CHAPTER SIX

POST-TREATMENT REVISION RESULTS

This chapter is divided into four main sections. In the first section, I will briefly summarize the findings yielded by the application of the three separate taxonomies described in chapter five to the post-treatment revisions by the participants. Cross-references between taxonomies will be left to the next two sections, which focus on the interpretation of the post-treatment revisions from the perspective of readability and feedback-independence. The last main section of the chapter advances some preliminary conclusions about the relationship between readability, feedback-independence and the subsequent diagnosis of writing instruction needs.

6.1 General Findings

My aim in this section is simply to summarize what changed and what should have changed but did not in post-treatment revisions. I will begin by reporting on the number of changes identified in the revisions, and by describing how

they were distributed among the individual participants. After that, I will describe how the changes were distributed according to the reading process, the writing product and the qualification categories, and will comment on factors which may have affected these distributions. Although at this stage I will not attempt to make any cross-references between taxonomies, the full details of the revisions are provided in appendix VI, which lists how each single change was coded according to the three taxonomies of the system.

6.1.1 Distribution of changes

A total of 496 single changes were identified in the eight revisions analysed. Of these, 431 changes were made by the participants themselves, and the remaining 65 changes were additional changes subsequently made by the native-speaker proofreaders. Figures 6.1 and 6.2 summarize how these changes were distributed among the eight participants in the group.

Figure 6.1: Distribution of changes made from T3 to T3*

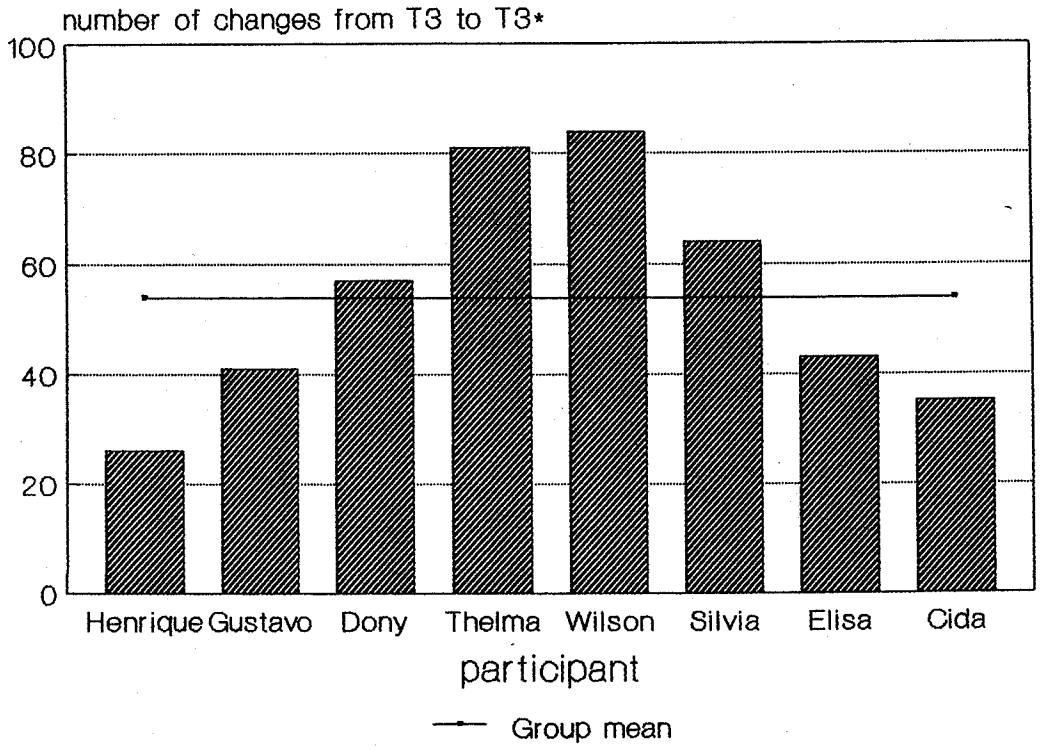
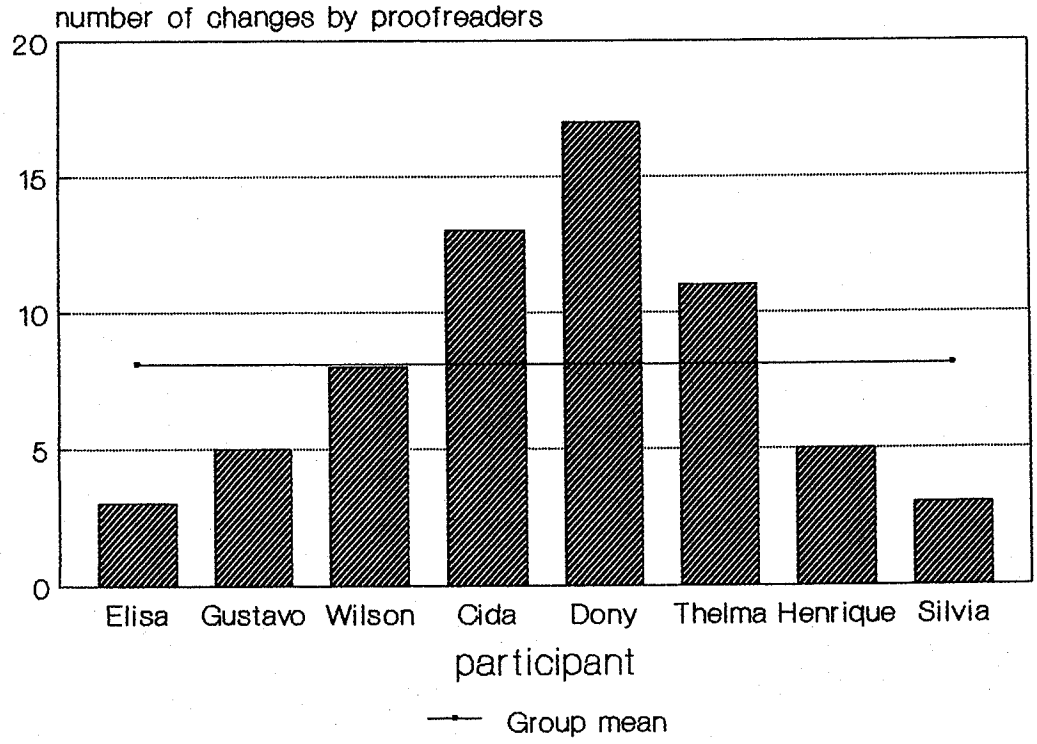


Figure 6.2: Distribution of changes added by proofreaders



As shown in figure 6.1, the number of changes per participant ranged from 26 (Henrique) to 84 (Wilson), the average being 53.9 changes, with a standard deviation of 21.3. A possible explanation for this rather large variance is that the participants began the post-treatment revisions at different starting points, i.e., some pre-treatment texts needed a lot more revision than others. In addition to this, it is also possible that after instruction had ceased some participants felt simply more critical than others about their pre-treatment texts.

Figure 6.2 indicates that the number of changes added by the native speakers after the participants had finished revising varied from 3 (Silvia and Elisa) to 17 (Dony), with an average of 8.1 and a standard deviation of 5.1. These differences at the "finishing-line" can in part be accounted for by the possibility that some participants left more parts of text unrevised than others. In addition to this, the differences shown in figure 6.2 could also be a consequence of some pairs of proofreaders having agreed more often than others, for, as explained in chapter five, only the changes which both native-speaker proofreaders agreed were necessary were taken into account. It is nevertheless worth recalling that none of the native speakers in question disagreed with one another in general terms, for, as said in chapter four, the accepted minimum

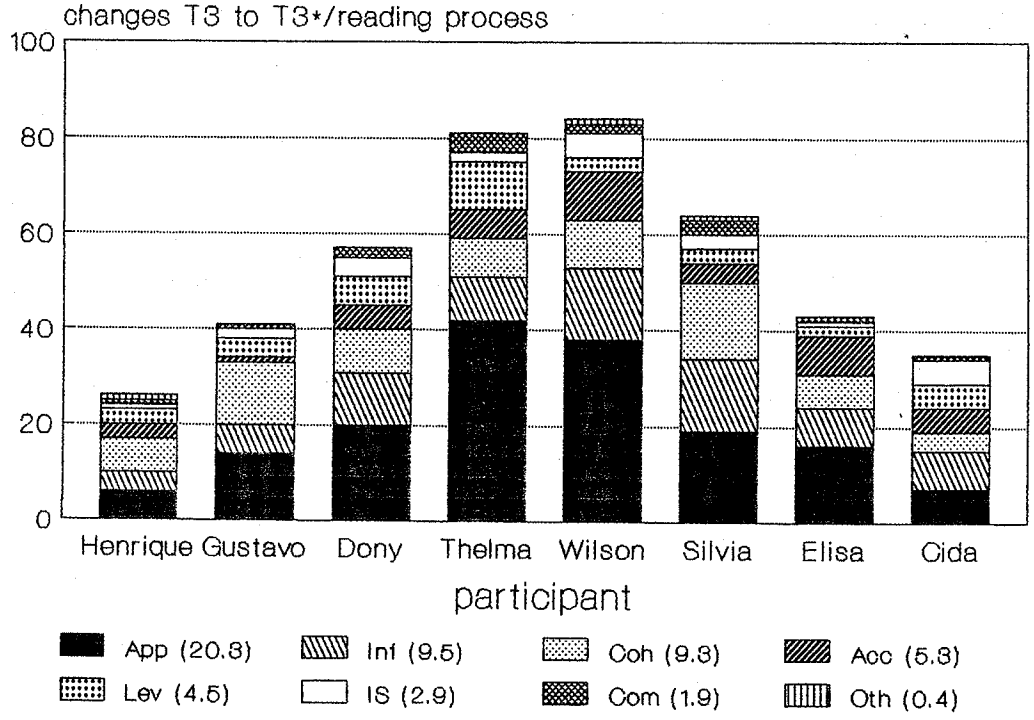
rate of agreement between readers as to impression judgements on the comparative readability of T1 to T6 was a correlation coefficient of +0.5.

To conclude this section, I should also mention that there is no significant relationship between the number of changes made by the participants and the number of changes then added by the proofreaders. That is to say, the correlation coefficient for the two distributions was +0.2, which means that the participants who left many parts of text unrevised were not necessarily those who made the fewest changes.

6.1.2 Distribution of changes according to the reading process categories

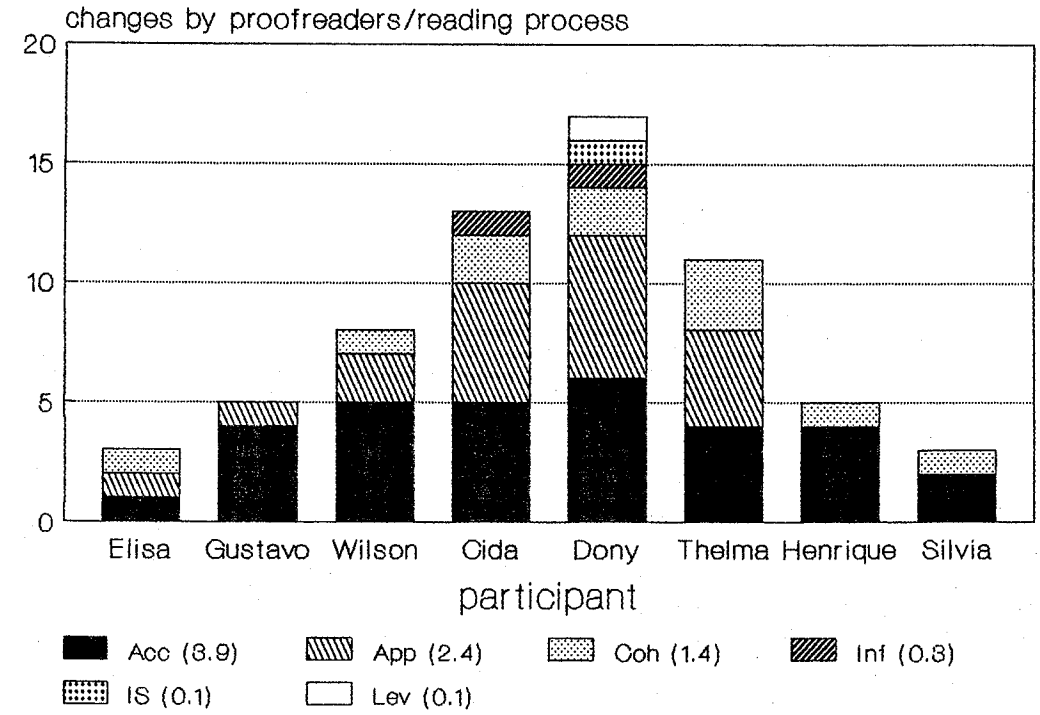
The distribution of the changes made by the participants according to the taxonomy for describing what changed in terms of reading process is summarized in figure 6.3, and figure 6.4 illustrates how the changes made by the native-speaker proofreaders were distributed according to the reading process categories.

Figure 6.3: Distribution of changes made from T3 to T3* according to the reading process categories



values in brackets = group means

Figure 6.4: Distribution of changes added by proofreaders according to reading process categories



values in brackets = group means

From figure 6.3 it is clear that, in almost all revisions, the great majority of changes made from T3 to T3* affected appropriateness. The second comparatively most frequently affected reading process category was informativity, which was closely followed by coherence. The average number of changes affecting accuracy was then much lower. Next came the changes affecting the reading process category for levels effect, and the sixth in the list was information-structure. The changes affecting commitment accounted for only a very small proportion of the changes made from T3 to T3*. Finally, the changes which could not be coded according to any of the above reading process categories, i.e., the changes coded "other", were the fewest of all.

