT WRITING: AN EMPIRICAL STUDY OF SECONDARY SCHOOL TEACHERS IN FIVE COUNTRIES*

Clare Furneaux, Amos Paran & Beverly Fairfax

Abstract

This study examines the feedback practices of 110 EFL teachers from five different countries (Cyprus, France, Korea, Spain, and Thailand), working in secondary school contexts. All provided feedback on the same student essay. The coding scheme developed to analyse the feedback operates on two axes: the stance the teachers assumed when providing feedback, and the focus of their feedback. Most teachers reacted as language teachers, rather than as readers of communication. The teachers overwhelmingly focused on grammar in their feedback and assumed what we called a Provider role, providing the correct forms for the student. A second role, Initiator, was also present, in which teachers indicate errors or issues to the learner but expect the learner to pick this up and work on it. This role was associated with a more even spread of feedback focus, where teachers also provided feedback on other areas, such as lexis, style and discourse.

1. Introduction

Most teachers and researchers would agree with Hyland (2003: 207) that "(f)eedback is central to learning to write in a second language." It is therefore not surprising that as part of the selection process for a teaching and teacher training post in a Japanese Senior High School, Nishimura (1997) asked candidates to provide feedback on a student essay entitled "Information of cancer" (sic), which dealt with the question of whether cancer patients should be told of the nature of their illness. The candidates were asked to respond to the essay and annotate it. Below is an example of what one candidate did:

INSERT FIGURE 1 HERE

The annotation 'GOOD!' in this example very clearly raises issues of what it is that teachers are responding to and commenting on when they read student essays. As this example shows, teachers may ignore the content of what they are reading and focus exclusively on linguistic features.

Research in this area in L2 dates from Zamel's (1985) pioneering study. Zamel noted that most L1 research had "revealed that teachers respond to most writing as if it were a final draft" (1985: 79) and that "teachers' marks and comments usually take the form of abstract and vague prescriptions and directives that students find hard to interpret." Looking at L2 teacher feedback, she found that the teachers in her study were responding to their students as language learners rather than apprentice writers and saw their texts "as final product to be edited" (Zamel 1985: 91).

Feedback on student writing, as Keh (1990: 294) put it, "is a fundamental element of a process approach to writing," which has a major formative role in

helping learners to revise and thus helping them to reach a better end product. It has, therefore, become a much researched topic since process writing approaches to the teaching of writing impacted on foreign language instruction, beginning in the early 1980s (see Ferris 2003 for a recent summary of research approaches and findings in the area of teacher response to L2 writing).

However, as Goldstein (2001) comments, there are still various tensions within this research base, with many areas remaining unexplored. The impetus for the study described here was the sharp dichotomy between the practice of providing feedback as revealed in the teacher's comments on the essay, and what we knew about research into feedback on L2 writing, and recommended feedback practice (eg Goldstein, 2004). We, therefore, set out to explore the issues of teacher response and teacher comments on student writing with EFL teachers from a variety of contexts.

2. Previous research on L2 teacher feedback

Two major strands stand out in the L2 writing research literature since Zamel's 1985 study. One strand looks at learners' reactions to feedback, either using questionnaires and interviews (Cohen and Cavalvanti 1990; Enginarlar 1993; Ferris 1995; Hedgcock and Lefkowitz 1994; Radecki and Swales 1988) and/or analysing subsequent student revision (Conrad and Goldstein 1999; Ferris 1997; Paulus 1999). In these studies, feedback emerges as an important factor in student revision. Conrad and Goldstein (1999) found that the type of problem that was highlighted by the teachers was crucial to the success of the revision; requests for revisions that focused on detail, coherence and cohesion, and examples, were likely to result in successful revisions, for example; in contrast "if the problem to be revised focused on explanation, explicitness, or analysis, the resulting revisions were almost never

successful" (Conrad and Goldstein 1999: 160). Some of these studies cast doubt on the successful implementation of certain aspects of the process approach in ESL contexts. For example, Nelson and Carson (1998) report that sometimes learners prefer negative comments, because they highlight problems in their writing that can be worked on. These negative comments were welcomed on any aspect of the essay, and at any level of the writing – word, sentence or rhetorical pattern. For similar reasons, the learners in the Nelson and Carson study (1998) preferred teacher comments to peer feedback.

The second strand of research focuses on teachers' feedback behaviour and its links with the learners' beliefs or actual writing practices. Thus Hedgcock and Lefkowitz (1994) found that there was a "close pairing between teachers' response behaviors and learners' beliefs about their effectiveness" (1994: 157). Some studies (Cohen 1987; Raimes 1985, 1987; Zamel 1983) have found that focus on form in teacher feedback has contributed to inadequate student writing strategies and revision strategies. Paulus (1999: 266) also argues that "(t)he way that teachers structure the writing classroom and the type of feedback they give will no doubt determine how their students approach the writing process, view feedback, and make revisions to their writing." This relationship works both ways: Truscott (1996) suggests that one of the reasons provided for grammar correction is that students believe in its efficiency (though he discounts this as a valid reason for grammar correction). Truscott's 1996 paper, in which he argued against grammar correction in writing, has been the starting point for a series of papers examining the evidence and attempting to decide this point (e.g. Ferris 1999; Truscott 1999). Ferris (2004) reaches the conclusion that the jury is still out on the efficacy of grammar correction, calling for more research on this topic. Bitchener, Young and Cameron (2005) is one recent

study which looked at the longitudinal (12 weeks) effect of different types of feedback corrections on student writing.

A third possible strand, that of the nature of teachers' annotations and comments on ESL writing, is relatively unexplored. Ferris (1997) examines marginal and end comments on student drafts and subsequent student revision. The analytic model developed coded for length of comment, comment type (functionally defined: eg ask for information/question), use of hedges and whether a comment was text-specific or generic. Conrad and Goldstein (1999) code one teacher's written comments for intended function, formal characteristics and type of problem to be revised (student revision being the focus of their study). Hyland and Hyland (2001) analyse teachers' comments at the end of student texts according to the function of these comments: praise, criticism, and suggestion. They also look at hedging and other forms of mitigation. Notably, however, all these studies focus on marginal and/or end comments, where teachers are mostly acting as readers of communication, with less of a focus on language errors (which is typically located in in-text feedback). Ferris (1997) includes a category focusing on Grammar and Mechanics, and specifically points out the need for a category that would deal with form rather than content. However, only verbal comments were analysed, although the teacher in the study did use correction symbols (see Ferris et al 1997). A fourth study, Lee (2004), examines Hong Kong secondary school teachers' error correction. 58 teachers' feedback on the linguistic errors in one student essay was examined, revealing that these teachers used either direct feedback, with error correction given, or indirect coded feedback.

