

Efficacy of Written Corrective Feedback on University Students' Writing

Sibanda Jabulani

*Rhodes University, P. O. Box 94, Grahamstown, South Africa
E-mail: jabusbnd@gmail.com*

KEYWORDS Assessment. Errors. Essay. Language of Attribution. L2 Writing

ABSTRACT The study investigated the efficacy of written corrective feedback (WCF) on five language of attribution error categories in three essays by 12 university students'. It utilised an experimental design where three students (above average, average, and below average performers) were given direct feedback; three, indirect feedback; three, metalinguistic feedback; and three (control group), no feedback on their errors. There was greater reduction in error incidences among the experimental groups than the control group, evidence of the efficacy of WCF on students' writing. Feedback effect was most manifest among the average performers and least among below average performers. Although decline in error incidences was not most manifest in the above average performers, it was evident that academic ability aided comprehension and actualisation of feedback. Indirect feedback registered the greatest effect and metalinguistic feedback the least effect, testimony to a correlation between type of feedback and resultant decline in error incidences.

INTRODUCTION

Sound academic writing evinces effectual scholarship, and enhancing students' academic writing competencies should be a priority for university lecturers and students alike. In fields like the humanities, there is a preponderance in the use of the essay to assess learners' content knowledge and its articulation. A cogent exposition of issues is considered a hallmark of scholarship and robust erudition should be manifest in both the content and its presentation. In the appraisal of students' essays, university lecturers often provide corrective written corrective feedback (WCF) to both the disciplinary knowledge and the linguistic, rhetorical and conventional aspects of writing. The latter are seen by Gray (2011) as setting academic writing apart from any other form of writing. This study investigates the efficacy of WCF on university students' writing.

Truscott's (1996) call for the abandonment of feedback on the basis of its supposed ineffectiveness and potential harm has been credited with providing the impetus to the raging debate on whether corrective feedback matters (Hyun, 2013; Wang and Jiang 2015). While lecturers view feedback provision as time consuming, students complain about the insufficiency and the delay in feedback provision (Goos et al. 2011). The large number of students the South African public universities have had to deal with, courtesy of the widening access to university

education, has most likely seen a reduction in the amount of feedback provision and an increase in the turnaround time for its provision. Within such a scenario, the question; 'Does written corrective feedback really matter?' is invaluable. Lack of consensus on the response to the question explains why feedback has in many instances been underutilised, withheld, underused or ignored.

Although this study is set within the South African Higher Education (HE) context, that the massification of university education, which has resulted in a linguistically and academically diverse cohort of university students, is not unique to South Africa, and that the linguistic conventions governing academic writing transcend national boundaries, accords the study global application. Wang and Jiang (2015) observe that the large body of research that has been conducted on the role of WCF has largely been in the area of second language acquisition. The present study takes the investigation onto university students' writing in a context where the majority of them were even first language users.

Theoretical Framework

The present study assumes a dialectical lecturer-student relationship where both parties actively produce, consume and transform each other's writing. What students write initiates the feedback the lecturer provides, which should in turn determine what the student will write next.

This dialectic explains the situation of this study within the academic literacies framework where language is socially situated and ideologically inscribed (Lillis 2003). The nature of the dialectic further explains how the study also draws from a skills and socialisation model, where there is need for explicit attention to discrete linguistic elements (skills model) and where students are implicitly inducted into established discourse and academic practices (socialisation model). The study therefore, draws on the whole range of approaches to student writing in HE as theorised by Lea and Street (2006) namely; the skills, socialisation, and academic literacies models. A theorisation and conceptualisation of feedback puts the study into perspective.

Theorising Feedback

Feedback in this context relates to recommendations made on a written product on the basis of its critical assessment with the express purpose of regulating and improving performance in the subsequent written products. Direct feedback on one hand, involves the provision of the correct linguistic form or structure close to the linguistic error (Bitchener et al. 2005) in the form of deletion of a superfluous form or structure, "...insertion of a missing word/phrase/morpheme, or the provision of the correct form or structure" (Bitchener 2008: 104). Indirect feedback on the other hand, merely signals the writer's commission of an error through "...underlining or circling the error; recording in the margin the number of errors in a given line; or using a code to show where the error has occurred and what type of error it is" (Bitchener 2008: 104). Direct corrective feedback's strength is in its unambiguity about the nature of both the linguistic deviation and the accurate form whereas that of indirect corrective feedback is in fostering hypothesis testing which possibly "...induce(s) deeper internal processing and promote(s) the internalization of correct forms and structures" (Bitchener 2008: 104).

Ellis (2009) adds another category of corrective feedback which, to me, straddles the direct and indirect feedback categories. This is the metalinguistic feedback where the assessor appends a numbered code in the margin which is followed-up on later at the end of the piece of writing with a brief description. The code (indirect feedback) could excite hypothesis testing which is later confirmed or refuted by the description at the end (direct feedback).

Storch (2010) joins several voices preceding him (like Bitchener 2008; Sheen 2007) in advising confinement of the assessment of the efficacy of feedback to a narrow range of treatable errors amenable to feedback. In this study, the efficacy of feedback was investigated on learners' employment of the language of attribution in their essay writing, an aspect of academic essay writing the present researcher found problematic in his experience assessing students' academic essays.