From figure 6.4, in turn, it can be seen that the great majority of changes which should have been made but were not had to do with accuracy. After that came the changes affecting appropriateness, which were closely followed by the ones in coherence. The native speakers then added only a very small number of changes in informativity, information-structure and levels effect, and no changes at all in commitment.

A number of factors may have affected the distribution of the changes according to the reading process categories. The first and most obvious one is that the analysis is

based on texts which may have needed different changes and on changes made by different participants and different proofreaders.

The second factor which may have affected the above results has to do with the experimental treatment itself, which may have placed more emphasis on some components of the reading process than others. For example, the different proportions of changes affecting coherence and accuracy could have to do with the fact that while coherence was explicitely discussed during the presentation of the course handout on connectives, comparatively very little attention was paid to accuracy. Although the experimental treatment factor could not have directly affected the changes by the proofreaders, it may have nevertheless affected what remained for them to change. In other words, there may have been more necessary changes which were unrelated to the treatment than necessary changes which were related to it.

A third factor which may have affected the distribution of the changes by the proofreaders but not the participants in terms of reading process is that only the changes which two different readers unfamiliar with subject-matter agreed were necessary were taken into account. This means that the changes which did not depend on idiosyncractic value-judgements, like probably all changes in accuracy, are a lot more likely to have been taken into account, and that

the proofreaders may have been reluctant to add changes which, like many changes in commitment, could have affected meaning in one way or another.

The last and less obvious of the factors which may have affected the above distributions is that some categories describe changes which can occur a lot less frequently than others in a text of limited length. For example, in a short text there can be many more changes pertaining to a category like appropriateness than changes pertaining to a category like levels effect: the number of changes in text-hierarchy (levels effect) which can be made in an essay which is roughly only two A4 pages long is simply a lot more limited than the number of changes in usage and style (appropriateness) which can be made in that same text.

The connection between what the proofreaders and what the participants changed in terms of reading process varied a lot from revision to revision. In Elisa's, Dony's and Thelma's revisions there was a certain amount of agreement between what the participants and the proofreaders changed in terms of reading process, for the correlation coefficients for the two distributions varied from +0.5 (Thelma's revision) to +0.7 (Elisa's revision). In the revisions by the remaining five participants, however, these same coefficients varied from +0.4 (Cida's revision) to -0.1 (Gustavo's and Silvia's revisions), indicating that

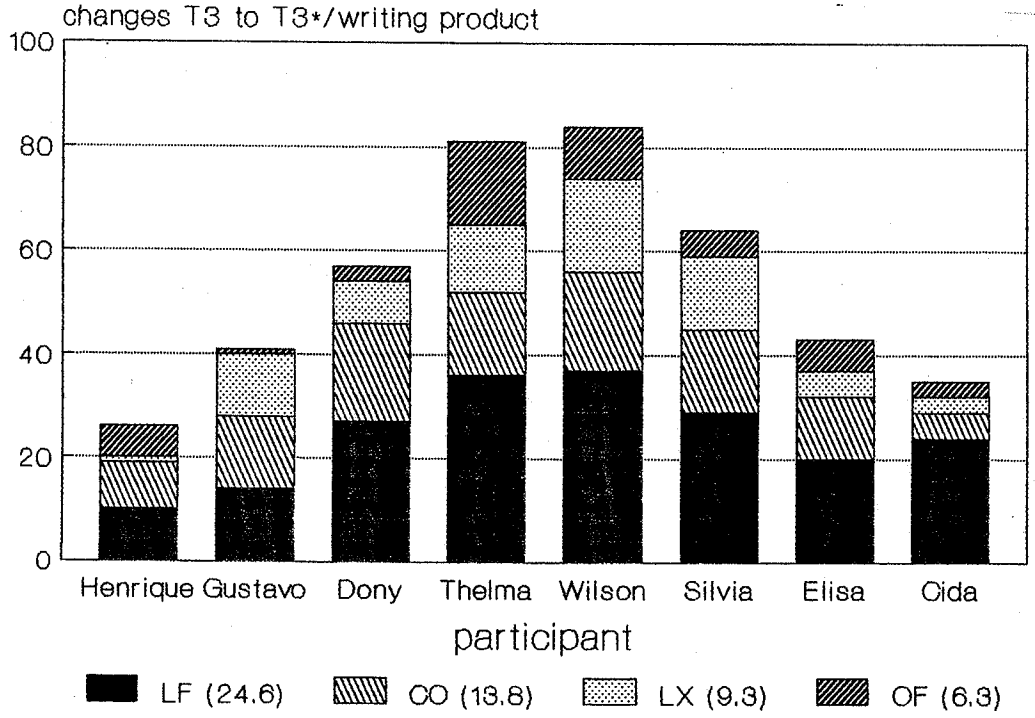
the changes in reading process by the participants and by the proofreaders were comparatively much more unrelated.

Having said this, in the next section I shall describe how these same changes were distributed according to the writing product categories.

6.1.3 Distribution of changes according to the writing product categories

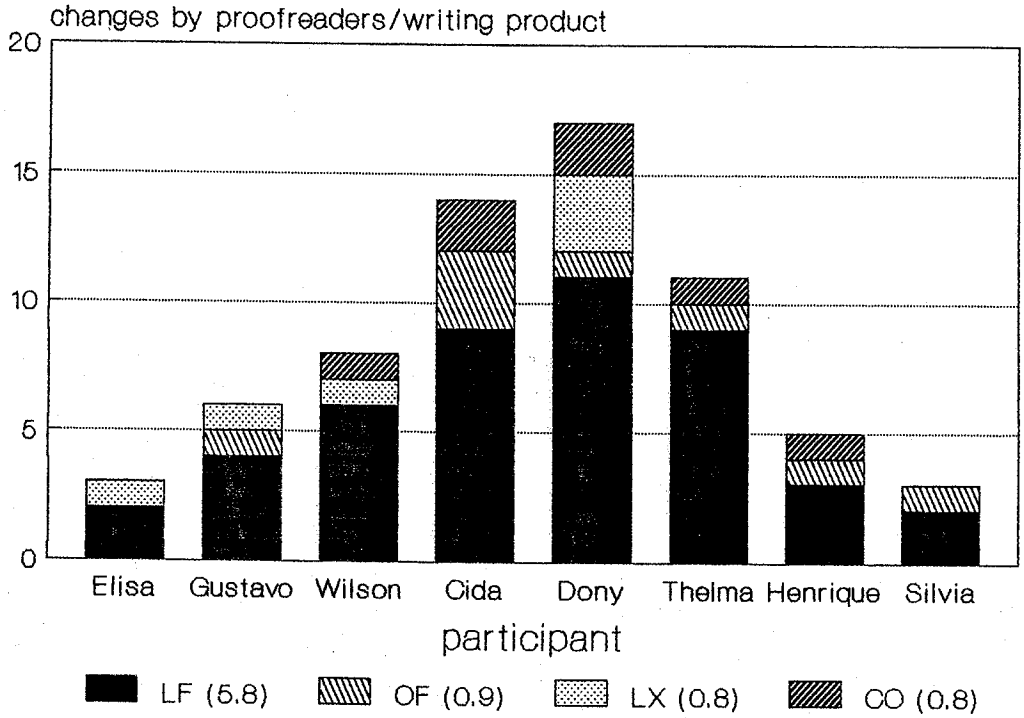
The distribution of the changes made from T3 to T3* according to the four macro-categories for describing revision from the viewpoint of writing product is summarized in figure 6.5, and the corresponding distribution of the changes by the proofreaders is shown in figure 6.6.

Figure 6.5: Distribution of changes made from T3 to T3* according to writing product



values in brackets = group means

Figure 6.6: Distribution of changes added by proofreaders according to writing product



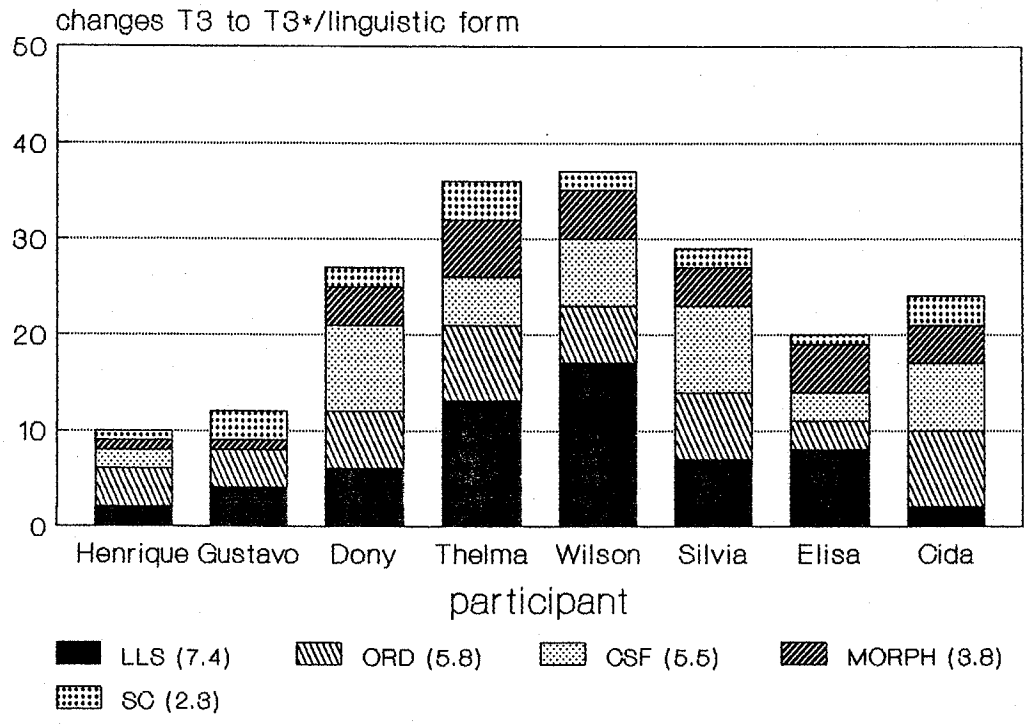
values in brackets = group means

Figure 6.5 shows that all participants gave priority to the revision of linguistic form. The second comparatively most frequent changes were those in content. Next came the changes in lexis, and the changes in orthographic form, in terms of group averages, were the least frequent ones of all. As to what should have been revised but was not, it can be seen from figure 6.6 that most changes introduced by the native speakers had to do with linguistic form. They then added an almost equal proportion of changes in content, lexis and orthographic form.

There seems to have been a very explicit connection between what the participants and what the proofreaders changed in terms of writing product, for the correlation coefficients for the two distributions varied from +0.5 (Gustavo's revision) to +1.0 (Cida's, Thelma's and Wilson's revisions). This means that the participants and the proofreaders tended to make the same general types of changes in writing product.

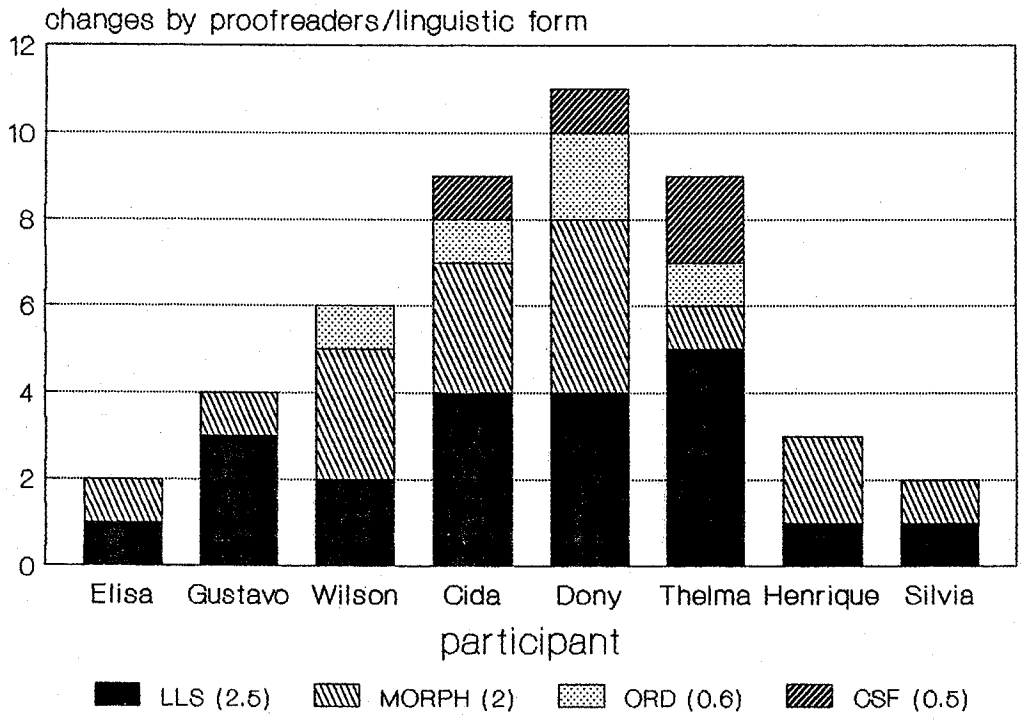
Going down the hierarchy for describing what changed in terms of writing product, the changes made from T3 to T3* pertaining to the sub-categories immediately below linguistic form were distributed as shown in figure 6.7 below. Figure 6.8 then summarizes the distribution of the changes in linguistic form which were subsequently added by the proofreaders.