It is important to note that, this latest study aside, most research to date has focussed mainly on process writing contexts such as those found in US college

composition courses and British EAP programmes. Thus studies of second language teacher feedback normally have the following characteristics:

1. they are mostly in ESL contexts (e.g. Chandler 2003; Conrad & Goldstein 1999; Ferris 1995, 1997; Ferris et al 1997; Hyland & Hyland 2001);
2. where studies are in EFL contexts, they focus on one (e.g. Ashwell 2000; Enginarlar 1993; Fazio 2001; Lee 2004) or two (e.g. Cohen & Cavalcanti 1990) contexts;
3. the vast majority of studies are in tertiary level institutions (eg Hyland & Hyland 2001) and also often in college composition courses (e.g. Ferris 1995, 1997; Ferris & Roberts 2001; Hedgecock & Lefkowitz 1994; Paulus 1999)
4. the number of teachers whose feedback practices are examined is small, ranging from one teacher (Ferris et al 1997 and Paulus 1999) to Zamel's fifteen (1985). The only exception is Lee 2004;
5. the majority of teachers are native speakers of English (NS), though Cohen & Cavalcanti 1990; Kobayashi 1992; Sheory 1986 and Takashima 1987 are notable exceptions. Sometimes, in fact, the researcher does not specify whether the teachers are native speakers or non-native speakers (NNS) (e.g. Lee 2004).

The characteristics of the research as described above indicate gaps in a number of areas in our knowledge of the teaching of writing. The first area we know little about is what happens in the context where the majority of the world's learners of English are to be found: secondary schools, with NNS teachers, in an EFL rather than ESL situation (see Lee 2004 for a similar viewpoint).

In addition, our experience working with hundreds of teachers from many countries indicates that most teachers at secondary school level do not look at drafts of work: work is normally handed in for assessment purposes; the feedback is provided on the final draft only; in fact, the final draft may be the only draft that is ever submitted. This indicates a second area where there is a gap: what happens where the student is not expected to re-work the same piece and resubmit it. Finally, we know little about what teachers actually do when providing feedback to their students.

These gaps in the literature, combined with our backgrounds and the teachers we worked with, led us to an interest in looking at what NNS teachers, working in a variety of EFL contexts worldwide within secondary education frameworks, do when they provide feedback on student writing. Although the NS/NNS distinction may be regarded as an artificial one, and the concepts themselves are controversial, there is some evidence that NS and NNS teacher behaviour may be different (e.g. Árva and Medgyes 2000; Hughes and Lascaratou 1981; Shi 2001). There are clearly other issues that come into play, such as educational background (teacher training in general and training in the teaching of writing in particular) and language proficiency. At this stage of the research, however, we believe that it is justified to look at this issue using this distinction. An additional distinction we decided to make was to examine to what extent NNS teachers from different national educational contexts differ in their feedback practices.

Our research questions were as follows:

1. Do different NNS teachers take on different roles when providing feedback on student essays?

2. Which aspects of student writing do teachers focus on? Do different teachers focus on different aspects when providing feedback?
3. Is country of origin a factor in teachers' role and correction focus?

3. Methodology

3-1 Text and Participants

Previous research in this area has focused on small numbers of teachers and students, taking an ethnographic approach in which highly contextualised data, collected using a variety of methods, is triangulated. What results then is a context-rich, emic understanding of the response and feedback practices of the teachers in the study and of students' reactions to these. Since we had access to a large number of teachers from different countries, we were interested in exploring the feedback practices of these different groups. We developed a research instrument which involved all participants in the study commenting on the same student essay (cf. Lee 2004). This is admittedly an artificial exercise, as the participants did not know the student who had produced the work, and commenting on the essay was not part of an ongoing course of instruction. However, what teachers do and how they comment on a student's piece of work will depend on a large number of factors, not least among which are the topic, the level of the language, the organisation, etc. This makes comparing practices between large numbers of teachers difficult. By asking all participants to comment on the same essay, we were able to approach the issue from an angle different from that normally taken in writing research.

The source for the text was Nishimura (1997), who presents a student essay telling the story of an aunt who had died of cancer. All essays lend themselves to commenting on both form and content, but we thought that the poignancy of the essay

would encourage both, in groups of teachers who, our experience suggested, would normally focus on form only.

Data were collected from 110 teachers who came from five countries (see Table 1). Like Lee (2004), we adopted convenience sampling: the French, Korean, Spanish and Thai groups were in England on short refresher teacher education courses commissioned by their Ministries of Education when the data were collected. These courses were run at the Centre for Applied Language Studies, the University of X, where all three authors were working at the time, with the aims of up-dating their general pedagogic practice as well as their language skills. The Cypriot group were attending a language teachers' Summer School in Nicosia focusing on the teaching of writing, co-run by the first author.

Participants ranged in age from 25 to over 56, with half in the age bracket 36-45. All were experienced teachers and all had experience teaching in secondary schools, with some also having taught in other contexts. A breakdown of their experience is given in Table 2.

INSERT TABLE 1 HERE

INSERT TABLE 2 HERE

3.2 Procedure

Participants received a pack containing a teacher background questionnaire and two tasks. An accompanying letter explained that the study explored teacher response to student writing, and guaranteed anonymity. In this paper we deal with the first task, which consisted of the student essay, unmarked. Participants were told that this was the final draft of an essay written by a secondary school student in Japan, and

were asked to mark as if it had been written by one of their students and they were going to return it the following day. They were asked to write on the paper any comments, symbols or grades that they would normally put on a script to be returned to a student. They were not provided with any marking scheme or guidelines. Of course, there is always the question of whether teachers in this situation do what they would do in their own classroom, with learners they know. However, within this study this was the nearest we could come to accessing our participants' normal behaviour.

3-3. The coding scheme

A variety of coding schemes has been used by researchers looking at teacher feedback. These were not applicable to our data, partly because they had been devised with the work of a small number of teachers in mind, and sometimes with a specific research agenda. We considered the influential Faigley and Witte (1981) coding scheme (used in Paulus 1999, for example), but since it focuses on student revision it was not suitable for our purposes. Ferris (1997; Ferris et al 1997) was the first to systematically analyse teachers' written commentary, using categories that focused on the aim or intent of the comment and its linguistic features (Ferris 1997). This works well where the teacher's response is expressed in fully articulated form, most often in complete sentences. In our case, a great deal of the feedback was through annotation and symbols (e.g. ticks, underlining, etc.), and verbal feedback, though sometimes present, was often quite short. Hyland and Hyland (2001), too, look at verbal comments, identifying praise, criticism and suggestions. This is more similar to the coding system that we developed, as we shall see below, but we felt we needed fuller differentiation between types of feedback that existed in our data. Lee (2004) coded for language error types, and effectiveness of correction. This scheme differentiates

only between direct and indirect feedback on language error alone; the coding scheme developed for the current study is more comprehensive and differentiates between a large number of feedback types.

Since none of the available systems of coding feedback seemed to fit our data, we developed our own coding scheme, through a grounded theory approach in which "categories are discovered by examination of the data" (Glaser & Strauss 1967: 3). This entails interrogating the data and coming up with successively more refined coding schemes, until one is found which was theorised to be valid and shown to be statistically reliable. (See Ferris et al 1997, for a similar approach).

A first trawl through our data indicated that teachers were assuming different roles in their comments and we therefore wanted our coding to enable us to look at teacher role in providing feedback, and specific feedback focus. The development of the coding scheme consisted of a recursive process of positing coding categories, refining them, and then positing new or amended ones, until we were able to reliably code the majority of teacher annotations and comments.

The final coding scheme operates on two axes. The first axis is that of teacher role. Six teacher roles were identified: Initiator, Supporter, Advisor, Suggester, Provider, and Mutator. The second axis is the feedback focus of each marking. The correction foci which we identified were: lexical, grammatical, stylistic, semantic, discoursal, mechanics. In addition, an unspecified category and an unclassifiable category were used. These categories are explained below.