The Error Category on Which Feedback was Applied

The language of attribution is the language the essay writer employs to signal to the reader that what follows is an enlisting of an external source's (author's) ideas to bolster one's own. These words, usually in the form of attributive tags or attributive verbs, should be so manipulated to ensure that they are consistent with the spirit and intent of the authorial voice. The essay should expertly demonstrate the interaction between and among the author, writer and textual material. Because attributive verbs have nuanced meanings, their ad hoc and imprecise use misrepresents the author's tone and stance. The categories related to the use of the language of attribution for which feedback was given and its efficacy assessed were:

- ♦ Attributive verb agreeing in person and number with authors.
- ♦ Tense consistence in attributive verbs.
- ♦ Consistence between words used to introduce a quote and the spirit and intent of the citation.
- ♦ Proper punctuation with the use of attributive language.
- ♦ The lexico-grammatical company each attributive verb keeps.
- ♦ Signal phrases fitting the syntax and grammar of author and writer's own words.

Apart from the preponderance of errors related to this linguistic category in students' writing, the choice of this error category was also actuated by the treatable nature of the errors, since use of language of attribution is largely rule governed.

Problem Statement

The study investigated the efficaciousness of targeted written corrective feedback on post-graduate certificate in Education (PGCE) univer-

sity students' writing over three pieces of essay writing. Focus was on the extent to which PGCE students of varying academic competence levels responded to direct or indirect WCF on a targeted aspect of academic essay writing; the language of attribution. The rationale was to determine the potential of feedback to feed forward to future academic writing products.

Research Questions

The study was guided by one major research question

1. Does WCF improve PGCE students' use of the language of attribution in academic essay writing, and two sub-questions namely:
 - a. Is there differential effect on the gains registered from different WCF options?
 - b. To what extent is the effect of WCF linked to differential student academic performance?

METHODOLOGY

Sample

The 32 (2014 cohort) University PGCE Foundation Phase students whom the researcher taught were the student population on which the study was conducted. Of these, only two were male. The linguistic demographic was 19 English speaking, 5 isiXhosa speaking, 4 Afrikaans speaking, and 3 other South Africa official language speaking, and 1 Shona speaking. Purposive sampling was employed to select 12 students for the study. The number 12 was necessitated by the need to have a student in the below average, average, and above average performance group for each of the feedback types namely; direct feedback, indirect feedback, metalinguistic feedback, and no feedback. Students' academic performance was determined by their scores in the first of the three essay assignments the researcher gave them. The best 10 performers were considered above average, the lowest 10, the below average performer and the remaining 12, average performers. Four students were selected from each of the three groups for the four feedback categories.

Because the target for the feedback was the language of attribution, it was imperative that the essays within the three performance categories which exhibited the greatest attributive lan-

guage errors selected. This would enable the tracing of the extent to which there was a reduction in the attributive language errors subsequent to the feedback. The initial appraisal of the first essay had provided other forms of WCF on a wide range of errors and not only language of attribution errors. The researcher then combed through the 32 essays for the preponderance of language of attribution error categories and selected 4 scripts per performance category. Names of students of the selected scripts in each performance category were placed in boxes and randomly picked and assigned to boxes reflecting the four feedback types to be provided. The pseudonyms given are alphabetic with student A-C being those given direct feedback; D-F, indirect feedback; G-I, metalinguistic feedback; and J-L, no feedback. The designations 1, 2 and 3 denote above average, average and below average performance respectively. Table 1 shows the resultant sample and the feedback provided.

Table 1: Participants and the feedback given to them

<i>Participant and academic performance above average (1), average (2), below average (3)</i>	<i>Feedback given</i>	<i>Home language</i>
Anna (1)	Direct	isiXhosa
Betty (2)	Direct	English
Catherine (3)	Direct	isiXhosa
Debra (1)	Indirect	English
Eric(2) the only male	Indirect	Afrikaans
Flora (3)	Indirect	English
Gladys (1)	Metalinguistic	Afrikaans
Helen (2)	Metalinguistic	English
Ivy (3)	Metalinguistic	English
Jane (1)	No feedback	isiZulu
Kate (2)	No feedback	English
Lucy (3)	No feedback	isiXhosa

Research Method and Procedure

Selected learners' essay scripts, which had been given feedback similar to that of the rest of the class were isolated for specific feedback on attributive language according to the four categories of feedback namely; in-text direct feedback, in-text indirect highlighted, coded or marginal feedback, metalinguistic feedback (which Bankier (2012) identifies as post-text feedback where the error incidence is identified in-text but its nature is noted and clarified at the end of the text), as well as feedback avoidance. The study

utilised the experimental design involving pre-test and post-tests. Documentation of errors in the use of the language of attribution in the first essay constituted the pre-test and the two subsequent essays represented the post-tests. Language of attribution errors for each student's three essays were expressed as a percentage of the total uses of the particular structure or form. If 12 attributive verbs were used to herald the advent of an external voice and 4 of these were inconsistent with the tone, spirit or intent of the authorial voice, the error incident would be $4/12 \times 100 = 33.3\%$. This would allow comparability on the particular error category across the three essays to determine the efficacy of feedback given. According to Cohen et al. (2007) 'reliability as equivalence' requires that equivalent test forms be employed.