Figure 6.7: Distribution of changes made from T3 to T3* according to sub-categories of linguistic form



values in brackets = group means

Figure 6.8: Distribution of changes added by proofreaders according to sub-categories of linguistic form



values in brackets = group means

From figure 6.7 it is clear that the changes in linguistic form made from T3 to T3* were predominantly those in lower-level syntax. The changes involving the reordering of elements in text were comparatively less frequent, and were closely followed by the changes in commutable syntactic forms. Next came the changes in morphology, and the least frequent changes of all were those in sentence-complexity. From figure 6.8 it can then be seen that most of the necessary changes in linguistic form had to do with lower-level syntax and morphology. Much less frequent were the necessary changes in order and commutable syntax, and there were no necessary changes in sentence-complexity.

When the distribution of the changes in linguistic form by the participants and the proofreaders were then compared, it was found that the changes in linguistic form by two of the participants (Elisa and Thelma) were proportionally very similar to those by the proofreaders, for the correlation coefficients for the two distributions were in both cases +0.9. Conversely, Henrique's and Cida's changes in linguistic form were relatively different from the changes in linguistic form by the proofreaders, for the two correlation coefficients were -0.5 and -0.6 respectively. The remaining coefficients were close to zero, which means that there was little or no connection between the changes in linguistic form by the other four participants and the proofreaders.

At this point it should be recalled that the reason why the categories within the taxonomy for describing changes in writing product were organized hierarchically is that I had predicted that the changes pertaining to some of the lowest-level categories in the hierarchy might be too infrequent to be analysed on their own right. It was therefore determined that the changes pertaining to these categories would only deserve separate attention later on in this chapter if they were represented by sixteen or more records, i.e., changes by the participants plus changes by the proofreaders, in the overall distribution. This means that there had to be an average of two or more records of those changes per revision for them to be considered representative enough to be analysed on their own right. The analysis of the categories which did not reach this criterion should be understood in the context of the analysis of the category immediately above it in the hierarchy which rules the taxonomy. For example, since the number of changes in spelling was below sixteen, the analysis of spelling is to be understood in the more general context of the analysis of orthographic form. Conversely, since the number of changes in punctuation was above sixteen, punctuation was considered representative enough to be analysed separately.

Figures 6.9 to 6.16 below summarize the overall distribution of the changes pertaining to the lowest-level categories within the taxonomy for describing the revision

of writing product. Attention is drawn to the pre-established criterion of sixteen, which determines which categories will later on be analysed separately and which will not.

Figure 6.9: Overall distribution of content changes

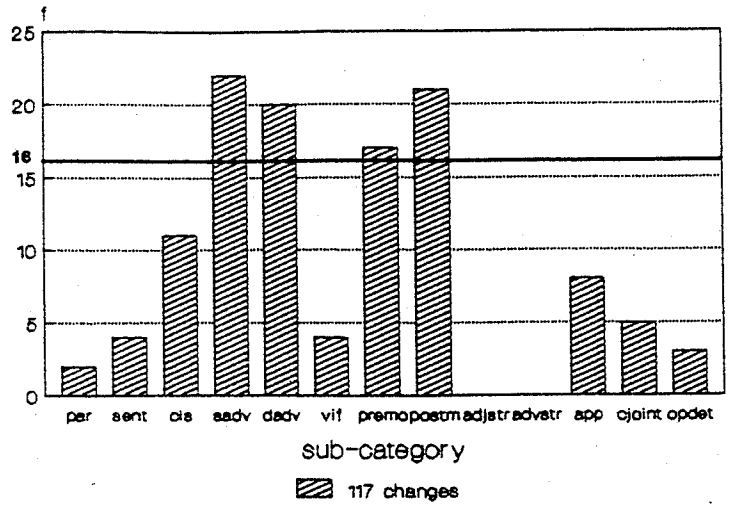


Figure 6.10: Overall distribution of changes in lexis

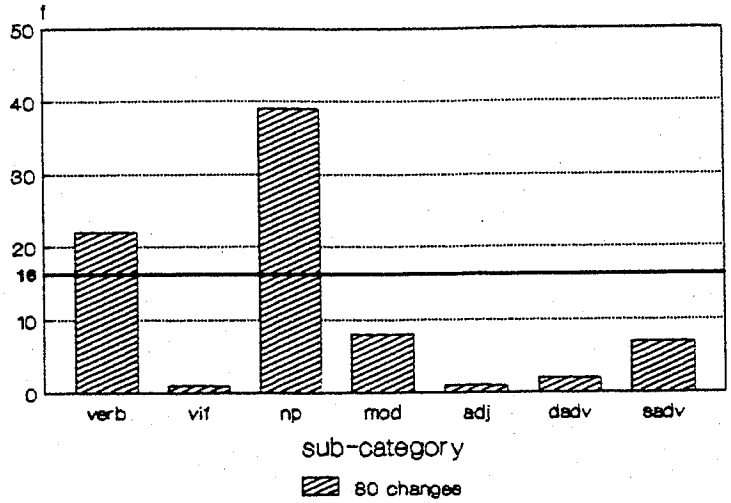


Figure 6.11: Overall distribution of changes in morphology

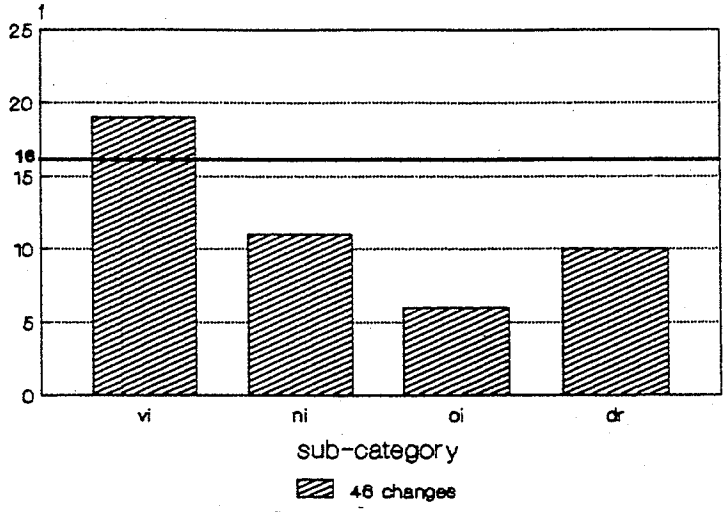


Figure 6.12: Overall distribution of changes in lower-level syntax

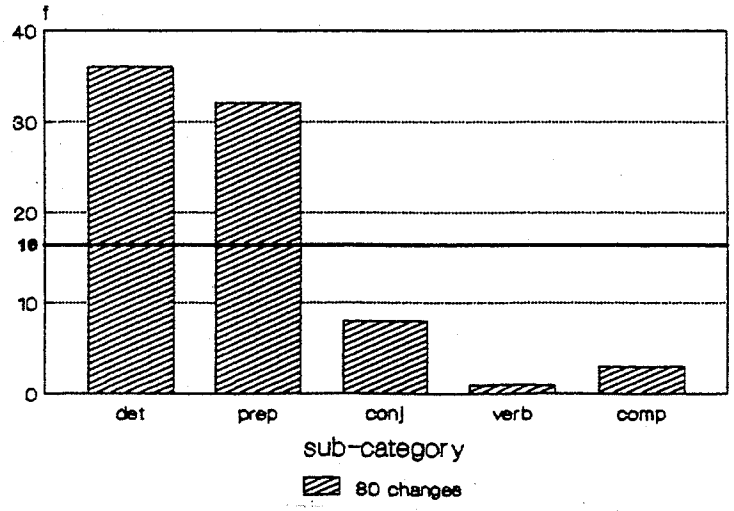


Figure 6.13: Overall distribution of changes in commutable syntax

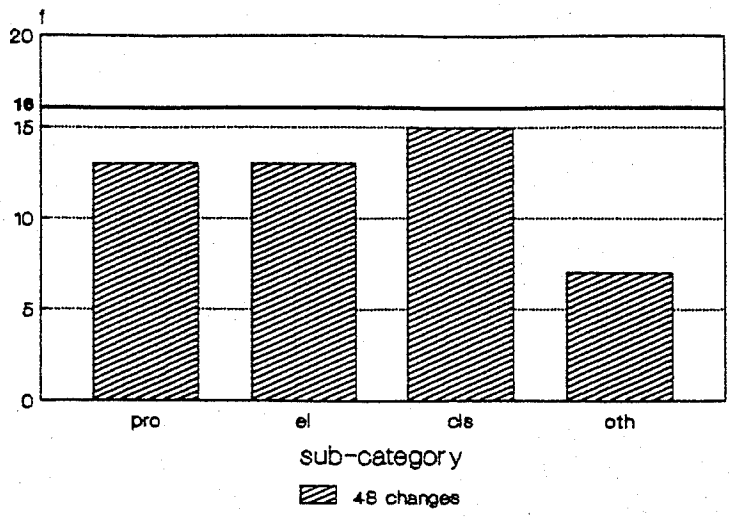


Figure 6.14: Overall distribution of changes in sentence-complexity

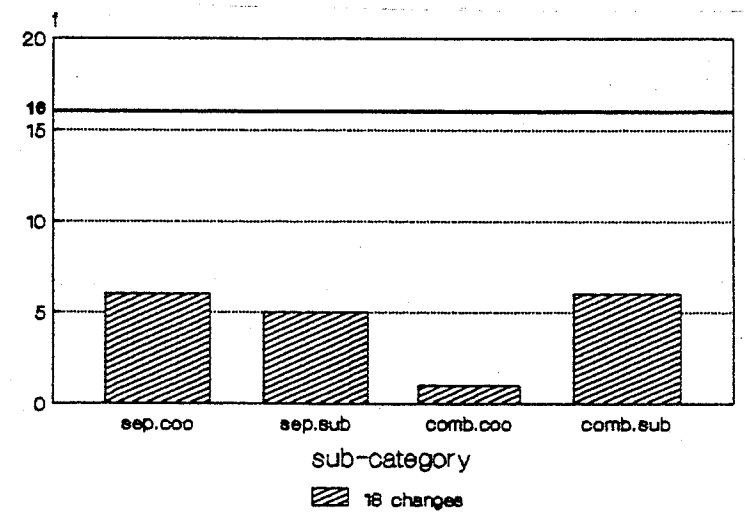


Figure 6.15: Overall distribution of changes in order

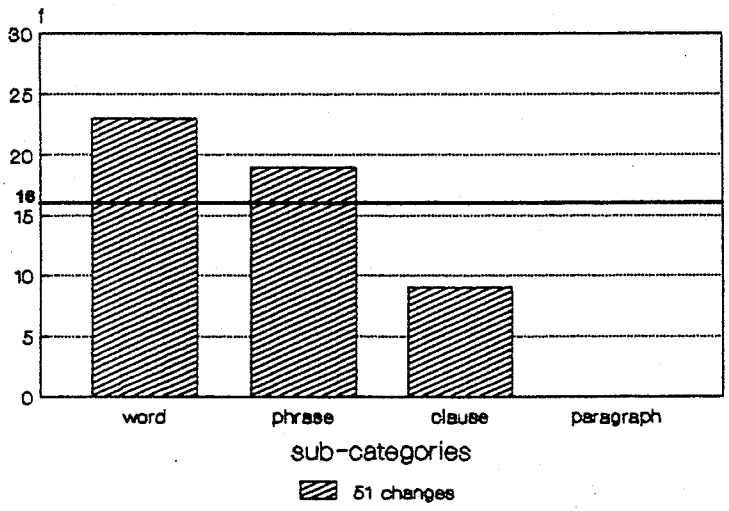
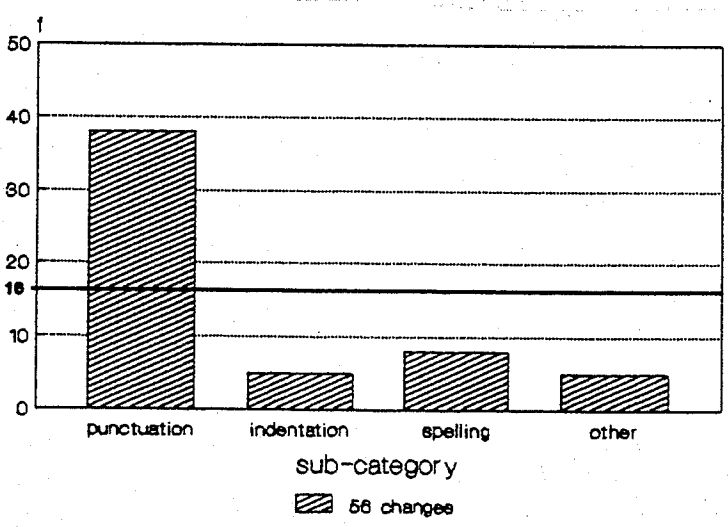


Figure 6.16: Overall distribution of changes in orthographic form



From the above it is clear that only the following lower-level categories were frequent enough to deserve separate attention later on in this study:

CONTENT - the addition or deletion of sentence and descriptive adverbials plus pre and postmodifiers;

LEXIS - verb and noun phrase lexis;

MORPHOLOGY - verb inflection;

LOWER-LEVEL SYNTAX - determiners and prepositions;

ORDER- word and phrase order;

ORTHOGRAPHIC FORM - punctuation.