3-3-1. Teacher Roles

The six roles are listed below, with the operational definitions used in the coding and the reliability calculation , and examples are provided for each.

Initiator: *Alerts by providing a specified (lexical, grammatical, stylistic, semantic, discoursal, mechanical) or unclassifiable (dotted lines, circle, question mark etc.) alert. The alert may take the form of a question or an explanation, provided there is no actual correction.*

INSERT FIGURE 2 HERE

Figure 2 shows a passage from the text, exemplifying an Initiator role. This category was so named because the teacher is providing a stimulus initiating a thought process in the learner, who needs to pick this up and take it further: the correction is usually fairly unspecific. In many cases, the teacher would either circle a word or underline it, indicating the type of problem with a code (e.g. T for tense; WW for Wrong Word; V for Vocabulary). In other cases the teacher would write a brief explanation of the problem near the underlined word (e.g. "no contraction" near "I'd").

Supporter: *Responds positively to the text with either symbols (++) or comments.*

INSERT FIGURE 3 HERE

This often took the form of the symbols + or ✔ or the word "good." In some cases the teacher would be explicit about the area of language being commented on, as in Example a in Figure 3. Example b shows a case where, although the teacher does not specify verbally what area is being praised, the location of the symbols makes it

clear that the comment refers to grammar. In other cases the aspect of the student's work that was being commented on would not be specified (Example c in Figure 3).

Advisor: *Identifies areas where the student needs to do further work, either on this particular text or in general. The teacher may offer to help the student work on a problem area. The note is clearly intended as advice*.

INSERT FIGURE 4 HERE

Here the teacher might suggest to the student that they should check the grammar of the sentence, or make a comment. Comments were sometime general - e.g. "some work on pronouns here" or "use an introductory sentence". In other cases the advice might be more direct, asking the student to rewrite a sentence and including the word with which they should start the sentence.

Suggester: *The teacher indicates advice by suggesting a better alternative in brackets (where elsewhere, for example, items are crossed out), or writes alternative(s) above the uncorrected original. (This contrasts mainly with Provider behaviour. In the case of Suggester the teacher does not indicate that what the student has written is actually wrong. This category was used only when the teacher clearly contrasted between Provider behaviour and Suggester behaviour and used a different system of correction for each.)*

INSERT FIGURE 5 HERE

The examples in Figure 5 may be less transparent than the examples in Figures 2-4, since identifying teacher feedback as belonging to the Suggester category is dependent on the existence of clear Provider behaviour (see below). Although ultimately this turned out to be a small category, it does capture a difference in feedback behaviour which justifies separating the two categories.

Provider*: Provides the correct form by substitution, addition, deletion or re-ordering of an item of language or punctuation. Such corrections do not change the meaning. May be accompanied by an explanation or identification of the problem.*

INSERT FIGURE 6 HERE

In this role the teacher provided the correct form for the student. Many corrections in this role left the learner no redress: the student was told exactly what to do and how to express various points, sometimes in ways which come across as quite strong, with the student's own words being crossed out and substitutes provided.

Mutator: *The teacher alters the text by deleting, adding or rewriting. Such alterations change the meaning.*

INSERT FIGURE 7 HERE

Figure 7 shows an example of a number of mutating changes within one passage in the text: the addition of "if", the substitution of "so" by "and", and the deletion of

"since I might not be cool" were all taken to entail a change in the meaning of the text.

Overall, then, the role axis describes the feedback behaviour of teachers on a continuum from open to closed and from general to specific. It moves from general hints about what needs to be done at the Initiator end of the continuum, through more specific prompts and suggestions, to outright corrections and rewriting of student text, ending with meaning-changing corrections at the Mutator end of the spectrum.

3-3-2. Correction Feedback

In addition to teacher roles, each piece of feedback was coded according to the point which the teacher was focusing on. The categories identified are listed below.

Lexical focus: cases where the teacher indicated that a wrong word was used. Identification of this focus would be through the code the teacher used (e.g. WW, V) or through the teacher's pointing out or providing alternative lexical items. (See example a in Figure 4 and example b in Figure 5).

Grammatical focus: cases where the teacher indicated a problem with grammar, including word order. Again, in some cases the teacher would use a code (T, WW where the focus is grammatical, such as use of the article) or provide the actual correction. (See examples a and b in Figure 3 and Example b in Figure 6).

Style focus: cases where there was no actual mistake but the teacher provided an alternative which was equivalent to the student's choice though different in terms

of formality or appropriacy. Initiator comments on style might be of the type "no contractions", or "would like" written next to "want." Some Advisor comments were judged as focusing on style – e.g. "Rewrite this sentence. Too many words." Some teachers included a comment on formality or appropriacy, as for example the comment in Figure 2.

Semantic focus: cases where the teacher's feedback focused on the meaning of what was being said and was conveying the message that "this is not the way to express this idea." The teacher's focus would not be on a lexical level, but on a more holistic level of expressing meaning. (See example b in Figure 4 and example c in Figure 6).

Discourse focus: cases where the teacher's comments focused on cohesion, organisation or paragraphing.

Mechanics: spelling, punctuation, capitalisation, or comments on handwriting.

In addition, an "unspecified" category was used where teachers used underlining, circles, or question marks (that is, identified a problem and the location of the problem), but did not specify what type of problem this was. Finally, an "unclassifiable" category was used for those cases where the teacher provided feedback but it was not possible to classify the response in any of the above categories.

3-4. Coding and Analysis Procedure

Once the coding scheme was finalised, the entire corpus of teacher corrections was coded. Problematic instances were discussed and agreement reached. An inter-rater reliability check resulted in a score of 90.13%, confirming the reliability of the coding.

3-5. Role attribution

The final stage in the analysis was classifying each teacher according to the main role categories which the majority of their corrections fell into. A teacher was classified as taking on a particular role if at least 50% of their corrections fell within that role, with a difference of at least 20% between that and the next most frequent role (e.g. a teacher whose profile was: 56% Initiator; 33% Provider; 8% Suggester and 3% Mutator was categorised as an Initiator). A teacher who did not fall into a clear category would be categorised according to the two main categories they belonged to (e.g. a teacher whose profile was 53% Provider, and 47% Initiator was classified as Provider/Initiator).

4. Results

A total of 4637 feedback annotations were identified in the entire sample. The distribution of these annotations was analysed using ANOVA. The variables analysed in the ANOVA were the main effects of role, focus and nationality, and the two-way interactions between them. It was not possible to analyse the three-way interaction of nationality, role, and focus, because there were too many possible combinations for which there were no instantiations. Because mutating annotations could not be coded

for focus, this role was not included in the statistical analysis; mutating annotations are discussed separately below. Since there were 270 mutating corrections, the total number of corrections analysed for the ANOVA was 4367. Table 3 presents the results of the ANOVA.

INSERT TABLE 3 HERE

The ANOVA indicates that the three main effects - of role, of correction focus and of nationality - are highly significant. So are two of the two-way interactions: role-focus, and role-nationality. The focus-nationality interaction is not significant. The main effects and the interactions are discussed below.

4-1. Role

Table 4 shows the number of corrections within each role category and the percentage of corrections that fell within each role.