The control group (students not given WCF on their language of attribution errors) enhanced the validity of the experimental procedure by allowing comparability of their percentage errors with those of the treatment groups (Bitchen 2008). That the essays were written under normal assessment conditions and to fulfil coursework requirements meant that they represented the best of students' efforts which made them valid measures of their proficiency levels. Because these were assignments where students had time to self-correct, the assumption was that all deviant uses of attributive language were errors and not mistakes. The researcher ensured that when students got back their marked scripts they had another essay assigned so that they would consider the WCF in the light of the new essay task. The need for immediate post-testing was actuated by the need to ensure that whatever reduction in the error incidence of particular forms or structures were attributable to the feedback and not to some extraneous variables.

At the university, PGCE students did not take an academic literacy course where structures like attributive tags could be taught and chances that their other lecturers focused on nuanced errors of attributive verbs were slim. This increased confidence that qualitative improvements students registered in the use of language of attribution could be attributed to the essay feedback rendered. An attempt was made to ensure comparability in the cognitive demand of the three assignment tasks, their length, the extent to which they needed extensive enlisting of scholars' and authors' ideas all of which added

to the validity with which the measurement of improvement could be made.

Ethical Considerations

Because the study coincided with the normal assessment of students' work by the researcher-cum-lecturer, access to students' essays was guaranteed. The use of pseudonyms was meant to protect their anonymity and the study did not pose any harm to the students. The potential negative effects of the consignment of some learners to a control group were mitigated by providing them detailed feedback on the other aspects of their writing apart from language of attribution aspects. Feedback denial for the control group was only on the study's targeted linguistic structure but not on the content and other linguistic forms and structures.

OBSERVATIONS AND DISCUSSION

Presentation of data was largely tabular and in accordance with error categories identified. The category was identified and the incidence of error noted across the 3 essays and percentage decreases (-a %) or increases (+a %) noted for each of the students. In some cases, exemplifications of the manifestations of the errors were made before an initial analysis of the data for the individual category.

The small number of students (9 for the experimental group and 3 for the control group) would preclude and compromise accuracy of statistical significance as not many statistical assumptions could be tested. This would compromise the credibility of the p-value or even render the t-test inappropriate. The small sample also brings fragility to the results where a minor change would alter or reverse the statistical significance. For these reasons, simple percentage error reduction (for categories 1-4) and simple error counts (category 5) were used in this study for comparisons.

The first category, represented in Table 2, considered the agreement between the attributive word (normally a verb) with the spirit, intent and tone of the quote or paraphrase.

Error Category 1

The first error category considered was lack of consistency between an attributive word and the tone and spirit of the citation. The efficacy of feedback was generally manifest in all three

Table 2: Consonance between attributive word and the citation's spirit, intent and tone

<i>Student and academic performance</i>	<i>No. and (%) of Essay 1 errors</i>	<i>Feed-back rendered</i>	<i>No. and (%) of Essay 2 E errors</i>	<i>% Difference</i>	<i>No. and (%) of Essay 3 errors</i>	<i>% Difference</i>
Ann 1	3/15 (20)	Direct	2/13 (15)	5	0/12 (0)	15
Betty 2	5/12 (41.7)	Direct	5/14 (35.7)	6	3/10 (30)	5.7
Cleo 3	5/10 (50)	Direct	6/13 (46)	4	5/11 (45)	1
Average	13/37 (35)		13/40 (32.5)	2.5	8/33 (24)	7.5
Debra 1	4/14 (28.6)	Indirect	2/12 (16.7)	11	0/10 (0)	16.7
Eric 2	6/13 (46)	Indirect	3/12 (25)	21	1/11 (9)	16
Flora 3	7/12 (58)	Indirect	5/11 (45)	13	3/13 (23)	22
Average	17/39 (43.6)		10/35 (28.6)	15	4/34 (11.8)	16.8
Gail 1	2/11 (18)	Metalinguistic	2/14 (14)	4	1/12 (8)	6
Helen 2	4/12 (33)	Metalinguistic	4/15 (26.7)	6.3	2/12 (16.7)	10
Ivy 3	5/10 (50)	Metalinguistic	6/12 (50)	0	7/11 (63.6)	-13.6
Average	11/33 (33)		12/41 (29)	4	10/35 (28.6)	0.4
Jane 1	4/16 (25)	No feedback	3/12 (25)	0	5/13 (38)	-13
Kate 2	6/13 (46)	No feedback	5/11 (45)	1	6/12 (50)	-5
Lucy 3	5/9 (55.6)	No feedback	5/10 (50)	5.6	6/13 (46)	4
Average	15/38 (39)	Control group	13/33 (39)	0	17/38 (44.8)	-5.8
Grand Total	41/109 (37.6)	For the 3 feed-back groups	35/116 (30)	7.6	22/102 (21.6)	8.4

types of feedback but most evident in indirect feedback, followed by direct feedback and lastly metalinguistic feedback. Storch (2010) asserts that indirect feedback is only efficacious where the learner is conversant with the problem structure to allow for self-correction. In the light of that observation, there would seem to be a disjuncture between the fact that indirect feedback spurred the greatest improvement and the study's assumption that the students had not had much exposure to the nuanced use of the language of attribution. The enhanced correct usage of the attributive verbs could be a result of avoidance of identified problematic attributive verbs and greater use of neutral verbs. The noticing function that indirect feedback afforded could have led to student-self investigation of the nuanced use of the problem verbs and the qualitative benefits of discovered knowledge over imparted knowledge can hardly be contested.