In contrast to this, the analysis of the changes pertaining to all other lower-level categories is to be understood in the context of the category immediately above them in the hierarchy which rules the taxonomy.

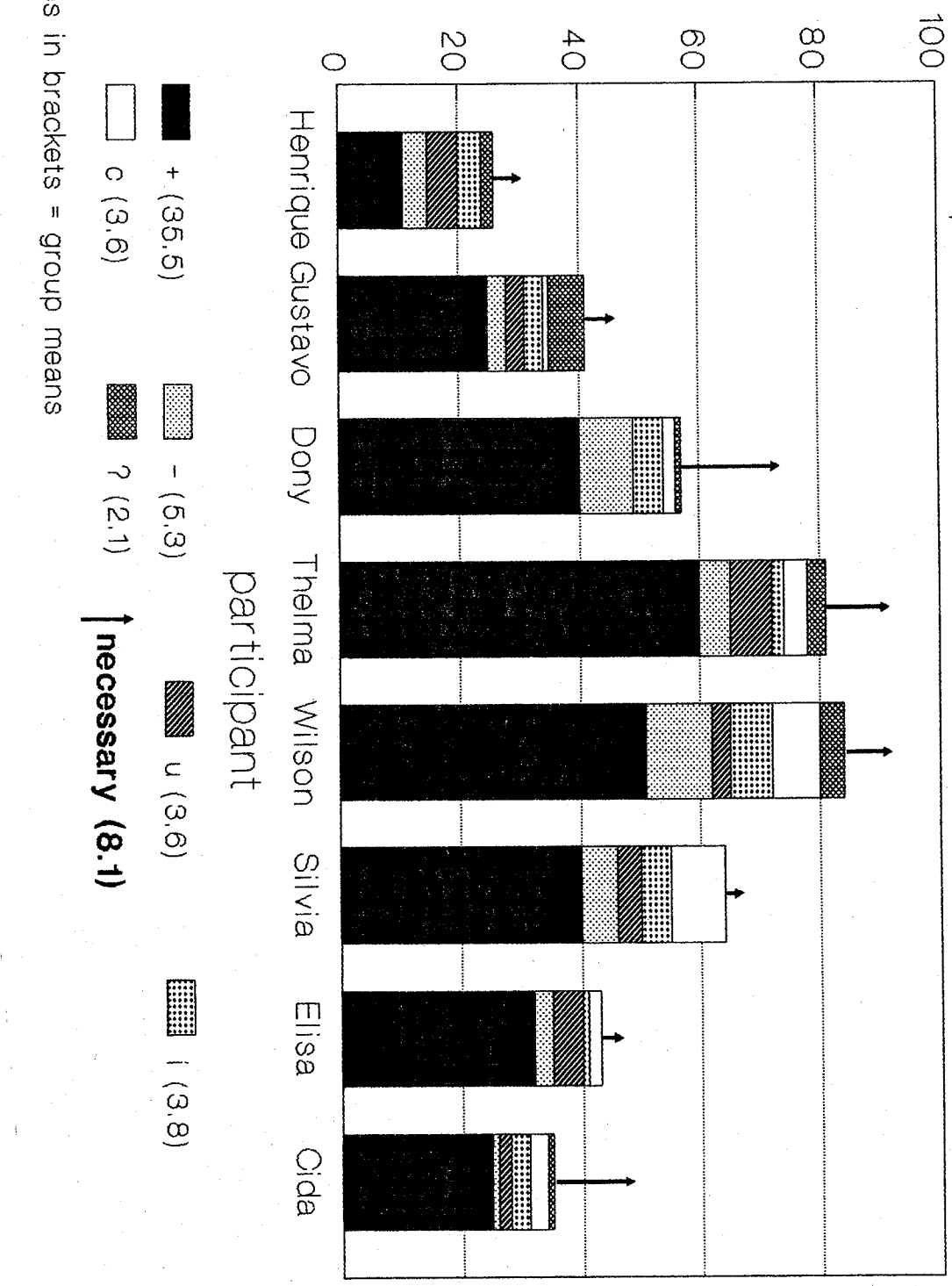
The factors which may have affected the distribution of the changes according to the writing product categories are similar in principle to the ones which may have affected the distribution of the changes according to the reading process categories, for it is only the categories, and not

the changes, that are different. Thus all one has to do is look at those factors from the perspective of writing product, as opposed to reading process. More specifically, the distribution of the changes according to the writing product categories may have been affected by the following: the pre-treatment texts may have needed different changes in writing product and the participants and proofreaders who decided what to change were different; the treatment gave probably more emphasis to some aspects of writing product than others; some writing product categories describe changes which can occur a lot less frequently than others in a short text; and only the additional changes which two different proofreaders unfamiliar with the subject-matter of the essays agreed were necessary were taken into account.

6.1.4 Distribution of changes according to the qualification categories

The analysis of the revisions from the perspective of the qualification categories alone is summarized in figure 6.17 below.

Figure 6.17: Distribution of the changes made from T3 to T3* and the changes by the proofreaders according to the qualification categories



As can be seen, the changes qualified as positive were by far the most predominant ones. In terms of group averages, it is also possible to say that the second most predominant changes were the necessary ones, that the negative changes came next, that the average number of ineffective, unnecessary and consequential changes was almost the same, and that the indeterminate changes were comparatively the fewest of all.

It is clear, however, that there was a lot of variability with regard to the qualification of changes in the individual revisions. In Cida's revision, for example, there were comparatively many changes which she should have made but did not (necessary), and only one change which did more harm than good (negative). This particular combination could indicate that Cida was the risk-avoider of the group.

Dony's revision stands out in that none of his changes were unnecessary. Since the unnecessary changes disclose the cases in which the participant was insecure as to whether revision was really necessary, Dony (the journalist in the group) seems to have behaved like a very confident writer. The pitfall was of course that the proportion of necessary changes in his revision was well above average.

Elisa's and Thelma's overall revisions were in turn markedly more positive and less ineffective than average. Silvia's revision then had the third greatest percentage of positive changes and only a very small proportion of necessary changes. In terms of overall qualification, their revisions seem to have been the three best in the group.

The two most average revisions in the group with respect to the qualification of the changes seem to have been Gustavo's and Wilson's. While the only marked feature in Gustavo's revision was high percentage of indeterminate changes, Wilson's revision was totally unmarked in terms of overall qualification.

In contrast to Gustavo's and Wilson's revisions, Henrique's was by far the most deviant one in the group in terms of overall qualification. On the one hand, his positive changes were comparatively a lot fewer than average, and he made no consequential changes at all. On the other hand, the proportions of negative, ineffective and unnecessary changes in his revision were well above what was average for the group. Clearly, Henrique's revision seems to have been the least successful one of all.

Many of the above differences can be accounted for by the possibility that some participants were better able to spot what needed revision in their pre-treatment texts and the possibility that, when they did see what needed revision,

some participants were simply better able to improve their texts. In addition to these individual variables, the experimental treatment too is likely to have been the cause of the above differences, for its relevance with regard to what needed revision in the first place was not the same for all pre-treatment texts. In other words, the revision of texts which contained many problems that were addressed during the treatment and the revision of texts which contained many problems that were not discussed during the treatment may have been qualitatively different. The last factor which may have affected the distribution of the changes according to the qualification categories has to do with the kind of changes made by the participants. Because only the changes which two different proofreaders agreed were necessary were taken into acount, it is likely that the proportion of necessary changes was greater in the revisions by participants who did not pay too much notice to the correction of certain elements, such as spelling and grammar, the necessity of which should not cause any disagreement between proofreaders.

In the next two sections of this chapter, cross-references which disclose information about changes in readability and feedback-independence will be made.

6.2 Readability

My objective in the present section is to compare the readability of the post-treatment revisions and pre-treatment final drafts by decomposing readability into the seven main reading process categories of the system of analysis[1]. This will enable me to find out how exactly readability changed from one version of text to the other, and hence test the following hyptheses, which are taken to be part of the more general hypothesis that the post-treatment revisions are more readable (H2):

a. The post-treatment revisions are more coherent than the pre-treatment final drafts;

b. The distinction between main points and supporting details of text is clearer in the post-treatment revisions;

c. The post-treatment revisions are less over or under-informative than the pre-treatment final drafts;

d. The degree of commitment to the truth of what is asserted in text is more convincing to the reader in the post-treatment revisions;

e. The reader's expectations as to the sequence of information in text are better fulfilled in the post-treatment revisions;

f. Usage and style are more appropriate in the post-treatment revisions;

g. The post-treatment revisions infringe fewer grammar and spelling conventions than the pre-treatment final drafts.

To test the above hypotheses, it was necessary to retrieve coding of the revisions according to both the reading process categories to which they are related, and the two qualification categories which have a directional effect upon readability, i.e., positive and negative. As said in chapter five, the ineffective changes do not affect readability because such changes are about infelicitous elements in text which were replaced by equally infelicitous equivalents. Likewise, the unnecessary changes do not influence readability in any specific direction because they are about felicitous elements in text which were replaced by other, equally felicitous ones. The indeterminate changes, in turn, have to be ignored if one wishes to obtain a realistic measure of what changed from one version of text to the other in terms of readability simply because they were changes which could not be evaluated. Also, the consequential changes cannot be included in the comparative analysis of the readability of the two texts because they were changes which were introduced as a result of other changes, which means that although they may affect the readability of T3* in relation to a comparative T3* without any consequential changes,

they do not affect the readability of T3* in relation to that of T3. Finally, the necessary changes must also be excluded from the interpretation of the comparative readability of T3 and T3* since they were changes which were added by the proofreaders after the participants had finished revising.

Thus of the total number of changes identified in the transcriptions, only the positive and negative changes, i.e., the changes which have a directional effect upon readability, were taken into account in the present part of the study. As shown in table 6.1 below, an average of 74.2% of the total number of changes per participant had a directional effect upon readability.

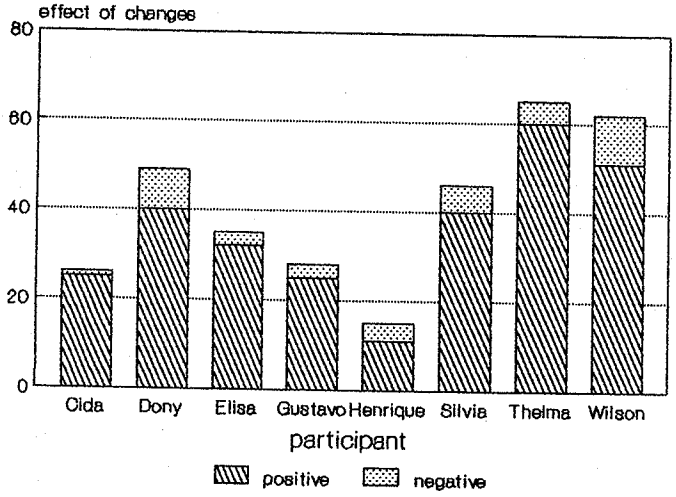
Table 6.1: Distribution of total number of changes per participant (T) and of changes with a directional effect upon readability (N)

PARTICIPANT	N	T	N/T%
Cida	26	35	74.3
Dony	49	57	86.0
Elisa	35	43	81.4
Gustavo	28	41	68.3
Henrique	15	26	57.7
Silvia	46	64	71.9
Thelma	65	81	80.2
Wilson	62	84	73.8
Total	326	431	
Mean	40.7	53.9	74.2
SD	17.8	21.3	8.8

Having determined which changes are relevant to the comparison of the readability of the post-treatment revisions and pre-treatment final drafts, the distribution

of such changes according to those which enhanced and those which hindered readability is summarized in figure 6.18 below.