INSERT TABLE 4 HERE

The ANOVA revealed the main effect of role was statistically highly significant ($F=78.9$, $p<0.001$), indicating that some of the roles occurred significantly more than others; the size of the MS indicated that the main effect of role was substantial enough to be considered on its own.[1] A Tukey Pairwise comparison was carried out to

[1] The MS is 'a measure of average variability' (Field 2005:738). In an analysis such as the one here it is important to compare the size of the MS for the main variables and the MS for the interactions between them. If the MS for any main effect is considerably larger than the MS for the interaction, then the main effect can be considered on its own, rather than as mediated through the interaction (see Mead 1988 for a full explanation of this point).

investigate where within the main effect of role this significance lay. Supporter, Advisor and Suggester showed no significant differences between them, and all three were significantly different from both Initiator and Provider. More importantly, there was a statistically significant difference between the findings for Initiator role and Provider role ($p<0.0001$).

4-2. Focus

Table 5 shows the mean number of feedback annotations by focus. The main effect of focus was also significant ($F=35.53$, $p<0.001$), indicating that there was a statistically significant difference between the number of corrections falling within the various foci. A Tukey Pairwise comparison was carried out to investigate where this significance lay. Here grammar proved to be the main source of the differences between the different foci. All comparisons of grammar with the other foci were highly significant ($p<0.0001$). Of the other pairwise comparisons, only two were significant: style and lexis ($p<0.05$) and style and unspecified ($p<0.01$).

INSERT TABLE 5 HERE

Unlike the main effect of role, however, the MS of the main effect of focus is of the same magnitude as the MS of the role and focus interaction, and the main effect of focus cannot therefore be interpreted on its own, but should be investigated in the context of the interaction. Tables 6 and 7 show the interaction between focus and role in raw numbers and in percentages, respectively, showing that providing comments and initiating comments are not spread equally but are concentrated on different foci. Providing comments are concentrated quite clearly on Grammar (61.6% of providing

comments fall into this category) with the next focus category far behind it. Initiating comments, on the other hand, are more evenly spread out; the highest number of initiating comments have an unspecified focus, but they take up only 33.8% of the comments in this role; the second highest focus within this role is Grammar with 26.5%.

INSERT TABLE 6 HERE

INSERT TABLE 7 HERE

Another phenomenon apparent from Table 7 is the way in which unspecified comments are clustered within Initiator and Supporter roles. It seems that the nature of this type of comment is such that it is possible to make them without specifying what the focus is – something which would be much more difficult where Provider comments are concerned.

4-3. Nationality

Table 8 presents the total number of corrections by nationality, as well as the mean number of corrections per teacher for each nationality. The main effect of nationality was highly significant ($F=6.93$, $p<0.001$). A Tukey Pairwise comparison was carried out to establish where the significance lay. The only comparisons that were significant were those of the Korean teachers with the Cypriot teachers ($p<0.02$), the French teachers ($p<0.0001$), and the Spanish teachers ($p<0.002$). There was no significant difference between the Koreans and the Thais.

INSERT TABLE 8 HERE

However, the MS for the main effect of nationality is not greatly different from the MS for the interaction of role and nationality, and the main effect of nationality therefore needs to be interpreted within that light. Table 9 and Table 10 present the distribution of annotations by role and nationality. Table 9 presents the totals, and Table 10 presents the percentages.

INSERT TABLE 9 HERE

INSERT TABLE 10 HERE

When we consider the significance of the main effect of nationality, the Tukey Pairwise comparisons between nationalities, and the interaction between nationality and role, the following picture emerges. It seems that there are no significant differences between teachers of different nationalities in the number of annotations they make, apart from the Korean group, who made significantly fewer corrections than three of the four other nationalities. However, different nationalities assume different roles when correcting, as Tables 9 and 10 indicate. The picture is clearest when looking at the Koreans - they very clearly exhibit providing behaviour, with 94% of their comments falling into this category. At the opposite end of the scale are the French teachers, who presented more initiating annotations than providing ones, with both categories being quite close to each other in number: 50% and 41% respectively. The two categories are also quite close to each other for the Cypriots. With the Thai and Spanish teachers, however, the two categories are further apart:

with the Thai teachers there is a ratio of providing to initiating annotations of 5:1, and for the Spanish teachers the ratio is 3:1.

4-4. Analysis of Teachers' Personal Role

Tables 11 and 12 show the results of role attribution; Table 11 presents the raw numbers, and Table 12 shows the percentages per nationality. Clearly, there are two roles which dominate - Provider and Initiator; there are two roles which appear only as second roles in combination with Initiator - Suggester and Supporter; and there are two roles which do not appear in the analysis as prevalent roles at all - Mutator and Advisor. Most teachers (a total of 95, i.e. 86% of the sample) fall clearly within either the Initiator profile or the Provider profile. Since a teacher was classified as a Provider or Initiator only if at least 50% of his or her corrections fell within the respective category, and only if the next category was at least 20% below the main category, this indicates quite a strong lack of variability in individual teacher response: of the whole sample, only 13 teachers (i.e. 12%) had a substantial percentage of responses in more than one category. Interestingly, one of these categories was always that of Initiator.

For two nationalities, French and Cypriots, there are more Initiators than Providers; for the other three nationalities there are more Providers than Initiators. The data for these two categories were analysed using a Likelihood Ratio Chi-square test (this test is more robust to small numbers than the Pearson Chi-square test and can deal with expected counts that are lower than 5). The differences were shown to be significant ($\chi_2 = 33.7$ df=4, $p<0.001$). Consideration of the expected values in the contingency table showed that the French teachers were expected to have fewer Initiators and more Providers than the actual numbers. For the Cypriots this was also

true, but to a far lesser extent. Among the other nationalities it was the Koreans who stood out as having more Providers than expected and fewer Initiators than expected (in fact, there were none).

INSERT TABLE 11 HERE

INSERT TABLE 12 HERE

4-5. Mutators

As mentioned above, it was not possible to analyse the role of Mutators together with the other roles. However, in terms of the overall performance it was the third most frequent category (see Table 4 above), after Provider and Initiator, and ahead of Suggester, Supporter and Advisor. It thus merits discussion, in spite of the lack of statistical tests for these data.

Table 13 presents the distribution of mutating corrections by nationality. The results indicate clearly that the Thai teachers included a far greater percentage of mutating corrections than the other nationalities in the study. A closer look at the data indicated that the majority of teachers, 65%, included mutating comments in their annotations. Although there were no teachers whose main role - or indeed secondary role - was that of Mutator, there were nevertheless some teachers whose comments included a fair amount of mutating corrections: 6 French teachers, 2 Koreans and 2 Spanish teachers had more than 10% of their comments as mutations. The Thais showed the greatest incidence of this, with 7 teachers who had more than 10% mutating corrections, 4 of whom had 33%, 30%, 28%, and 21% respectively.

INSERT TABLE 13 HERE

4-6. Margin and End Comments

We chose the essay for our study because of its content, but very few teachers responded to this content either through making margin comments or through end comments (see below for a discussion of the numbers). Of those teachers who did make margin and end comments, very few related to the story which the student tells. Only one teacher in the whole sample wrote "I'm very sorry about your aunt's death"; another wrote "A very moving story" but went on to add "and a very appropriate illustration", thus turning this into a comment on the writing rather than on the experience. Others who commented on the personal experience did so only within the general context of providing a personal example in an essay, and often in a non-committed way: "You tried to write many things based on a real experience"; "I enjoyed reading about your conversation with your parents and your personal reactions"; "I see you care about other people's feelings".