On average, in both direct and indirect feedback, there were marginally more gains between the second and third essays than between the first and second essays. This could be indicative of the efficacy of feedback being more marked when sustained over time and not taken as a once-off instalment. Overall, absence of WCF initially led to no change in the error incidence between the first two essays but later resulted in change for the worse when more errors

were committed in the third essay than in the first and second essays. With the exception of Flora (a below average performer who registered significant improvement across the two essays), the efficacy of feedback was more pronounced among the average and above average students than among the below average students. Feedback seemed to profit the more able students possibly because their academic aptitude predisposed them to better and fuller understanding of feedback, and to self-correction. A computation of the feedback errors and gains made across the pieces of writing is summarised in Table 3.

The figures exclude the control group. The average group had better gains overall than the below and above average student groups. That the above average group exhibited less gains than the average group meant the efficacy of feedback, while correlating with performance (on account of the below average students registering the least gains), was not only dependant on students' general academic abilities. The above average students possibly got excited with the high scores and did not pay close attention to the feedback provided.

Error Category 2

The next error category related to the incompatibility between attributive words and their

Table 3: Summary of error incidence on the match between attributive word and citation's spirit, intent and tone

<i>Variable considered</i>		<i>Essay 1</i>	<i>Essay 2</i>	<i>Essay 1 and 2 Difference</i>	<i>Essay 3</i>	<i>Essay 2 and 3 Difference</i>
<i>Student Performance</i>	Above average	9/40 (22.5%)	6/39 (15%)	7.5%	1/34 (2.9%)	12.1%
	Average	15/37 (40.5%)	12/41 (29%)	11.5%	6/33 (18%)	11%
	Below average	17/32 (53%)	7/36 (47%)	6%	15/35 (42.9%)	4.1%
<i>Feedback Provided</i>	Direct	13/37 (35%)	13/40 (32.5%)	2.5%	8/33 (24%)	7.5%
	Indirect	17/39 (43.6%)	10/35 (28.6%)	15%	4/34 (11.8%)	16.8%
	Metalinguistic	11/33 (33%)	12/41 (29%)	4%	10/35 (28.6%)	0.4%

accompanying lexico-grammatical terms and the unwarranted addition or omission of the latter. The term 'that' was preponderantly used to accompany the bulk of the attributive verbs that were employed. In some cases, this introduced an error of addition where 'that' was made to accompany a word which did not need any accompaniment as in 'explores that', 'highlights that', 'advocates for'. In other instances, the attributive verb used warranted accompaniment but the wrong accompaniment was used as in 'supports that', 'contradicts that', or the notorious 'discusses about'. Table 4 shows the extent of the errors per student across the 3 essays. The percentage error occurrence was a measure of the number of errors as a percentage of the total number of attributive words used in an essay.

For this category of errors, metalinguistic feedback registered the greatest gains and direct feedback had a marginal edge over indirect feedback overall. This was a reversal of the order in which the three forms of WCF impacted subsequent pieces of writing. The implication then was that the impact of a type of feedback could also be contingent upon the nature of the error being corrected. Category 1 error was largely semantic whereas category 2 error was largely grammatical and the distinction potentially accounted for the reversal in the impact of the three forms of WCF on students' writing. That the least percentage error reduction was realised by the control group speaks to the general efficacy of feedback in error diminution in students' writing.

Table 4: Errors in the lexico-grammatical terms accompanying attributive words

<i>Student and academic performance</i>	<i>No. and (%) of Essay 1 errors</i>	<i>Feed-back rendered</i>	<i>No. and (%) of Essay 2 errors</i>	<i>% Difference</i>	<i>No. and (%) of Essay 3 errors</i>	<i>% Difference</i>
Ann 1	2/15 (13)	Direct	2/14 (14)	-1	1/12 (8)	6
Betty 2	3/11 (27)	Direct	1/11 (9)	18	1/14 (7)	2
Cleo 3	5/12 (41.7)	Direct	4/12 (33)	8.7	4/13 (30.8)	2.2
<i>Average</i>	<i>10/38 (26)</i>		<i>7/38 (18)</i>	<i>8</i>	<i>6/39 (15)</i>	<i>3</i>
Debra 1	2/11 (18)	Indirect	1/13 (7.7)	10.3	2/15 (13)	-5.3
Eric 2	3/14 (21)	Indirect	1/14 (7)	14	0/13 (0)	7
Flora 3	4/11 (36)	Indirect	3/12 (25)	11	3/12 (25)	0
<i>Average</i>	<i>8/36 (22)</i>		<i>5/39 (12.9)</i>	<i>9.1</i>	<i>5/40 (12.5)</i>	<i>0.4</i>
Gail 1	1/13 (7.7)	Metalinguistic	2/15 (13)	-5.3	0/12 (0)	13
Helen 2	4/12 (33)	Metalinguistic	2/15 (13)	20	1/13 (7.7)	5.3
Ivy 3	5/11 (45)	Metalinguistic	3/11 (27)	18	3/10 (30)	-3
<i>Average</i>	<i>10/36 (27.8)</i>		<i>7/41 (17)</i>	<i>10.8</i>	<i>4/35 (11)</i>	<i>6</i>
Jane 1	1/15 (6.7)	No feedback	1/14 (7)	-0.3	2/12 (16.7)	-9.7
Kate 2	3/13 (23)	No feedback	3/14 (21)	2	1/13 (7.7)	13.3
Lucy 3	5/13 (38)	No feedback	4/14 (28.6)	9.4	3/12 (25)	2.4
<i>Average</i>	<i>9/41 (22)</i>	<i>Control group</i>	<i>8/42 (19)</i>	<i>3</i>	<i>6/37 (16)</i>	<i>3</i>
Grand Total	28/110 (25)	For the 3 feedback groups	19/118 (16)	9	15/114 (13)	3

The fact that the percentage error decrease was more marked between the first and second essays than between the second and third essays could be construed as indicative of some errors being more prone to WCF than others, thereby registering greater gains from initial feedback than others. In that sense, the efficacy of WCF is dependent on the nature of the error category.