Figure 6.18: Distribution of changes with a directional effect upon readability



As can be seen, although readability was affected to different extents in the different revisions, the total number of positive changes, which enhanced readability in one way or another, was greater than the total number of negative changes, which hindered readability, in all eight revisions. Having said this, it should be noted that figure 6.18 only gives a very vague, if not distorted, idea of how readability was generally affected by the revisions, for it converges changes which affected different aspects of readability, some of which may carry more weight than others. The effect upon readability of, for instance, ten positive changes in accuracy and five negative changes in

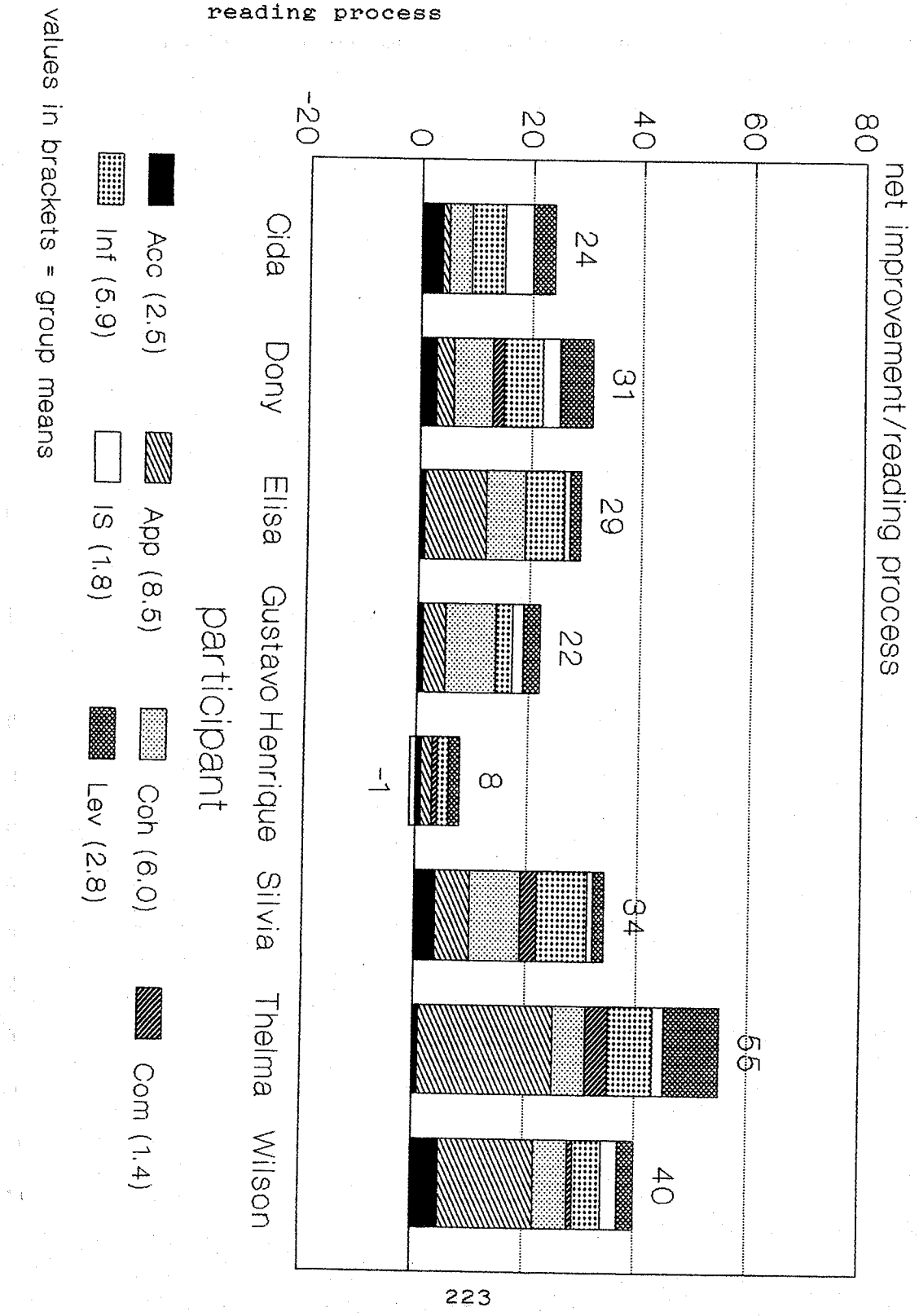
coherence can be very different from that of ten positive changes in coherence and five negative changes in accuracy.

To obtain a more accurate picture of how readability was affected and then test hypotheses (a) to (g), it was necessary to distribute the positive and negative changes by each participant according to the system's reading process categories, and compare the number of positive and negative changes for each separate category. This of course implies that for convenience I am assuming that the positive and negative changes carry equal weight once they have been distributed according to the reading process categories. That is to say, if on the one hand it is misleading to assign equal weight to the positive and negative changes pertaining to categories so diverse as, say, coherence and accuracy, it is on the other hand legitimate to compare the positive and negative changes in coherence alone in order to find out whether or not coherence improved. Thus while it is meaningless to compare a positive change in coherence with a negative change in accuracy when assessing their combined effect upon readability, it seems operationally reasonable to assume that one positive and one negative change in coherence cancel each other out.

Figure 6.19 below therefore summarizes the revisions' net effect upon readability after the positive and negative changes by each participant were sorted out according to

the reading process categories. The values shown were calculated by subtracting the negative changes for each category from the positive ones.

Figure 6.19: Net effect upon readability of changes in reading process



From the above it can be seen that although the revisions seem to have had a generally positive effect upon the various reading process categories in which readability was decomposed, there was a considerable amount of variability with respect to the ways in which readability was affected in the different revisions. I will leave the differences between reading process categories to a later part of the analysis, and will comment on the differences between participants first. This will enable me to check whether the present atomistic approach towards comparing the readability of T3 and T3* is consistent with the corresponding holistic impression judgements on readability that, as explained in chapter five, eight pairs of readers had been asked to supply.

6.2.1 Differences between participants

Thelma's revision was the one with the greatest number of changes with a directional effect upon readability. Most of those changes tied in with appropriateness, which, like levels effect, informativity and commitment, improved more than average in Thelma's revision. In addition to this, Thelma's changes pertaining to the remaining reading process categories into which readability was decomposed were also predominantly positive. In accordance with these results, the pair of readers responsible for comparing her texts agreed that the revision was more readable.

Wilson's revision too had a comparatively very large number of changes affecting readability and, like in Thelma's revision, resulted in an above average improvement in appropriateness. Accuracy and information-structure also improved more than average in Wilson's revision, and the effect of his changes pertaining to the four other reading process categories was more often positive than negative. In agreement with this, the two readers who evaluated Wilson's revision found it more readable than his pre-treatment version of the same text.

Dony's positive and negative changes were fewer than Thelma's and Wilson's but were nevertheless greater in number than the changes with a directional effect upon readability in the other revisions. When Dony's positive and negative changes were distributed according to the reading process categories, it became clear that in his revision coherence, informativity, levels effect, information-structure, accuracy and commitment had improved more than in the average revision. Also, Dony made more positive than negative changes in appropriateness. Accordingly, the readers who assessed his revision found it more readable than the pre-treatment text.

Silvia's changes with a directional effect upon readability were almost as frequent as Dony's. Not only was there improvement with respect to all reading process categories, but also accuracy, commitment, coherence and informativity

improved more than average in Silvia's revision. Although this indicates readability must have certainly improved, one of the readers responsible for comparing Silvia's revision with her pre-treatment final draft disagreed with his co-reader and thought the revision was less readable. There seems to be nothing in the actual revision to justify this divergent opinion, and it is contradictory that the reader who found Silvia's revision less readable added very few extra "necessary" changes when asked to revise and proofread the text so as to enhance its readability. There is however a plausible external explanation for his negative impression judgement: Silvia's essay was about pharmacology, and the reader in question, as he belatedly informed me, was an experienced teacher of medical English. When asked whether he had evaluated readability as required, he admitted it being possible that his evaluation was based on the content of the courses he taught more than on readability alone.

The revision by Elisa contained a below average number of changes with a directional effect upon readability, and none of them had to do with commitment. However, there was improvement with respect to all other categories into which readability was decomposed, and appropriateness, coherence and informativity improved more than in the average revision. This time there was no disagreement between

readers, both of whom, in accordance with the present analysis, found Elisa's revision more readable than her pre-treatment text.

The changes affecting readability in Gustavo's revision were slightly fewer and, like the ones by Elisa, did not affect commitment. The changes pertaining to all other reading process categories were more frequently positive than negative, and what was particularly marked in Gustavo's revision was an above average improvement in coherence. Again, in agreement with the present analysis, the pair of readers responsible for comparing the two versions of text found the post-treatment revision more readable. Without having been asked to do so, one of the readers even emphasized the point in affirming that the revision was "a lot more readable". This could mean that Gustavo's changes in coherence played a very important role in enhancing readability.

Cida did not make many changes which affected readability, but of these, all except one were positive. Commitment again did not change, but there was improvement with respect to all other reading process categories. In addition to this, accuracy, informativity and information-structure improved more than average. Accordingly, the two readers who evaluated Cida's texts felt her post-treatment revision was more readable than her pre-treatment final draft.

Finally, Henrique's revision affected readability to a much lesser extent than that of his colleagues, and very few of his changes were positive. The only two changes in coherence in his revision cancelled each other out (one was positive but the other, negative), and his changes in information-structure had an overall negative effect upon readability. Although there was apparently some improvement with respect to the remaining five reading process categories, it was below average if compared with the improvement in the revisions by the other participants.

The two native speakers who compared Henrique's pre-treatment final draft with his post-treatment revision did not perceive any overall improvement, and actually found the earlier version of text more readable. A possible explanation for their failure to detect an overall improvement in readability - despite the fact that the changes pertaining to five reading process were more frequently positive than negative - is that the differences between the two was so small that their combined effect upon overall readability was imperceptible. And indeed, when matched t-tests were applied in order to compare the positive and negative reading process changes by the participants, Henrique's were the only ones which were not significantly different at the 0.05 level. Because the differences between Henrique's positive and negative

reading process changes were not significant, the latter may have influenced his readers' perceptions slightly more than the former, which explains why the native speakers actually found Henrique's revision **less** readable.

To summarize, although the participants changed readability to different extents and in different ways from T3 to T3*, seven of the eight post-treatment revisions were considered to be more readable than the pre-treatment final drafts according to both the system's atomistic analysis of the revisions and the holistic impression judgements supplied by fifteen out of sixteen different NS readers (the only discrepant reader being the medical English teacher responsible for evaluating Silvia's texts). In the next section, I will focus on the differences between categories so as to determine which changes in reading process were more successful and find out which of those changes could have actually contributed towards improved readability.

6.2.2 Differences between categories

To begin with, if one refers back to figure 6.19, it can be seen that when the negative reading process changes were subtracted from the positive ones, the category with the

biggest average net improvement was appropriateness (+8.5) and the one with the smallest average net improvement was commitment (+1.4). However, these were also the categories with the greatest and smallest average number of changes with a directional effect upon readability. In fact, if one correlates average net improvement per category with average number of positive and negative changes per category, the coefficient obtained is +Ø.9. This means that the more positive and negative changes there were, the bigger was the net improvement observed. It therefore does not make sense to compare one reading process category with another in terms of net improvement, for, as said in section 6.1.2, some categories describe reading process changes which can occur a lot less frequently than others in a text of limited length.

It makes a lot more sense to compare one category with another in terms of positive/negative ratios for, irrespective of the number of positive and negative changes per category, they tell us how many more positive than negative changes there were for each category. The results obtained are summarized in table 6.2 below.

Table 6.2: Distribution of positive and negative changes according to reading process categories

	APP		ACC		COH		COM	
PARTICIPANT	+	-	+	-	+	-	+	-
Cida	1	0	4	0	4	0	0	0
Dony	10	7	4	1	7	0	2	0
Elisa	11	0	4	3	7	0	0	0
Gustavo	6	2	1	0	10	1	0	0
Henrique	3	1	2	1	1	1	1	0
Silvia	8	2	4	0	10	1	3	0
Thelma	26	2	3	2	7	1	4	0
Wilson	23	6	6	1	6	0	1	0
MEAN	11	2.5	3.5	1	6.5	0.5	1.4	0
SD	9.0	2.6	1.5	1.1	3.0	0.5	1.5	0
+:- overall ratio	4.4		3.5		13		*	

Table 6.3 (continued):

	INF		IS		LEV	
PARTICIPANT	+	-	+	-	+	-
Cida	7	1	5	0	4	0
Dony	8	1	3	0	6	0
Elisa	7	0	1	0	2	0
Gustavo	3	0	2	0	3	0
Henrique	2	0	0	1	2	0
Silvia	12	3	1	0	2	0
Thelma	8	0	2	0	10	0
Wilson	8	3	4	1	3	0
MEAN	6.9	1	2.3	0.3	4	0
SD	3.1	1.3	1.7	0.5	2.8	0
+:- overall ratio	6.9		9		*	

As can be seen, there was a considerable amount of variability in the overall ratios of positive and negative changes for each separate reading process category, although for all types of reading process changes the total number of positive changes was greater than the total

number of negative changes. When it comes to comparing one category with another in terms of overall ratios, at the least fortunate extreme are the changes in accuracy and appropriateness. The changes which improved accuracy were only 3.5 times more frequent than the changes which compromised accuracy, and the positive changes in appropriateness were only 4.4 times more frequent than the negative ones. The overall results for informativity were slightly better, for there were 6.9 times as many changes which made the post-treatment revisions as informative as was required than changes which made them either less or more informative than necessary. The positive/negative ratios were much higher for information-structure and coherence. In terms of information-structure, the changes which made the sequencing of ideas in text more predictable to the reader were 9 times more frequent than the changes with the opposite effect; in terms of coherence, the changes which made the post-treatment revisions more coherent than the pre-treatment final drafts were 13 times more frequent than the changes which made T3* less coherent. Finally, it is notable that none of the changes in levels effect or commitment hindered readability.

Although the above results give some idea of which changes in reading process were more and which were less successful, it must be recalled that they are based exclusively on group totals, and may therefore flatten out individual profiles in an unrealistic way. Table 6.3

therefore summarizes the results for the statistical comparison of the positive and negative means for each separate reading process category.

Table 6.3: Comparison of positive and negative means for each separate reading process category (* all values are significant at the 0.05 level for one-tailed test)

CATEGORY	T-matched*
app	2.921
acc	4.183
coh	5.796
com	2.376
inf	6.905
is	3.190
lev	4.073

From the above it can be seen that for all categories the positive changes were significantly more frequent than the negative ones at the 0.05 probability level. This implies the changes pertaining to all seven reading process categories into which readability was decomposed must have enhanced more than hindered overall readability.