Many of the other comments are in fact baffling. One such comment was "You could have avoided repeating the same idea, thinking through, such as I would like to be told if I have cancer because I would fight against it. This idea was repeated two or three times". Another comment was "You seem to mention the word cancer in every other line" – quite a shocking comment, taking into account the topic of the essay.

From a quantitative point of view, the first question of interest was to see how many of the teachers actually provided margin or end comments. Table 14 presents the numbers of teachers who provided comments and those who did not provide comments by nationality. This was analysed using a Likelihood Ratio Chi-square test. This was significant (χ_2 = 15.9 df=4, p=0.003). An examination of the observed and

expected counts revealed that the only group where there was a major difference was the Thai teachers, most of whom made no comments.

INSERT TABLE 14 HERE

We also looked at the margin or end comments according to the role which the teacher assumed. Table 15 presents the numbers of teachers who provided margin or end comments according to their main role, taking into account only the Initiators and the Providers, i.e. the two groups where role attribution was clear cut (see Table 11 above). An examination of these data indicates that although in both groups the teachers tended to provide additional comments, there was a relatively higher proportion of Initiators who did so. Although when these data were analysed using a Likelihood Ratio Chi-square test, the difference was not found to be significant (χ_2 = 3.0, df=1, p=0.082), the findings suggest that with a larger number of respondents this might have achieved significance.

INSERT TABLE 15 HERE

5. Discussion

5-1. Limitations of this study

First, a word of caution is in order regarding the limitations of this study. One limitation is that the teachers were all self-selected, in that they had chosen to attend in-service courses. Other limitations may be more of a concern. All our teachers were NNS, but we were unable to control for proficiency, which may have been an influencing factor. This is especially the case with the Korean teachers, who were felt

by their instructors to have been of lower proficiency than the other teachers in this study. Since the nationality main effect seems to be mainly due to the Koreans, this may in fact be due to proficiency.

Another factor is that the teachers were examining learner language which they might not have come across in real life. Any errors may therefore have intruded more, and skewed the results. This could have been solved by asking each group to correct an essay written by a student in their own country. However, as pointed out earlier, that would have obliterated any possibility of comparing across cultural groups.

Finally, we might ask whether, in fact, a simulated exercise such as this will access teachers' real behaviour, since 'teachers respond to students in their comments, and not just to texts' (Hyland and Hyland 2006: 4). Some teachers might not make a personal comment on a topic such as this in writing to the student, but would talk to the student in person, in a break. Another possibility is that teachers will annotate less (or more!) than they would in real life, or that they will annotate according to their perception of what it is that the researchers are interested in. One could also argue that possibly teachers cannot picture the student to whom they are giving the feedback; this is not a known individual. They may feel that that they are marking in a context that is unknown to them. But this is precisely the type of individual variability which (though interesting) will have less of an effect when researchers look at a large sample such as the one we examined. In addition, such objections do not in fact detract from the validity of the present study, since they would apply to all the teachers in our study, irrespective of their nationality.

5-2. Summary and conclusions

The first research question we posed was whether different teachers take on different roles when providing feedback on student essays. We found that of the six roles we identified, the Provider role was more prevalent. This is supported both by the number of providing annotations, which was significantly higher than the number of any other type of annotation, and also by the role attribution analysis, where more than half our sample were quite unambiguously Providers. However, initiating behaviour was also present in our sample, with a high percentage of teacher annotations falling within this role; this was significantly different from the other roles. In the role attribution analysis, too, this was the second most frequent role. This is in line with the findings in Lee (2004), where secondary school teachers provided explicit correction for more than half of the errors (a direct parallel to our Provider category), and where the only other type of feedback identified was indirect coded feedback, corresponding to our Initiator category. In our study, possibly because of the nature of the task, we were able to identify additional categories of feedback. In addition, we found different patterns in different nationalities.

Our second question asked whether different teachers focus on different aspects of student writing when providing feedback. The answer here was quite unambiguous: our teachers very clearly respond to language, unlike those in ESL studies of teacher feedback we reported on above, where there is a greater tendency for teachers to respond as readers of communication. Within the strongly language-oriented feedback behaviour exhibited by the teachers in our study, grammar came very much to the fore. Grammar-oriented annotations accounted for 45% of the annotations analysed in Table 5 above. The vast majority of the annotations were done through a Provider role. Indeed, this type of annotation – grammar provision – accounted for 33.5% of all annotations, though this was only one out of 35 possible combinations. The teachers

in the study thus exhibited what we feel is quite a traditional feedback pattern. Another important finding relates to the focus of the annotation in the case of different roles. We found that where initiating behaviour is concerned the spread of the teachers' focus is much more even than where Provider roles are concerned.

Our third research question asked whether patterns of response vary from country to country. The picture here was slightly more complex. Overall, few differences were found, and the only differences that were significant were due to one nationality group – the Koreans, who corrected far less than the other groups. Looking at the interaction of nationality and role, it is also clear that the French teachers were different from the other groups, in that they presented more Initiator than Provider annotations. However, we feel that the feedback practices of our groups of teachers have more in common with each other than with the feedback given by NS teachers in ESL contexts described in earlier studies. We feel this is true even taking into account the differences in methodology between our study and previous ones.

Our study also reaffirms Zamel's (1985) findings that, indeed, teachers sometimes mutate student text. Although we were unable to analyse mutating annotations together with the other annotations, it is still worth noting that mutating corrections were the third largest category of annotations, and that here, too, there seemed to be an effect of nationality, in that the Thai group included far more mutating annotations than the other nationalities in this study.

5-3. Further research

As far as we are aware, this is the only study, apart from Lee (2004), which has looked at this area, and the only one that has attempted international comparisons. This is, therefore, an area where replications in other contexts would be particularly

important. One obvious area for replication is the use of our coding scheme. This was devised with this text, and these teacher corrections, in mind. It would be interesting to know whether the coding scheme works in other contexts. A second, profitable area of research would be to look at different populations of teachers, for example, comparing NS and NNS teacher behaviour, or the behaviour patterns of teachers of other foreign languages. One could also examine the behaviour of content teachers who are not teaching language, but who read student work in EFL or ESL contexts – e.g. content teachers in secondary schools, or subject specialists in ESP contexts (see for example, Leki 2006, which explored a similar question in university setting). In addition, it would be interesting to compare teachers involved in process writing, such as those teaching on British EAP courses or American composition courses, with those taking a more product-oriented approach: it would stand to reason that the behaviour of teachers from these two different approaches would be different. A third area of investigation would be teacher feedback on different drafts. There is no reason to assume that teachers focus on the same areas when looking at drafts at different stages. Do teachers adopt different feedback roles on different drafts? Are the foci different in different drafts? Finally, one could combine the quantitative methodology we adopted with a qualitative one, asking practitioners such as the EFL secondary school teachers we studied to discuss their feedback practices to explore the reasons they give the feedback they do.