The pattern manifest in the first error category where the greatest gains were made among the average performers and the least gains were realised among the below average performers is repeated in this category as Table 5 shows. Average students seemed to profit more from WCF than the other two performance groups.

Error Category 3

The third error category was the lack of agreement between the attributive verb and the number of authors. The most prevalent of these er-

rors included reference to multiple authors, particularly where one was named and et al. used to denote the others. In these cases, the attributive verb in the simple present almost always carried an 's'. Another manifestation of subject-verb agreement errors was the citation of a single author who is cited by multiple authors or vice versa. Students, in this case, had challenges navigating the co-occurrence restrictions.

In Table 6, the percentage error occurrence is expressed over the number of attributive verbs used. The impact of initial and subsequent feedback was ambivalent when it came to the types of WCF proffered. While indirect feedback had the greatest impact in the initial feedback (between Essay 1 and 2) direct feedback had the greatest impact in the subsequent feedback (between Essay 3 and 4), and metalinguistic feedback had the second greatest impact in both the initial and subsequent feedback. That overall, the control group manifested the least gains and

Table 5: Summary of error incidence on lexico-grammatical terms accompanying attributive words

<i>Variable considered</i>		<i>Essay 1</i>	<i>Essay 2</i>	<i>Essay 1 and 2 Difference</i>	<i>Essay 3</i>	<i>Essay 2 and 3 Difference</i>
<i>Student Performance</i>	Above average	5/39 (12.8%)	5/42 (11.9%)	0.9%	3/39 (7.7%)	4.2%
	Average	10/37 (27%)	4/40 (10%)	17%	2/40 (5%)	12%
	Below average	14/34 (29%)	10/35 (28.6%)	0.4%	10/35 (28.6%)	0%
<i>Feedback Provided</i>	Direct	10/38 (26%)	7/38 (18%)	8%	6/39 (15%)	3%
	Indirect	8/36 (22%)	5/39 (12.9%)	9.1%	5/40 (12.5%)	0.4%
	Metalinguistic	10/36 (27.8%)	7/41 (17%)	10.8%	4/35 (11%)	6%

Table 6: Lack of agreement between the attributive verb and the number of authors

<i>Student and academic performance</i>	<i>No. and (%) of Essay 1 errors</i>	<i>Feedback rendered</i>	<i>No. and (%) of Essay 2 E errors</i>	<i>% Difference</i>	<i>No. and (%) of Essay 3 errors</i>	<i>% Difference</i>
Ann 1	0/9 (0)	Direct	1/11 (9)	9	0/12 (0)	9
Betty 2	3/11 (27)	Direct	2/10 (20)	7	0/11 (0)	20
Cleo 3	4/13 (30.8)	Direct	4/11 (36)	-5.2	3/10 (30)	6
<i>Average</i>	<i>7/33 (21)</i>		<i>7/32 (21.9)</i>	<i>-0.9</i>	<i>3/33 (9)</i>	<i>12.9</i>
Debra 1	1/12 (8)	Indirect	0/13 (0)	8	0/10 (0)	0
Eric 2	2/10 (20)	Indirect	2/13 (15)	5	1/9 (11)	4
Flora 3	4/10 (40)	Indirect	3/9 (33)	7	3/10 (30)	3
<i>Average</i>	<i>7/32 (21.9)</i>		<i>5/35 (14)</i>	<i>7.9</i>	<i>4/29 (13.8)</i>	<i>0.2</i>
Gail 1	1/12 (8)	Metalinguistic	1/10 (10)	-2	0/12 (0)	10
Helen 2	4/12 (33)	Metalinguistic	2/9 (22)	11	1/11 (9)	13
Ivy 3	3/12 (25)	Metalinguistic	3/11 (27)	-2	2/10 (20)	7
<i>Average</i>	<i>8/36 (22)</i>		<i>6/30 (20)</i>	<i>2</i>	<i>3/33 (9)</i>	<i>11</i>
Jane 1	1/13 (7.7)	No feedback	0/9 (0)	7.7	1/12 (8)	-8
Kate 2	2/11 (18)	No feedback	2/12 (16.7)	1.3	2/10 (20)	-3.3
Lucy 3	4/12 (33)	No feedback	4/11 (36)	-3	3/9 (33)	3%
<i>Average</i>	<i>7/36 (19)</i>	<i>Control group</i>	<i>6/32 (18.8)</i>	<i>0.2</i>	<i>6/31 (19)</i>	<i>-0.2</i>
Grand Total	22/101 (21.8)	For the 3 feedback groups	18/97 (18.6)	3.2	10/95 (10.5)	8.1

even a retrogression between the initial and subsequent feedback gains, demonstrates the efficacy of WCF on students' writing (See Table 7).