6.2.3 Are the post-treatment revisions more readable than the corresponding pre-treatment final drafts?

The results presented in 6.2.1 indicate that the post-treatment revisions by seven of the eight participants were more readable than their corresponding pre-treatment final

drafts. The results supplied in 6.2.2, in turn, indicate that the changes pertaining to all reading process categories were significantly more positive than negative. Although some reading process changes may have helped enhance readability more than others, and although some revisions were probably more successful than others, it is not possible to determine exactly which revisions were more successful and which reading process changes helped enhance readability more. After all, the participants made different changes in reading process, revised some components of reading process better than others, and some reading process changes carry simply more weight than others. Still, since the globality of the results point towards improved readability in seven individual revisions plus improvement with respect to all categories into which readability was decomposed, my overall conclusion is that after instruction had ceased the participants were able to improve the readability of their pre-treatment final drafts. Moreover, the fact that the above conclusion - which was reached via the system of analysis developed in chapter five - is in accordance with the impression judgements by fifteen out of sixteen different native-speaker readers conversant with the discourse of English expository prose seems to constitute proof that the system's atomistic approach towards the data is consistent with holistic impression judgements on readability.

6.3 Feedback-independence

In this section the pre-treatment final drafts and the post-treatment revisions will be interpreted from the viewpoint of feedback-independence. More specifically, my first concern is to find out whether the revisions disclose evidence of increased feedback-independence, and my second concern is to investigate in what respects feedback-independence may have increased. The former will enable me to test H3, i.e., that the revisions contain evidence of an increase in feedback-independence, and the latter will help diagnosing the kind of feeback needed by the participants.

It is already known that the results presented in the previous section indicate that after instruction had ceased the participants were generally able to revise their own essays in a way which improved overall readability. Such evidence must not, however, be equated with evidence of an increase in feedback-independence. This claim is based on the following considerations:

I. Is feedback-independence analysable only in terms of the end-product of revision?

The interpretation of the results from the perspective of readability in 6.2 drew upon only the outcome or product of the revisions, for only the changes which enhanced or hindered the readability of the end-product were relevant

to the analysis. In the interpretation of the revisions from the perspective of feedback-independence, however, it makes sense to consider the revision process as well. The ineffective and the unnecessary changes must therefore also be retrieved, for although qualitatively they do not affect the product of the revisions, they disclose information which is relevant to the efficiency of the revision process and to the understanding of the kind of feedback the participants needed. While the changes qualified as ineffective yield important information about what the participants tried, but failed, to improve, the changes qualified as unnecessary disclose important information about what the participants changed, but did not have to, probably because they felt unsure about the quality of certain parts of their pre-treatment final drafts. In contrast to this, the changes qualified as consequential and the ones qualified as indeterminate must be excluded from the interpretation of the revisions from the perspective of feedback-independence just as they were excluded from the interpretation of the revisions from the viewpoint of readability. The consequential changes cannot be included because these changes were subordinated to other changes, which makes it impossible to tell whether the consequential changes mean that learning has taken place, or whether the participants would have already been able to introduce the consequential changes before the treatment had they been necessary at that point. The indeterminate changes, in turn, cannot be included in the

interpretation of the post-treatment revisions from the viewpoint of feedback-independence simply because those changes could not be evaluated.

II. Is feedback-independence analysable on the basis of the changes made by the participants alone?

In the interpretation of the revisions from the viewpoint of readability it was only necessary to examine the parts of the pre-treatment final drafts which were revised by the participants themselves. After all, what the participants left unchanged could not have affected readability. To understand the revisions from the perspective of feedback-independence, however, it is important to take into account what the participants left unrevised, for this kind of information is essential to the understanding of the feedback the participants needed. Thus besides having to retrieve the changes coded positive, negative, ineffective and unnecessary, it was also important to retrieve additional information outside the revisions about what the participants left unrevised. In view of this, the "necessary" changes introduced by the native-speaker proofreaders, which, as said in chapter five, are taken to disclose precisely this kind of information, were also accessed.

III. Is feedback-independence analysable only in terms of reading process?

As pointed out in the beginning of this section, besides trying to determine whether or not feedback-independence increased, my second concern was to attempt to find out in what respects it increased. The idea being of course to try and diagnose the kind of feedback needed by the participants. Unlike the interpretation of the revisions from the viewpoint of readability, in which only cross-references with the reading process categories were considered, in the diagnosis of the kind of feedback needed by the participants it is important that the coding of the revisions according to writing product also be accessed. After all, if a given type of writing product change can affect different components of the reading process, and if the same change in reading process can be generated by different changes in writing product, then it is obvious that being independent from feedback presupposes being able to revise not only reading process, but also writing product. If the writer has difficulties in revising certain aspects of his own prose, the feedback he needs may sometimes have more to do with helping him understand how to manage a given component of the reading process and the writing product changes it requires, and sometimes it may have more to do with helping him understand how to manage writing product so that he can address different components of the reading process. For the diagnosis of feedback-

independence to serve a practical pedagogical purpose, it must therefore take both reading process and writing product into account.

Thus to find out whether feedback-independence increased from T3 to T3*, it was necessary to retrieve the coding of all changes by the participants - except for the consequential and indeterminate ones - and of all additional "necessary" changes made by the proofreaders. To find out in what respects feedback-independence increased, it was in turn necessary to examine the above from the dual perspective of reading process and writing product. In numeric terms, this means this part of the study is based on a corpus of 450 observations (385 positive, negative, ineffective and unnecessary changes made by the participants themselves, and 65 necessary changes made by the native speakers). Such observations shall be referred to as feedback-independence observations (FIO).

To explain how increased feedback-independence was measured, I must first of all make it clear my interpretation of increased feedback-independence is based on the following set of assumptions:

ASSUMPTION I: The FIO are observations which signal that learning has been to a greater or lesser extent sufficient (Learning-sufficient observations, i.e., LSO), and observations which signal that learning, even if partial, has been insufficient (Learning-insufficient observations, i.e. LIO). Hence FIO = LSO + LIO.

ASSUMPTION II: The positive changes signal that learning has been to a greater or lesser extent sufficient. Hence LSO = positive changes:

- The changes qualified as positive are FIO which indicate that after the treatment the writer was able to revise with full or partial success parts of text which he was not able to revise on his own at a pre-treatment point. Learning was to a greater or lesser extent sufficient.

ASSUMPTION III: All other FIO are signs of insufficient learning. Hence LIO = negative, ineffective, unnecessary and necessary changes:

- The negative changes indicate that the writer probably needed feedback telling him that his post-treatment revision intuitions were ill-founded and actually made certain parts of text less, rather than more, readable. Learning, even if partial[2], was insufficient.

- The ineffective changes focus on the parts of text the writer may have noticed needed revision, but which he was nevertheless unable to revise successfully when drawing on his own resources. Therefore, the writer probably needed feedback telling him that his attempted revision was unsuccessful and which showed him how to revise what he correctly perceived needed revision. Again, even if partial, learning was insufficient.

- The unnecessary changes indicate that the writer needed feedback in the form of external reassurance from a person able to point out that certain parts of text did not need any revision in the first place. Learning was once more insufficient.

- The necessary changes point towards the need for feedback alerting the writer to the parts of text which should have been revised but were not, either because the writer was unaware those parts needed revision, or because he realized those parts needed revision but for some reason or other avoided revising. In this case too, learning was insufficient.

The operational consequence of the above set of assumptions is that the FIO which indicate that learning was sufficient can be measured and compared with the FIO which indicate that learning was insufficient. This measure, in turn, is the one which seems most logical to use when attempting to

find out whether the post-treatment revisions contain evidence of an increase in feedback-independence. After all, apart from the fact that neither before nor after the treatment the participants were given any cues as to what in their texts might have needed revision, the pre-treatment final drafts represent the best version of text the participants were able to arrive at on their own at a pre-treatment point. In addition to this, as seen in chapter four, the learning which took place during the treatment, whatever it might have been, was maintained in T4, T5 and T6. Evidence that the learning-sufficient observations are significantly more frequent than the learning-insufficient observations will therefore be interpreted as a sign of increased feedback-independence.

In the next three sections I will concentrate first on the results obtained for the overall comparison of learning-sufficient and learning-insufficient observations in order to test H3, i.e., that the post-treatment revisions disclose evidence of increased feedback-independence. After that, cross-references between the learning-sufficient and learning-insufficient observations and the reading process categories will be utilized in order to find out more about feedback-independence from the perspective of reading process, and the consequent reader-oriented feedback the participants might need; cross-references between the learning-sufficient and learning-insufficient observations and the writing product categories will then be used in

order to examine feedback-independence from the perspective of content, lexis, and linguistic and orthographic form, and the consequent writing-product support the participants might need.

6.3.1 Do the post-treatment revisions contain evidence of an increase in feedback-independence?

In order to find out simply whether or not the post-treatment revisions disclose evidence of increased feedback-independence, all that is strictly necessary is to retrieve the 450 FIO relevant to this part of the study, and distribute them according to those which indicate that learning was sufficient and those which indicate that learning was insufficient. Table 6.4 summarizes the results obtained for such a distribution.

Table 6.4: Distribution of feedback-independence observations according to those which signal that learning was sufficient (LSO) and those which signal that learning was insufficient (LIO)

PARTICIPANT	LSO	LIO
Cida	25	19
Dony	40	31
Elisa	32	12
Gustavo	25	14
Henrique	11	18
Silvia	40	18
Thelma	60	25
Wilson	51	29
MEAN	35.5	20.8
SD	15.6	6.9

LSO:LIO overall ratio = 1.7

From the above it is clear that there was much variability with regard to the LSO:LIO ratios for the different participants. Elisa is the participant whose revision disclosed the greatest overall evidence of increased feedback-independence (2.7 LSO for every LIO), and in Henrique's revision, the learning-insufficient observations were actually more frequent than the learning-sufficient ones (0.6 LSO for every LIO).

For the group as a whole, the total number of learning-sufficient observations was almost two times greater than the total number of learning-insufficient observations. When the two were then compared via a matched t-test, it was found that the observations signaling that learning had been sufficient were, at the 0.05 level, significantly more frequent than the observations which pointed towards insufficient learning (t-matched = 3.270). From this it was concluded that the post-treatment revisions hold evidence to a very likely overall increase in feedback-independence.

In the next two sections these feedback-independence observations will be analysed from the perspective of reading process and writing product so as to find out in what respects feedback-independence increased and consequently determine what kind of feedback is still, or no longer, needed.

6.3.2 Feedback-independence and reading process

The first step in the interpretation of feedback-independence from the viewpoint of reading process was to distribute the learning-sufficient and learning-insufficient observations according to system's reading process categories*. The results are summarized in table 6.5.

Table 6.5: Distribution of feedback-independence observations which signal that learning has been sufficient (LSO) and that learning has been insufficient (LIO) according to the reading process categories

CATEGORY/	Acc		App		Coh		Com	
PARTICIP.	LSO	LIO	LSO	LIO	LSO	LIO	LSO	LIO
Cida	4	5	1	10	4	2	0	0
Dony	4	7	10	15	7	4	1	0
Elisa	4	4	11	5	7	1	0	1
Gustavo	1	4	6	7	10	2	0	0
Henrique	2	5	3	3	1	6	1	0
Silvia	4	2	8	7	10	5	3	0
Thelma	3	7	26	14	7	4	4	0
Wilson	6	7	23	12	6	4	1	0
MEAN	3.5	5.1	11	9.1	6.5	3.5	1.3	0.1
SD	1.5	1.8	9.0	4.3	3.0	1.7	1.5	0.4
LSO:LIO overall ratio	0.7		1.2		1.9		10	

Table 6.5 (continued):

CATEGORY/	Inf		IS		Lev	
PARTICIP.	LSO	LIO	LSO	LIO	LSO	LIO
Cida	7	2	5	Ø	4	Ø
Dony	8	3	3	1	6	1
Elisa	7	1	1	Ø	2	Ø
Gustavo	3	Ø	2	Ø	3	1
Henrique	2	2	Ø	1	2	Ø
Silvia	12	3	1	Ø	2	Ø
Thelma	8	Ø	2	Ø	10	Ø
Wilson	8	5	4	1	3	Ø
MEAN	6.9	2	2.3	Ø.4	4	Ø.3
SD	3.1	1.7	1.7	Ø.5	2.8	Ø.5
LSO:LIO overall ratio	3.4		6		16	

As can be seen, the average number of changes which signal that learning was sufficient was greater than the average number of changes which signal that learning was insufficient for all reading process categories except accuracy. The LSO:LIO ratios in turn indicate that for every learning-insufficient observation in levels effect, there were as many as 16 learning-sufficient observations. Commitment also scored high in this respect, and the LSO:LIO ratios for information-structure, informativity and coherence were not too low. For both appropriateness and accuracy, however, there was almost a one to one correspondence between the total number of learning-sufficient and learning-insufficient observations.

The above ratios give some idea of the differences between between categories, but do not take individual differences into account. They therefore do not tell us in which

respects feedback-independence actually increased for the group as a whole. The results obtained for the comparison of the learning-sufficient and learning-insufficient means for each reading process category are shown in table 6.6 below.

Table 6.6: Results for the comparison of reading process learning-sufficient and learning-insufficient means (not significant (*) significant (**) for one-tailed test: 0.05 level)

CATEGORY	T-MATCHED
Acc	-2.303**
App	0.719*
Com	1.938**
Coh	2.201**
IS	3.071**
Lev	3.837**
Inf	4.754**

As can be seen, the results obtained for one-tailed tests at the 0.05 probability level reveal that the learning-sufficient observations were significantly more frequent than the learning-insufficient observations in coherence, commitment, informativity, information-structure and levels effect, but not in accuracy and appropriateness. The results therefore suggest that following the experimental treatment there was an overall increase of feedback-independence with respect to the former. In contrast to this, there does not seem to be sufficient proof of increased feedback-independence in terms of appropriateness and accuracy. The fact that the learning-sufficient observations pertaining to accuracy were actually

significantly **less** frequent than the learning-insufficient ones suggests that the participants are still particularly far from being independent from feedback in this respect. But this does not rule out the possibility that the participants may have nevertheless learned something about accuracy, for results might have been even less favourable had there been no treatment[4].