Hyland and Hyland (2001: 185) rightly point out that "Providing written feedback to students is one of the ESL writing teacher's most important tasks, providing the kind of individual attention that is otherwise rarely possible under normal classroom conditions". We also know that students are deeply affected by the messages this feedback gives them about what matters in writing. Process writing

approaches have influenced the feedback practices of teachers in ESL tertiary-level contexts, but it would seem that this may not the case for some EFL secondary school teachers: the focus of their feedback and the stance they adopt indicate that they have other concerns, and their feedback practices are therefore quite strikingly different, as indicated by the findings of our study.

References

Arva, V. & Peter Medgyes (2000) Native and non-native teachers in the classroom. *System*, 28(3): 355-372.

Ashwell, Tim (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: is content feedback followed by form feedback the best method? *Journal of Second Language Writing,* 9 (3): 227-257.

Bitchener, John, Stuart Young and Denise Cameron (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing.* 14 (3): 191-205.

Chandler, Jean (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing,* 12 (3): 267-296.

Cohen, Andrew. D. (1987). Student processing of feedback on their compositions. In *Learner strategies in language learning,* Anita Wenden & Joan Rubin (eds.), 57-69. Englewood Cliffs, NJ: Prentice Hall.

Cohen, Andrew D. and Marilda C. Cavalcanti (1990). Feedback on compositions: teacher and student verbal reports. In *Second Language Writing: Research Insights for the Classroom*, Barbara Kroll (ed.), 155-177. New York: Cambridge University Press.

Conrad, Susan M. and Lynn M. Goldstein (1999). ESL student revisions after teacher-written comments: text, contexts and individuals. *Journal of Second Language Writing* 8 (2): 147-180.

Enginarlar, Hüsnü (1993). Student response to teacher feedback in EFL writing. *System,* 21 (2): 193-204.

Faigley, Lester and Stephen Witte (1981). Analyzing revision. *College Composition and Communication,* 32 (4): 400-414.

Fazio, Lucy L. (2001). The effect of corrections and commentaries on the journal writing accuracy of minority- and majority-language students. *Journal of Second Language Writing*, 10 (4): 235-249.

Ferris, Dana (1995). Student reactions to teacher response in multi-draft composition Classrooms. *TESOL Quarterly*, 29 (1): 33-53.

- (1997). The influence of teacher commentary on student revision. *TESOL Quarterly*, 31 (2): 315-339.

- (1999) The Case for Grammar Correction in L2 Writing Classes: A Response to Truscott (1996). *Journal of Second Language Writing*, 8 (1): 1-11.

- (2003). Responding to writing. In *Exploring the Dynamics of Second Language Writing*, Barbara Kroll (ed), 119-140. Cambridge: Cambridge University Press.

- (2004). The *"Grammar Correction" Debate in L2 Writing*: Where are we, and where do we go from here? (and what do we do in the meantime…?). *Journal of Second Language Writing* 13(1): 49-62.

Ferris, Dana, Susan Pezone, Cathy R. Tade, and Sharee Tinti (1997). Teacher commentary on student writing: descriptions & implications. *Journal of Second Language Writing*, 6 (2): 155-182.

Ferris, Dana and Barrie Roberts (2001). Error feedback in L2 writing classes: how explicit does it need to be? *Journal of Second Language Writing,* 10 (3): 161-184.

Field, Andy. 2005. *Discovering Statistics Using SPSS. Second Edition*. London: Sage Publications.

Glaser, Barnie G. and Anselm L. Strauss (1967). *The Discovery of Grounded Theory.* Chicago: Aldine de Gruyter.

Goldstein, Lynn. (2001). For Kyla: What does the research say about responding to ESL writers? In *On Second Language Writing,* Tony Silva and Paul Kei Matsuda (eds), 73-89. Mahwah, NJ: Lawrence Erlbaum.

Goldstein, Lynn M. (2004). Questions and answers about teacher written commentary and student revision: teachers and students working together. *Journal of Second Language Writing,* 13 (1): 63-80.

Hedgecock, John and Natalie Lefkowitz (1994). Feedback on feedback: assessing learner receptivity to teacher response in L2 composing. *Journal of Second Language Writing,* 3 (2): 141-163.

Hughes, Arthur and Chryssoula Lascaratou (1982). Competing criteria for error gravity. *ELT Journal*, 36 (3): 175-182.

Hyland, Ken (2003). *Second Language Writing*. Cambridge: Cambridge University Press.

Hyland, Fiona and Ken Hyland (2001). Sugaring the pill: praise and criticism in written feedback. *Journal of Second Language Writing,* 10 (3): 185-212.

Hyland, Ken and Fiona Hyland (2006). Contexts and issues in feedback on L2 writing: An introduction. In *Feedback in Second Language Writing*, Ken Hyland and Fiona Hyland (eds.), 1-19. Cambridge: Cambridge University Press.

Keh, Claudia L. (1990). Feedback in the writing process: a model and methods for implementation. *ELT Journal*, 44 (4): 294-304.

Kobayashi, Toshihiko (1992). Native and non-native reactions to ESL compositions. *TESOL Quarterly,* 26 (1): 81-112.

Lee, Icy (2004). Error correction in L2 secondary writing classrooms: The case of Hong Kong. *Journal of Second Language Writing,* 13 (4): 285-312.

Leki, Ilona (2006). “You cannot ignore”: L2 graduate students’ response to discipline-

based written feedback. In *Feedback in Second Language Writing*, Ken Hyland and Fiona Hyland (eds.), 266-285. Cambridge: Cambridge University Press.

Mead, Roger (1988). *The Design of Experiments: Statistical Principles for Practical Applications*. Cambridge: Cambridge University Press.

Nelson, Gayle L. & Carson, Joan G. (1998). ESL students' perceptions of effectiveness in peer response groups. *Journal of Second Language Writing*, 7 (2): 113-132.

Nishimura, Naoto (1997). *Innovation in English Language Teaching at Tokai Senior High School.* Unpublished MA TEFL Dissertation, University of X, UK.

Paulus, Trena M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing*, 8 (3): 265-289.

Radecki, Patricia M. and Swales, John M. (1988). ESL Student Reaction to Written Comments on their Written Work. *System*, 16 (3): 355-365.

Raimes, Ann (1985). What unskilled ESL students do as they write: a classroom study of composing. *TESOL Quarterly,* 19 (2): 229-258.

Raimes, Ann (1987). Language proficiency, writing ability and composing strategies: a study of ESL college student writers. *Language Learning*, 37 (3): 439-468.

Sheory, Ravi (1986). Error perceptions of native-speaking and non-native-speaking teachers of ESL. *ELT Journal,* 40 (4): 306-312.

Shi, Ling (2001) Native- and nonnative-speaking EFL teachers' evaluation of Chinese students' English writing. *Language Testing,* 18 (3): 303-325.

Takashima, Hideyuki (1987). To what extent are non-native speakers qualified to correct free composition? – a case study, *The British Journal of Language Teaching,* 25 (1): 43-48.

Truscott, John (1996). The case against grammar correction in L2 writing classes, *Language Learning,* 46 (2): 327-369.

Truscott, John (1999). The case for the case against grammar correction in L2 writing

classes: A response to Ferris. *Journal of Second Language Writing* 8 (2): 111-122.

Zamel, Vivian (1983) The composing processes of advanced ESL students: six case studies. *TESOL Quarterly*, 17 (2): 165-187.

Zamel, Vivian (1985). Responding to student writing*, TESOL Quarterly,* 19 (1): 79-101.