Error Category 4

The next error category relates to the misfit between the attributive tag's and the writer's/author's syntax and grammar. The percentages given in Table 8 are over the number of constructions with attributive language

There were fewer error manifestations in this category than in the preceding categories. No gains were made in terms of error incidence reduction among the control group. In fact, on the whole, there was a retrogression in performance relating to the fit between the attributive tag's and the writer's/author's syntax and grammar

with each new essay. Although all experimental groups registered error reduction, it was modest on account of the error incidences being generally low even in the initial essay. Where a single error was manifest in the first essay on the category, its reduction in the next essay could only be by 1; which made the percentage differential performance between the two essays small. On the whole, indirect WCF had the greatest error-reductive impact followed by metalinguistic feedback.

From Table 9, average learners profited most from the WCF than did the other performance groups with the above average group registering the least gains. The fact that the above average group started with few error incidences accounted for their low percentage growth in the second essay (where no error was noted). The

Table 7: Summary of error incidence on lack of agreement between attributive verb and number of authors

Variable considered		Essay 1	Essay 2	Essay 1 and 2 Difference	Essay 3	Essay 2 and 3 Difference
Student performance	Above average	2/33 (6%)	2/34 (5.9%)	0.1%	0/34 (0%)	5.9%
	Average	9/33 (27%)	6/32 (18.8%)	8.2%	2/31 (6%)	12.8%
	Below average	11/35 (31%)	10/31 (32%)	-1%	8/30 (26.7%)	5.3%
Feedback Provided	Direct	7/33 (21%)	7/32 (21.9%)	-0.9%	3/33 (9%)	12.9%
	Indirect	7/32 (21.9%)	5/35 (14%)	7.9%	4/29 (13.8%)	0.2%
	Metalinguistic	8/36 (22%)	6/30 (20%)	2%	3/33 (9%)	11%

Table 8: Misfit between the attributive tag's and the writer's/author's syntax and grammar

Student and academic performance	No. and (%) of Essay 1 errors	Feedback rendered	No. and (%) of Essay 2 errors	(%) Difference	No. and (%) of Essay 3 errors	% Difference
Ann 1	0/15 (0)	Direct	0/13 (0)	0	0/12 (0)	0
Betty 2	1/12 (8)	Direct	1/14 (7)	1	0/10 (0)	7
Cleo 3	2/10 (20)	Direct	2/13 (15)	5	2/11 (18)	-3
Average	3/37 (8)		3/40 (7.5)	0.5	2/33 (6)	1.5
Debra 1	1/14 (7)	Indirect	0/12 (0)	7	0/10 (0)	0
Eric 2	2/13 (15)	Indirect	1/12 (8)	7	0/11 (0)	8
Flora 3	3/12 (25)	Indirect	2/11 (18)	7	2/13 (15)	3
Average	6/39 (15)		3/35 (8.6)	6.4	2/34 (5.9)	2.7
Gail 1	1/11 (9)	Metalinguistic	0/14 (0)	9	1/10 (10)	-10
Helen 2	3/12 (25)	Metalinguistic	2/14 (14)	11	0/12 (0)	14
Ivy 3	2/10 (20)	Metalinguistic	3/12 (25)	-5	3/11 (27)	-2
Average	6/33 (18)		5/40 (12.5)	5.5	4/33 (12)	0.5
Jane 1	1/16 (6)	No feedback	1/12 (8)	-2	1/13 (7.7)	0.3
Kate 2	2/13 (15)	No feedback	2/11 (18)	-3	2/12 (16.7)	1.3
Lucy 3	2/9 (22)	No feedback	3/10 (30)	-8	4/13 (30.8)	-0.8
Average	5/38 (13)	Control group	5/33 (15)	-2	7/38 (18)	-3
Grand Total		For the 3 feedback groups				

Table 9: Summary of error incidence on lack of agreement between attributive verb and number of authors

<i>Variable considered</i>		<i>Essay 1</i>	<i>Essay 2</i>	<i>Essay 1 and 2 Difference</i>	<i>Essay 3</i>	<i>Essay 2 and 3 Difference</i>
<i>Student Performance</i>	Above average	2/40 (0.5%)	0/39 (0%)	0.5%	1/32 (3%)	-3%
	Average	6/37 (16%)	4/40 (10%)	6%	0/33 (0%)	10%
	Below average	7/32 (21.9%)	7/36 (19%)	2.9%	7/36 (19%)	0%
<i>Feedback Provided</i>	Direct	3/37 (8%)	3/40 (7.5%)	0.5%	2/33 (6%)	1.5%
	Indirect	6/39 (15%)	3/35 (8.6%)	6.4%	2/34 (5.9%)	2.7%
	Metalinguistic	6/33 (18%)	5/40 (12.5%)	5.5%	4/33 (12%)	0.5%

sole error occurrence by the above average group in the third essay then signified a retrogression. Where error occurrences were quite low, percentage computations tended to mask gains made.

Error Category 5

The shift in verb tenses was recorded differently. The researcher determined the prevalent tense in each of the student's essays and noted the instances of departure from the tense. These were not recorded in percentage terms but in the actual number of instances of departure from the tense. The most prevalent shifts were from the simple present to the past tense. Mixing the two tenses sometimes within the same paragraph created both recency (occasioned by the simple present tense) and distance (engendered by the

past tense) which was confounding. Where time frame changes in the essay presentations warranted a shift in the attributive verb tenses, these were not recorded as departures from the main verb tense employed in the essay.