According to the above diagnosis, future instruction should certainly give more emphasis to helping the participants handle accuracy and appropriateness, and also to helping them gain further feedback-independence in terms of coherence and informativity, for which the learning-insufficient observations were still comparatively frequent. When cross-references with these reading process learning-insufficient observations and the macro-categories for writing product were made, it was found that 87.8% of the feedback on accuracy needed by the participants had to do with linguistic form, and that the remaining 12.2% had to do with orthographic form. Feedback regarding appropriateness should focus mostly on linguistic form (53.4%) and lexis (37%), but should not underrate the importance of orthographic form (8.2%). In order to help the participants become more independent from feedback in terms of coherence, instruction should focus mainly on content (60.7%) and linguistic form (28.6%), and to a lesser extent on lexis (7.1%) and orthographic form (3.6%).

The feedback on informativity, in turn, should pay special attention to content (87.5%).

The amount of emphasis future instruction should assign to the remaining components of the reading process, i.e., levels effect, information-structure and commitment, can probably be reduced since the very small number of learning-insufficient observations pertaining to them indicates that the participants seem to have acquired by now reasonable standards with which to evaluate their own prose in these respects.

6.3.3 Feedback-independence and writing product

In this section feedback-independence will be interpreted from the viewpoint of writing product. The same 450 observations examined from the perspective of reading process in the previous section were therefore sorted out according to the system's taxonomy for describing the revision of writing product.

Since this taxonomy contains categories which are embedded within larger categories, a top-down approach to the analysis was adopted. This means that cross-references between the learning-sufficient and learning-insufficient observations and the macro-level categories at the top of

the taxonomy were accessed first; after that, cross-references between the feedback-independence observations and the sub-categories immediately under linguistic form were retrieved; the details relative to the categories at the lowest level of the hierarchy of the taxonomy, as explained in the beginning of this chapter, were only examined from the viewpoint of feedback-independence if the categories were represented by a minimum of 16 records in the overall distribution.

Thus to begin with, table 6.7 below summarizes the results obtained for the distribution of the learning-sufficient and learning-insufficient observations according to content, lexis, linguistic and orthographic form. Table 6.8, in turn, shows the results obtained for the statistical comparison of means.

Table 6.7: Distribution of learning-sufficient and learning-insufficient observations according to content, lexis, linguistic and orthographic form

CATEGORY/	Content		Lexis		Ling.form		Orth.form	
PARTICIPANT	LSO	LIO	LSO	LIO	LSO	LIO	LSO	LIO
Cida	3	3	2	0	17	14	3	2
Dony	15	6	4	7	18	17	3	1
Elisa	11	1	2	4	15	6	4	1
Gustavo	10	0	2	9	12	5	1	0
Henrique	3	7	1	0	4	7	3	4
Silvia	10	5	7	3	20	7	3	3
Thelma	14	2	7	2	25	18	14	3
Wilson	8	8	9	5	24	16	10	0
MEAN	9.3	4	4.3	3.8	16.9	11.3	5.1	1.8
SD	4.5	2.9	3.0	3.2	6.8	5.5	4.5	1.5
LSO:LIO overall ratio	2.3		1.1		1.5		2.9	

Table 6.8: Results for the comparison of learning-sufficient and learning-insufficient means for content, lexis, linguistic and orthographic form (not significant (*) significant (**) for one-tailed test: 0.05 level)

CATEGORY	T-MATCHED
Content	2.507**
Lexis	0.339*
Ling.form	3.156**
Orth.form	2.091**

From table 6.7 it can be seen that when the 450 feedback-independence observations were distributed according to the four macro-categories of writing product, the differences between the total number of learning-sufficient and learning-insufficient observations were a lot more evenly balanced than when these same observations were distributed according to the reading process categories. While the LSO:LIO ratios for reading process varied from 16 (levels effect) to 0.7 (accuracy), the same ratios for the writing product macro-categories varied only from 2.9 (orthographic form) to 1.1 (lexis). A possible explanation for this could be that the writing product macro-categories are so ample that the finer differences underlying them become flattenned out when grouped together into categories as general as content, lexis, linguistic and orthographic form. It is also possible, however, that the amount of emphasis assigned during the treatment to the different components of the reading process was a lot less evenly balanced than the amount of emphasis conferred to content,

lexis, linguistic and orthographic form. In fact, because the instruction provided was above all discourse-oriented, rather than focus exclusively on content or lexis or linguistic or orthographic form, it touched a bit of everything. In contrast to this, the treatment must have obviously paid much greater attention to the more discoursal components of the reading process (coherence and information-structure, for example) than to its less discoursal components (accuracy and appropriateness)[55].

Still, from table 6.8 it is possible to see that at least one important distinction in writing product has surfaced: the learning-sufficient observations in lexis were not significantly more frequent than the learning-insufficient observations. On the one hand, it is therefore unlikely that there has been an increase in feedback-independence with regard to lexis. On the other hand, however, there appears to have been an increase in feedback-independence in terms of content, linguistic and orthographic form.

The differences between learning-sufficient and learning-insufficient observations became much less even when the macro-category for linguistic form was decomposed into the sub-categories for morphology, lower-level syntax, commutable syntactic forms, sentence complexity and order. Tables 6.9 and 6.10 below summarize the results obtained.

Table 6.9: Distribution of learning-sufficient and learning-insufficient observations according to morphology, lower-level syntax, commutable syntax, sentence complexity and order

CATEG.	MORPH		LLS		CSF		SC		ORD	
PART.	LSO	LIO	LSO	LIO	LSO	LIO	LSO	LIO	LSO	LIO
Cida	3	3	2	4	3	5	2	Ø	7	2
Dony	2	5	4	6	6	3	2	Ø	4	3
Elisa	3	2	7	2	2	1	1	Ø	2	Ø
Gust.	1	1	4	3	1	Ø	2	1	4	Ø
Hen.	1	2	1	2	Ø	1	1	Ø	1	2
Silvia	4	1	6	1	6	2	Ø	1	4	1
Thelma	2	4	1Ø	8	2	5	4	Ø	7	1
Wilson	4	4	11	6	3	3	1	1	5	2
MEAN	2.5	2.8	5.6	4	2.9	2.5	1.6	Ø.4	4.3	1.4
SD	1.2	1.5	3.6	2.4	2.1	1.9	1.2	Ø.5	2.1	1.1
LSO:LIO overall ratio	Ø.9		1.4		1.2		4.3		3.1	

Table 6.10: Results for the comparison of learning-sufficient and learning-insufficient means for morphology, lower-level syntax, commutable syntax, sentence complexity and order (not significant (*) significant (**) for one-tailed test: Ø.Ø5 level)

CATEGORY	T-MATCHED
Morph	-Ø.386*
Lls	1.476*
Csf	Ø.444*
Sc	2.376**
Ord	3.643**

From the above it is clear that the biggest difficulties in the revision of linguistic form had to do with morphology. The post-treatment revisions also did not disclose enough evidence of increased feedback-independence in terms of lower-level and commutable syntax. However, the same does not apply to sentence complexity and order. In terms of sentence complexity, there were as many as 4.3 learning-

sufficient observations for every learning-insufficient one, and in terms of order the total number of learning-sufficient observations was 3.1 times greater than the total number of learning-insufficient ones. In both cases, the fact that the means for learning-sufficient observations were significantly greater than the means for learning-insufficient ones suggests that there was an increase in feedback-independence.

Going further down the hierarchy of the writing product taxonomy, tables 6.11 and 6.12 below summarize the results obtained for feedback-independence and the sub-categories of content which were represented by 16 or more records in the overall distribution.

Table 6.11: Distribution of learning-sufficient and learning-insufficient observations according to the addition or deletion of sentence and descriptive adverbials, and pre and postmodifiers.

CATEGORY/	SADV		DADV		PREMOD		POSTMOD	
PARTICIPANT	LSO	LIO	LSO	LIO	LSO	LIO	LSO	LIO
Cida	0	1	1	0	1	0	1	0
Dony	0	2	4	0	1	1	3	2
Elisa	1	0	1	1	1	0	3	0
Gustavo	2	0	0	0	2	0	1	0
Henrique	0	2	1	2	0	1	0	1
Silvia	0	2	2	1	3	0	3	0
Thelma	3	0	2	0	6	0	1	1
Wilson	2	3	3	2	0	1	1	1
MEAN	1	1.3	1.8	0.8	1.8	0.4	1.6	0.6
SD	1.2	1.2	1.3	0.9	2.0	0.5	1.2	0.7
LSO:LIO overall ratio	0.8		2.3		4.7		2.6	

Table 6.12: Results for the comparison of learning-sufficient and learning-insufficient means for the addition or deletion of sentence and descriptive adverbials, and pre and postmodifiers (not significant (*) significant (**) for one-tailed test: 0.05 level)

CATEGORY	T-MATCHED
SAdv	-0.357*
DAdv	1.871*
Premod	1.672*
Postmod	2.000**

The above figures indicate that the addition or deletion of postmodifiers was the only one of the sub-categories of content with 16 or more records in the overall distribution for which evidence of increased feedback-independence was accepted as being sufficient. It is however interesting to note is that even though the LSO:LIO ratio for the addition or deletion of premodifiers was comparatively the highest, the mean for learning-sufficient observations was not significantly greater than the mean for learning-insufficient observations. The large amount of individual variability with respect to the category explains this apparent contradiction.

Tables 6.13 and 6.14 below summarize the results obtained for cross-references between the learning-sufficient and learning-insufficient observations, and verb and noun phrase lexis, which were the only two sub-categories of lexis with 16 or more records in the overall distribution.

Table 6.13: Distribution of learning-sufficient and learning-insufficient observations according to verb and noun phrase lexis

CATEGORY/ PARTICIPANT	VERB LEX LSO	VERB LEX LIO	NP LEX LSO	NP LEX LIO
Cida	Ø	Ø	2	Ø
Dony	3	4	1	2
Elisa	1	Ø	1	2
Gustavo	Ø	2	Ø	4
Henrique	Ø	Ø	1	Ø
Silvia	2	Ø	4	3
Thelma	3	1	2	1
Wilson	2	2	7	3
MEAN	1.4	1.1	2.3	1.9
SD	1.3	1.5	2.3	1.5
LSO:LIO overall ratio	1.2		1.2	

Table 6.14: Results for the comparison of learning-sufficient and learning-insufficient means for verb and noun phrase lexis (not significant (*) for one-tailed test: Ø.Ø5 level)

CATEGORY	T-MATCHED
Verb lexis	Ø.5Ø9*
NP lexis	Ø.444*

According to the figures in table 6.13, the total number of learning-sufficient observations in verb and noun phrase lexis was almost the same as the total number of learning-insufficient observations. The values in table 6.14 then confirm that the evidence of increased feedback-independence in terms of verb and noun phrase lexis was as unsatisfactory as that of increased feedback-independence in terms of lexis in general.

Only five of the lower-level categories of linguistic form had enough records to justify a more detailed interpretation of their relationship with feedback-independence. The first one, verb-inflection, was a sub-category of morphology; the next two, determiners and prepositions, were sub-catogories of lower-level syntax; and the last two, word and phrase-order, were sub-categories of order in general. The results derived from cross-references between these lower-level categories of linguistic form and the learning-sufficient and insufficient observations are summarized in tables 6.15 and 6.16 below.

Table 6.15: Distribution of learning-sufficient and learning-insufficient observations according to verb inflection, determiners, prepositions, word and phrase order

CATEGORY/	MORPH.vi		LLS.det		LLS.prep		ORD.word		ORD.phr	
PARTICIP.	LSO	LIO	LSO	LIO	LSO	LIO	LSO	LIO	LSO	LIO
Cida	1	2	1	4	Ø	Ø	3	2	4	Ø
Dony	2	2	2	2	2	3	Ø	2	4	Ø
Elisa	2	1	5	1	1	1	1	1	Ø	Ø
Gustavo	Ø	1	2	Ø	1	3	2	Ø	1	Ø
Henrique	1	1	Ø	Ø	1	1	1	Ø	Ø	2
Silvia	Ø	Ø	2	Ø	3	2	1	1	1	Ø
Thelma	Ø	1	5	5	2	2	3	1	1	Ø
Wilson	2	Ø	3	4	7	2	2	2	3	Ø
MEAN	1	1	2.5	2	2.1	1.8	1.6	1.1	1.8	Ø.3
SD	Ø.9	Ø.8	1.8	2.1	2.2	1.Ø	1.1	Ø.8	1.7	Ø.7
LSO:LIO overall ratio	1		1.3		1.2		1.4		7	

Table 6.16: Results for the comparison of learning-sufficient and learning-insufficient means for verb inflection, determiners, prepositions, word and phrase order (not significant (*) significant (**) for one-tailed test: 0.05 level)

CATEGORY	T-MATCHED
MORPH.vi	0.000*
LLS.det	0.661*
LLS.prep	0.513*
ORD.word	1.080*
ORD.phr	2.049**

The above figures indicate that there was an exact one-to-one correspondence for the total number of learning-sufficient and insufficient observations in verb-inflection. According to the criteria adopted in the present interpretation of the results, this means that there is not enough evidence of an increase in feedback-independence insofar as verb-inflection is concerned. The learning-sufficient means for determiners, prepositions and word order were also not very different from the corresponding learning-insufficient means, which again implies that the data holds no evidence to an increase in feedback-independence in those respects. For phrase order, however, there were as many as seven learning-sufficient observations for every learning-insufficient one, and the statistical comparison of means led me to the conclusion that there was enough evidence of an increase in feedback-independence.