Table 1

Participants: numbers and countries of origin

Nationality	N
French	41
Cypriot	24
Spanish	17
Korean	17
Thai	11

Table 2

Background and experience of the teachers in the study

	0-1 year	2-5 years	6-15 years	16+ years
Teaching English	0	9%	46%	45%
Teaching EFL at secondary school	0	11%	48%	41%
Teaching writing	5%	13%	45%	37%

Table 3

ANOVA of corrections according to role, focus and nationality

Source	df	Seq. SS	Adj. SS	Adj. MS	F	p
Role	4	5456.12	4295.77	1073.94	78.9	0.000
Focus	7	4691.16	3385.65	483.66	35.53	0.000
Nationality	4	377.12	377.12	94.28	6.93	0.000
Role*Focus	28	12816.23	12816.23	457.72	33.63	0.000
Role*Nationality	16	1146.06	1146.06	71.63	5.26	0.000
Focus*Nationality	28	428.45	428.45	15.30	1.12	0.325
Error	112	1524.57	1524.57	13.61		
Total	199	26439.71				

Table 4

Number and percentage of corrections by role

Role	No. of corrections	% of total corrections
Initiator	1704	37
Supporter	50	1
Advisor	22	0.4
Suggester	211	4.5
Provider	2380	51
Mutator	270	6
Total	4637	100

Table 5

Number and percentage of corrections by focus

Focus	Total	%
Grammar	1973	45.2
Lexis	495	11.3
Semantics	448	10.25
Discourse	273	6.25
Style	27	0.6
Mechanics	391	9.0
Unspecified	613	14.0
Unclassifiable	147	3.4
Total	4367	100

Table 6

Total number of corrections by role and focus

	grammar	lexis	discourse	semantics	style	mechanics	unclassifiable	unspecified
Initiator	452	196	93	135	7	104	141	576
Supporter	6	5		2				37
Advisor	7	1	7	4	1		1	
Suggester	42	86	18	56	5	4		
Provider	1465	207	155	251	14	283	5	

Table 7

Percentage of corrections by focus for each role

	grammar	lexis	discourse	semantics	style	mechanics	unclassifiable	unspecified	total %
Initiator	26.5	11.5	5.4	7.9	0.4	6.1	8.3	33.8	100
Supporter	12	10		4				74	100
Advisor	31.8	4.5	31.8	18.2	4.5		4.5		100
Suggester	19.1	40.8	8.5	26.5	2.4	1.9			100
Provider	61.6	8.7	6.5	10.5	0.6	11.9	0.2		100

Table 8

Total and mean number of annotations by nationality

	Total annotations	Number of teachers	Mean and standard deviation of annotations per teacher
French	2013	41	49.1 (12.92)
Cypriot	929	24	38.7 (9.58)
Spanish	789	17	46.4 (11.91)
Korean	255	17	15.0 (10.63)
Thai	381	11	34.6 (15.73)

Table 9

Total number of annotations by role and nationality

	Initiator	Supporter	Advisor	Suggester	Provider
French	1010	28	9	121	844
Cypriot	449	5	0	1	473
Spanish	179	4	12	61	533
Korean	15	0	0	0	240
Thai	52	10	1	28	290

Table 10

Percentage of annotations by role for each nationality

	Initiator	Supporter	Advisor	Suggester	Provider	Total
French	50.2	1.4	0.5	6	41.9	100
Cypriot	48.3	0.5		0.1	50.9	100
Spanish	22.7	0.5	1.5	7.7	67.6	100
Korean	5.9				94.1	100
Thai	13.6	2.6	0.3	7.3	76.1	100

Table 11

Role attribution by nationality

	French	Cypriot	Spanish	Korean	Thai	Total
Initiator	23	12	3	0	1	39
Provider	12	9	12	15	8	56
Initiator/Provider	1	3	1			5
Provider/Initiator	5		1			6
Initiator/Suggester					1	1
Initiator/Supporter					1	1
No Attribution				2		2

Table 12

Role attribution by nationality expressed in percentages

	French	Cypriot	Spanish	Korean	Thai
Initiator	56.1	50	17.6		9.1
Provider	29.3	37.5	70.6	88.2	72.7
Initiator/Provider	2.4	12.5	5.9		
Provider/Initiator	12.2		5.9		
Initiator/Suggester					9.1
Initiator/Supporter					9.1
No Attribution				11.8	
Total	100	100	100	100	100

Table 13

Number and % of mutating annotations by nationality

	French	Cypriot	Spanish	Korean	Thai
Total mutating annotations	101	19	48	20	82
Total annotations	2013	929	255	789	381
% of all annotations	5%	2%	7.8%	6%	21.5%

Table 14

Presence of margin and end comments by nationality

	French	Cypriot	Spanish	Korean	Thai
Comments	29	20	13	10	2
No Comments	12	4	4	7	9

Table 15

Presence of margin and end comments by teacher role

	Initiators	Providers
Comments	29	32
No Comments	10	24

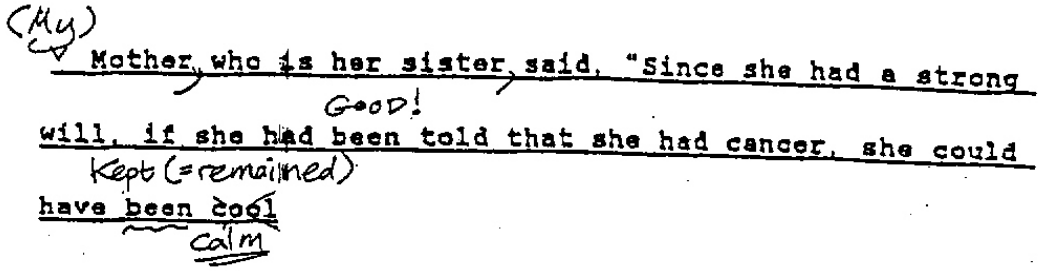


Fig. 1. Example of teacher comment on a student essay (Nishimura 1997: 220)