Table 10 shows that for this category, the control group had greater gains than the experimental groups between essay 1 and 2 and the reverse was true for essay 2 and 3 where the feedback groups had higher gains than the control group. Again, because of the low error numbers, the differences were not very marked. The summary of the data presentation and its analysis is given in Table 11.

Summary and Analysis

Table 11 provides a summary of the percentage error reduction between the first and sec-

Table 10: Number of unwarranted departures from the main attributive tense of the essay

<i>Student and academic performance</i>	<i>No. of Essay 1 errors</i>	<i>Feedback rendered</i>	<i>No. of Essay 2 errors</i>	<i>Difference</i>	<i>No. of Essay 3 errors</i>	<i>Difference</i>
Ann 1	2	Direct	0	2	0	0
Betty 2	3	Direct	1	2	0	1
Cleo 3	3	Direct	1	2	1	0
<i>Average</i>	<i>3</i>		<i>0.7</i>	<i>2</i>	<i>0.3</i>	<i>0.3</i>
Debra 1	2	Indirect	0	2	0	0
Eric 2	3	Indirect	2	1	0	1
Flora 3	4	Indirect	2	2	1	1
<i>Average</i>	<i>3</i>		<i>1</i>	<i>2.5</i>	<i>0.3</i>	<i>0.7</i>
Gail 1	1	Metalinguistic	1	0	0	1
Helen 2	1	Metalinguistic	1	0	1	0
Ivy 3	2	Metalinguistic	1	1	0	1
<i>Average</i>	<i>1.3</i>		<i>1</i>	<i>0.3</i>	<i>0.3</i>	<i>0.7</i>
Jane 1	2	No feedback	1	1	2	-1
Kate 2	2	No feedback	2	0	2	0
Lucy 3	3	No feedback	2	1	3	-1
<i>Average</i>	<i>2.3</i>	<i>Control group</i>	<i>1.7</i>	<i>0.7</i>	<i>2.3</i>	<i>-0.7</i>
Grand Total	2.3	For the 3 feedback groups	1	1.3	0.3	0.6

Table 11: Feedback effect per student per error category across the 3 essays

Student and performance	Feedback rendered	Error category 1			Error category 2			Error category 3			Error category 4			Error category 5		
		1and2 % Difference	2and3 % Difference	1and3 % Difference	1and2 % Difference	2and3 % Difference	1and3 % Difference	1and2 % Difference	2and3 % Difference	1and3 % Difference	1and2 % Difference	2and3 % Difference	1and3 % Difference	1and2 % Difference	2and3 % Difference	1and3 % Difference
Ann.1	Direct	5	15	-1	6	9	9	9	0	0*	2	0*	2	0*	0	
Betty 2	Direct	6	5.7	18	2	7	20	1	7	7	2	1	2	1	1	
Cleo 3	Direct	4	1	8.7	2.2	-5.2	6	5	-3	2	2	0	2	0	0	
Debra 1	Indirect	11	6.7	10.3	-5.3	8	0*	7	0*	2	2	0*	2	0*	0*	
Eric 2	Indirect	21	16	14	7	5	4	7	7	1	1	1	1	1	1	
Flora 3	Indirect	13	22	11	0	7	3	7	7	2	2	1	2	1	1	
Gail 1	Metalinguistic	4	6	-5.3	13	-2	10	9	-10	0	0	1	0	1	1	
Helen 2	Metalinguistic	6.3	10	20	5.3	11	13	11	14	0	0	0	0	0	0	
Ivy 3	Metalinguistic	0	-13.6	18	-3	-2	7	-5	-2	1	1	1	1	1	1	
Jane 1	None	0	-13	-0.3	-9.7	7.7	-8	-2	0.3	1	1	-1	-1	-1	-1	
Kate 2	None	1	-5	2	13.3	1.3	-3.3	-3	1.3	0	0	0	0	0	0	
Lucy 3	None	5.6	4	9.4	2.4	-3	3	-8	-0.8	1	1	-1	-1	-1	-1	

Actual
count not %

ond as well as between the second and the third essays on each of the four error categories with the fifth category identifying only the actual error count differences. Table 11 shows the extent to which the feedback rendered led to error reduction with each new essay in the different error categories. The instances where there was no change in the percentage errors in each category, and instances where after the feedback errors increased in the subsequent essay, are highlighted in bold.

Of the 9 students \times 10 comparisons (instances in which essay 1 and 2 and 2 and 3 were compared for the 5 error categories), only 24/90 (26.7%) showed either no change in error quantity or an increase in the errors from one essay to the next within particular error categories. Exactly half of these instances showed neither increase nor decrease in the error incidence. The 0* denotes students who had no error in both essays being compared on a particular category and so could not attain anything better than 0 percent. Considering these under non-improvement would be misleading and so the five 0* were removed from the 24 to remain with 19/90 (21%) instances of non-improvement despite feedback provision. An overwhelming 79 percent instances of error reduction testified to the efficacy of feedback.

Following the same pattern of analysis, 5/30 (16.7%) of direct feedback instances did not lead to error reduction, 2/30 (6.7%) of indirect feedback instances did not result in error reduction, and 12/30 (40%) of metalinguistic feedback led to no reduction in errors. It was apparent that indirect feedback was the most efficacious of the three feedback forms and metalinguistic feedback was the least effectual. This vindicates Bitchener's (2008: 104) assertion that:

Once the error has been noted, indirect feedback has the potential to push learners to engage in hypothesis testing—a process which Ferris (2002) and others (see Doughty and Williams, 1998) suggest may induce deeper internal processing and promote the internalization of correct forms and structures.