The last lower-level category of writing product which was frequent enough to be examined from the viewpoint of feedback-independence was punctuation. The results derived from cross-references between the category and the LSO and LIO are shown in tables 6.17 and 6.18 below.

Table 6.17: Distribution of learning-sufficient and learning-insufficient observations according to punctuation

PARTICIPANT	PUNCTUATION	
	LSO	LIO
Cida	1	1
Dony	2	Ø
Elisa	4	Ø
Gustavo	1	Ø
Henrique	3	3
Silvia	1	2
Thelma	8	2
Wilson	9	Ø
MEAN	3.6	1
SD	3.2	1.2
LSO:LIO overall ratio	3.6	

Table 6.18: Results for the comparison of learning-sufficient and learning-insufficient means for punctuation (significant (**) for one-tailed test: Ø.Ø5 level)

CATEGORY	T-MATCHED
punct	2.145**

The above indicates not only that for every learning-insufficient observation in punctuation there were as many as 3.6 learning-sufficient observations, but also that the latter were significantly more frequent than the former. There is therefore evidence to suggest that there was an increase in feedback-independence in terms of punctuation.

To conclude this section, according to the above interpretation of feedback-independence from the viewpoint of writing product, it appears that the participants would benefit from further instruction which gave special emphasis to lexis, morphology (especially verb inflection), lower-level syntax (especially determiners and prepositions), commutable syntax, adverbials, premodifiers and word order.

When cross-references between the learning-insufficient observations pertaining to lexis and the reading process categories were made, it was found that 90% of these observations had to do with appropriateness. It therefore seems that feedback on lexis would greatly help the participants manage appropriateness on their own. The learning-insufficient observations pertaining to morphology and lower-level syntax tied in with mostly accuracy and appropriateness. Feedback on morphology and lower-level syntax would therefore probably help the participants produce more accurate and more appropriate texts. The learning-insufficient observations pertaining to commutable syntax affected mainly appropriateness and coherence, which means feedback on commutable syntax would probably have a positive effect on these two components of the reading process. From cross-references between reading process and the learning-insufficient observations pertaining to adverbials in general, it appears that teaching the participants more about their use will enhance coherence

and informativity. Finally, the LIO pertaining to premodifiers had a one-to one correspondence with informativity, and the word-order LIO had to do with appropriateness, accuracy, information-structure and levels effect. Teaching the participants more about premodifiers and word-order would therefore probably help them improve the above components of the reading process.

The participants would obviously also benefit from instruction which helped them become even more independent from feedback regarding punctuation, phrase-order, sentence-complexity and postmodifiers, although the few learning-insufficient observations pertaining to these categories indicate that by now the participants seem to have acquired reasonable standards with which to evaluate their own prose in these respects, and that the amount of emphasis assigned to these parts of writing product can consequently be reduced.

6.4 Conclusions

The first conclusion about the interpretation of the post-treatment revisions from the perspective of readability and feedback-independence is that after instruction had ceased the participants seem to have been generally able to improve the readability of their pre-treatment final

drafts, and that feedback-independence appears to have generally increased.

Notwithstanding these coinciding results, the second conclusion reached is that it is misleading to assume that evidence of improvement in the readability of the end-product of revision can be equated with evidence of increased feedback-independence. A learner's self-sufficient ability to improve the readability of his text may at times distract one from seeing that he has not in fact gained feedback-independence.

With regard to the present data, it is notable that although the post-treatment revisions were found to be both more accurate and more appropriate than the corresponding pre-treatment final drafts - and hence the changes in accuracy and appropriateness must have enhanced more than hindered readability -, evidence of increased feedback-independence in these respects was inconclusive, for there were comparatively too many feedback-independence observations in accuracy and appropriateness indicating that learning had been insufficient. In addition to this, even though there was evidence of improved readability **and** increased feedback-independence for the remaining five components of the reading process, the overall positive:negative ratios for each category were always much higher than the corresponding learning-sufficient:learning-insufficient ratios. In other words, the changes in reading

process which enhanced readability outnumbered the corresponding changes which hindered readability to a much greater extent than the learning-sufficient observations in reading process outnumbered the learning-insufficient ones. The reason why this was so is that certain changes which did not in fact hinder the readability of the end-product were nevertheless taken to be signs of insufficient learning.

Thus even if writers are able to improve the readability of their texts on their own, there may still be much more for them to learn before feedback becomes unnecessary. It should not be forgotten, however, that feedback-independence may increase in some respects without the overall result being improved readability if what increases in terms of feedback-independence contributes only very slightly towards improved readability. Traditional product-oriented instruction, for example, may result in increased feedback-independence in accuracy, which is unlikely to in itself correlate with a general improvement in readability. Depending on the kind of instruction provided, the following four combinations of changes in readability and feedback-independence may result:

1. + readability + feedback-independence
2. + readability - feedback-independence
3. - readability + feedback-independence*
4. - readability - feedback-independence

(*in accuracy or other factors unlikely to contribute much towards improved readability)

If the goal of writing instruction is to help writers rely less and less on cues from the writing teacher, then it follows that the analysis of feedback-independence is more basic to one's understanding of writers' needs than the analysis of the readability of the texts they produce. If feedback-independence increases in all respects, or at least in terms of what is important to readability, then it is natural that readability should also improve.

The third and last conclusion reached is that for the analysis of feedback-independence to serve a practical pedagogical purpose, it is vital that it be understood from the dual perspective of reading process and writing product. Although certain correspondences between the two are not unlikely, when the learning-insufficient observations pertaining to the different reading process categories were sorted out according to the macro-categories for writing product, and when the learning-insufficient observations pertaining to the different writing product categories were conversely sorted out according to reading process, different combinations of the two occurred. Figures 6.20 and 6.21 below summarize the results obtained.

Figure 6.20: Distribution of learning-insufficient observations in reading process according to writing product

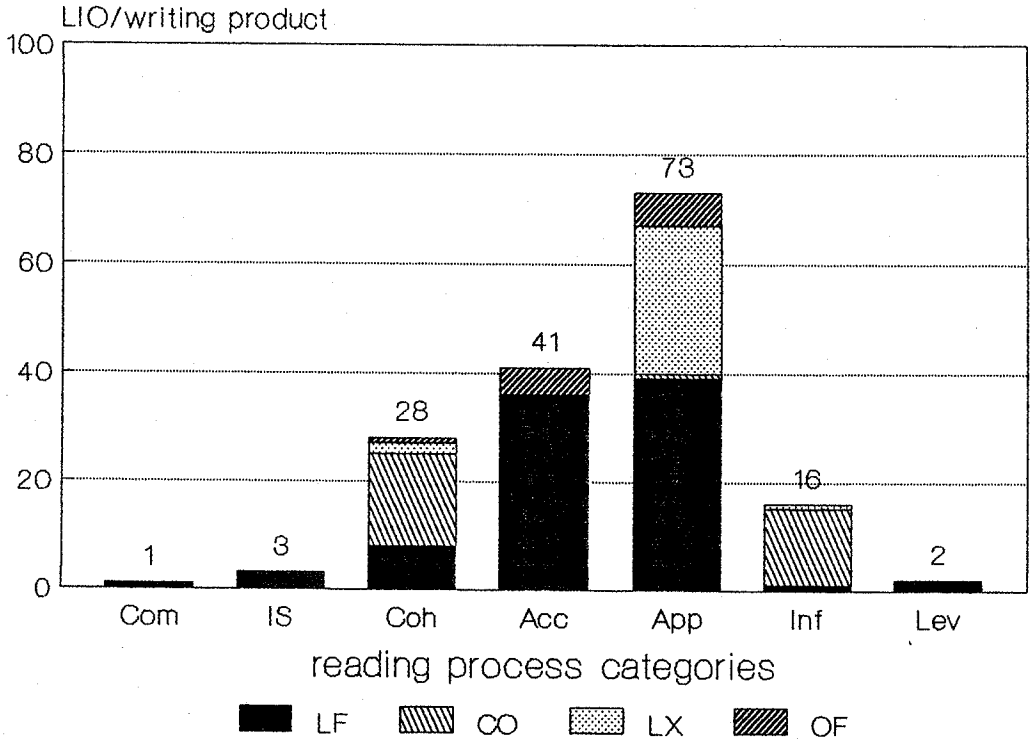
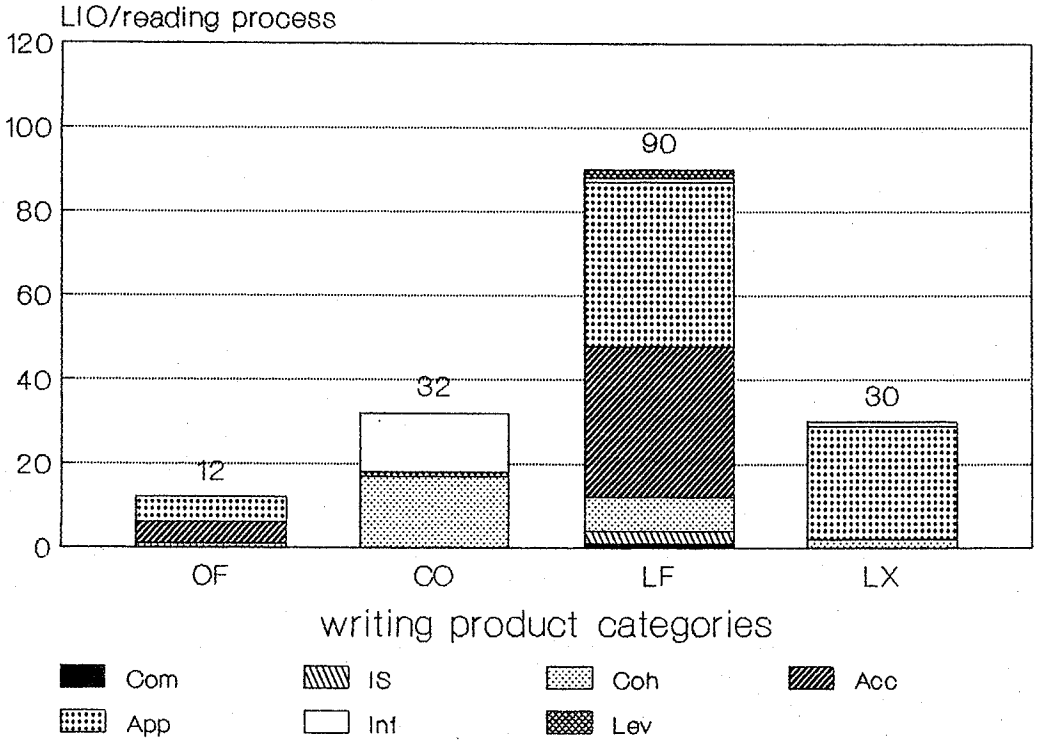


Figure 6.21: Distribution of learning insufficient observations in writing product according to reading process



From the above it is clear that in order to help the participants gain feedback-independence in terms of different components of the reading process, future instruction should provide them with different types and amounts of writing product feedback. For the participants to gain further feedback-independence in terms of coherence and informativity, for example, considerable emphasis may have to be placed on content, which is however likely to be of little or no consequence to an increase in feedback-independence with respect to the remaining components of the reading process. Conversely, in order to help the participants gain feedback-independence with regard to different aspects of the writing product, future instruction should provide them with different types and amounts of reading process feedback. For example, before the participants can do without feedback on lexis, they will have to learn a lot more about appropriateness and comparatively very little else about the remaining components of the reading process. Keeping reading process and writing product apart from one another is therefore extremely important when it comes to choosing the right focus for future writing instruction, especially if the instructional period is short and decisions have to be made as to what needs be addressed most urgently.

Notes to chapter six

1. The changes coded according to the discourse category `other', which added up to only Ø.4% of the total number of changes identified in the transcriptions, were not taken into account inasmuch as these changes do not affect readability in any perceptible or identifiable way.

2. Negative changes should not be categorically interpreted as signs of irreversible backsliding. On the contrary, in the context of revision following a short instructional period they seem to be typical indicators of what may occur in Stage Two of Kellerman's (1983;1987) U-shaped behaviour language acquisition thesis, whereby a form which was error-free in Stage One becomes deviant in Stage Two and returns to the norm in Stage Three. In the words of Kellerman (1987:215), "... the appearance of deviant forms in Stage Two should not be seen as evidence of attrition in linguistic competence, but as a cognitive advance..." Thus if Stage Two is seen as part of the path towards second language development, then it seems perfectly plausible that some of the negative changes in the revisions be signs insufficient, albeit partial, learning. Frawley and Tolf (1985:41) have similar views: "errors may not be errors as such, but may well represent a speaker's [or, more broadly, a learner's] attempt to gain control of a task".

3. The reading process changes coded "other" were again excluded from the analysis.

4. As explained in the beginning of chapter five, the changes made from the pre-treatment final drafts to the post-treatment revisions could not be compared with changes made in the absence of the experimental treatment in equal terms, which makes it impossible to determine whether there would have been even more LIO pertaining to accuracy had there been no treatment.

5. It is not my yet intent to examine treatment effect. This will be left to chapter seven. At this point it seems nevertheless appropriate to anticipate that the changes with an explicit connection with the treatment tended to tie in with the more discoursal reading process categories.