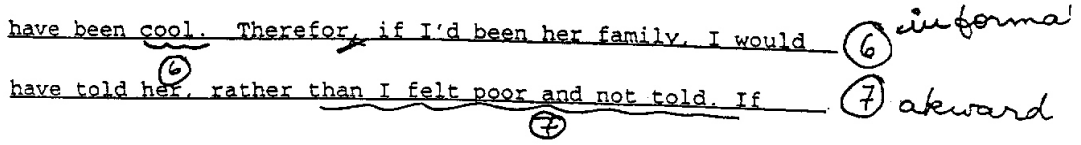


Fig. 2. Examples of Initiator annotations

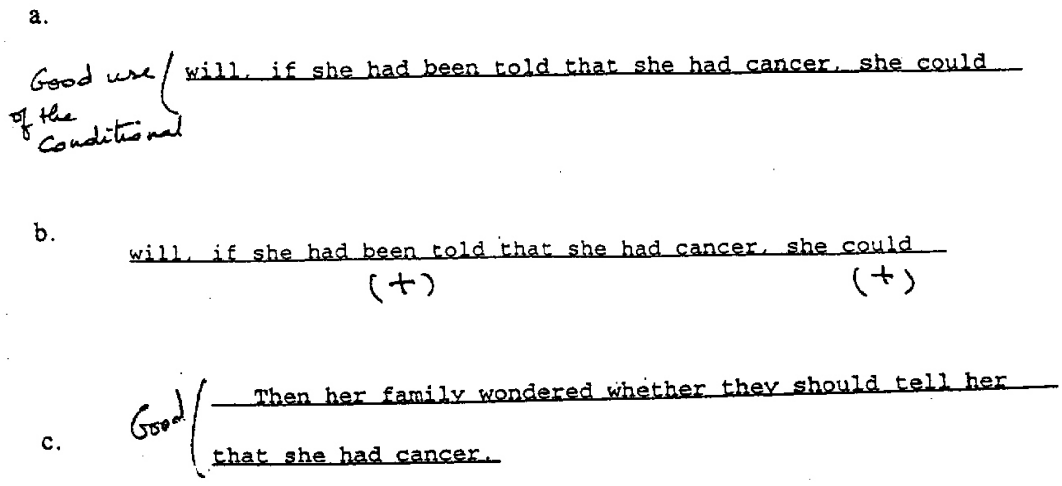


Fig. 3. Examples of Supporter annotations

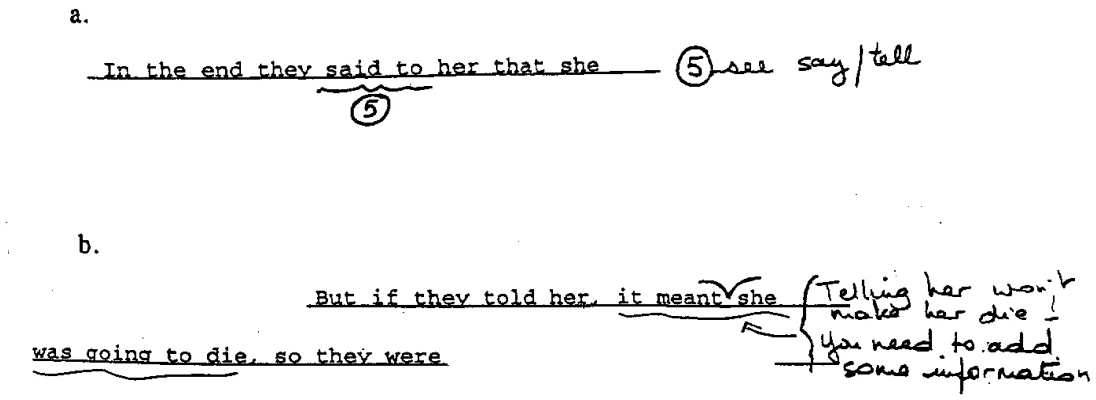
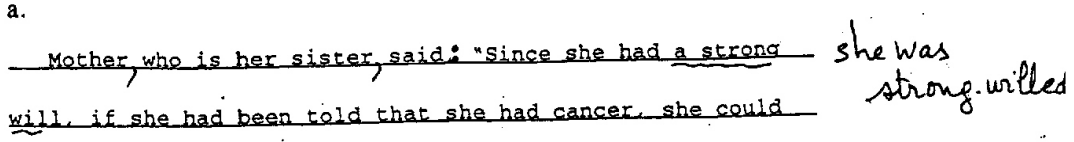


Fig. 4. Examples of Advisor annotations

a.



b.

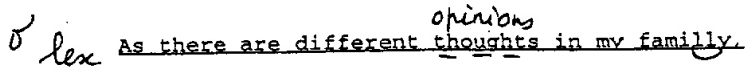


Fig. 5. Examples of Suggester annotations

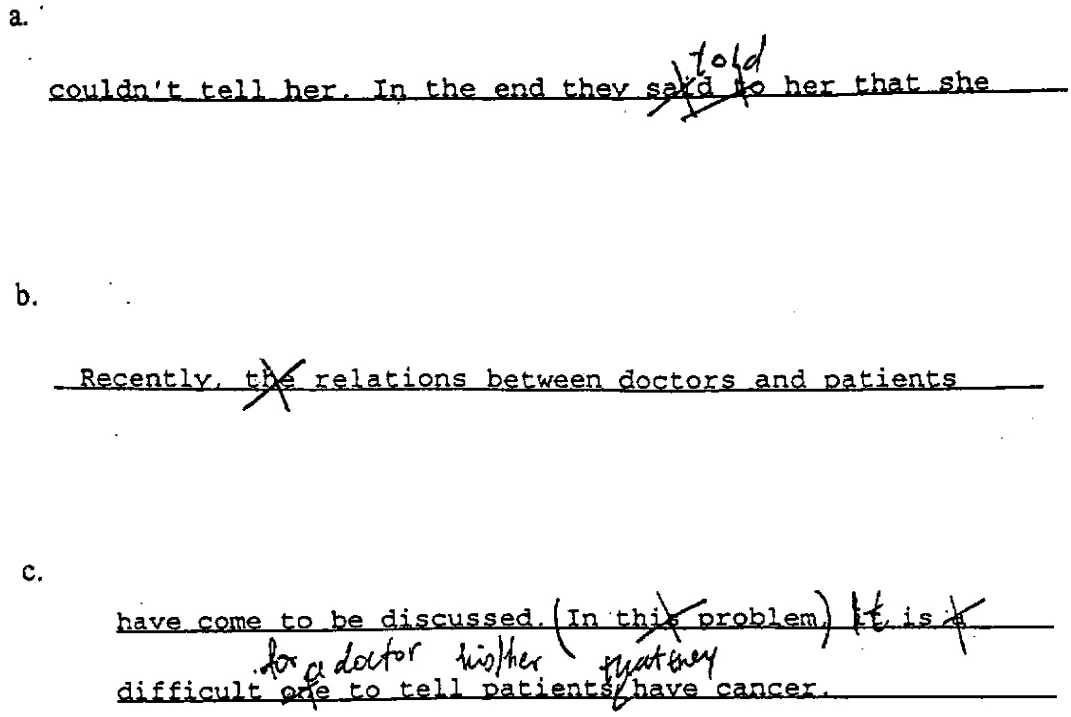


Fig. 6. Examples of Provider annotations

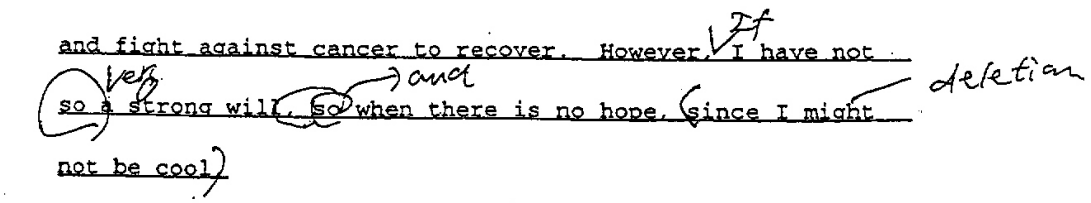


Fig. 7. Examples of Mutator annotations

Word count: 8051 words

* We would like to thank Claire Booth and Maura O'Regan for their work on coding the data and carrying out the statistical analysis and Sandro Leidi from the School of Applied Statistics at the University of Reading for his advice and help in analysing the data. We are indebted to the Centre for Applied Language Studies (now the Department of Applied Linguistics) at the University of Reading for an internal grant that supported the research. We would also like to thank two anonymous reviewers for their comments on an earlier draft of this paper.