Storch's 2009 study reported in Storch (2010), found that on short writing tasks (150-200 words), direct WCF was more efficacious than indirect WCF whereas longer writing tasks (250-300 words) were more amenable to indirect WCF. The findings would explain the edge indirect feedback had over direct feedback in the

present study where recommended essay length was 4000-6000 words. The explanation was that in short writing tasks, the direct feedback could be held in memory whereas in longer pieces of writing the student would not be able to keep in memory the direct feedback given before. The need for deeper processing mentioned earlier which is occasioned by indirect feedback would then be most applicable for longer texts.

Most students seemed not to benefit from lengthy feedback characteristic of metalinguistic feedback, if at all they took time to read through the feedback. Another possible explanation was that the adjacency of feedback to the point of error had improved chances of association between the two and comprehending the feedback better than where feedback was left until at the end of the piece of writing.

The control group which did not receive the WCF fared badly in terms of error reduction from one essay to the next with 16/30 (53%) of the instances showing no error reduction compared to the 21 percent for the experimental group. This was a strong argument in favour of feedback in improving students' writing.

It was also important to determine the extent to which the efficacy of feedback colligated with students' general academic performance. Above average performers (with 1 to their names) had 7/30 (23%) instances of non-improvement, average performers had 2/30 (6.7%) instances of non-improvement and the below average performers had 10/30 (33%) instances where they did not register error reduction. The assumption was that the below average students were prone to misunderstand the feedback or to fail to act on the feedback and that the above average students were less likely to consider feedback closely once they had scored highly. Concerning indirect feedback, Storch (2010) notes that it is based on the assumption that what the student needs is to notice the error and because the structure is known, the student would be able to work out the correct form. That observation supports the present study's perceived link between the efficacy of feedback and the academic potential of the student.

There was also an extent to which the different error categories were susceptible to feedback. In the first error category all students registered a reduction in errors for both essay 1 and 2 and 2 and 3 comparisons. The instances of

non-improvement in the other categories were: 5/18 (27.8%) category 2; 3/18 (16.7%) category 3; 5/18 (27.8%) category 4; and 4/18 (22%) category 5.

CONCLUSION

From the foregoing, it is apparent that feedback generally has potential to lead to improvement in students' writing and that indirect feedback holds the greatest promise for transforming students' writing. The efficacy of feedback however, depends on the extent of students' academic potential to profit from it. There was sufficient evidence to conclude that generally, feedback can feed forward to students' subsequent writing. The widely used "Input-Interaction-Output Model" in feedback assessment and provision with no regard for the context of language use and learning is limiting and limited. The need for contextualising the efficacy of feedback when advised the need to conclude on the efficacy of feedback on the basis of context-specific study. Some scholars warn against attempting to generalise about the efficacy of WCF on the basis of a narrow range of structures. An attempt to universalise the efficacy of particular feedback for all learners and for all contexts is untenable as the efficacy of feedback types is relative to diverse variables.

RECOMMENDATIONS

In view of the study's findings, the study observes that the debate about which feedback type is effective is a retrogressive one as different studies would always produce conflicting findings because of the particular contexts. The study therefore, recommends that researchers be context specific in their conclusions about the efficacy of feedback and avoid a one-size-fits-all claim. Those who provide feedback to students (lecturers and teachers) should consider diverse variables which give insight on what feedback type profits what structural or

semantic aspects of students' writing and employ it. There is, therefore, need for action research within one's context in order to discover what feedback types best match what writing aspects. The time thus invested would go a long way to ensuring that valuable time is not invested in feedback types which would marginally, if not, hardly enhance students' writing on particular aspects.

REFERENCES

- Al Ajam A 2015. The effect of written corrective feedback on Omani students' accuracy in the use of English prepositions. *Advances in Language and Literary Studies*, 6(1): 61-71.
- Bankier J 2012. Post-text and in-text corrective feedback. *Language Education in Asia*, 3(1): 85-95.
- Bitchener J 2008. Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17(2): 102-118.
- Bitchener J, Young S, Cameron D 2005. The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 9: 227-258.
- Cohen L, Manion L, Morrison K 2007. *Research Methods in Education*. 6th Edition. London: Routledge.
- Ellis R 2009. A typology of written corrective feedback types. *ELT Journal*, 63(2): 97-107.
- Gray BE 2011. *Exploring Academic Writing through Corpus Linguistics: When Discipline Tells Only Part of the Story*. Doctoral Dissertation. USA: Northern Arizona University (UMI No. 3490519).
- Lea MR, Street BV 2006. The 'Academic Literacies' Model: Theory and applications. *Theory into Practice*, 45(4): 368-377.
- Lillis T 2003. Student writing as academic literacies: Drawing on Bakhtin to move from critique to design. *Language and Education*, 17(3): 192-207.
- MacLure M 2006. The bone in the throat: Some uncertain thoughts on baroque method. *Qualitative Research in Education*, 19(6): 729-745.
- Ostovar-Namaghi SA, Shakiba K 2015. Theorizing EFL teachers' perspectives and rationales on providing corrective feedback. *The Qualitative Report*, 20(6): 727-745.
- Sheen Y 2007. The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 4: 255-283.
- Storch N 2010. Critical feedback on written corrective feedback research. *International Journal of English Studies*, 10: 29-46.