RESEARCH ARTICLE

ISSN
INTERNATIONAL STANDARD SERIAL NUMBER INDIA

2395-2636 (Print): 2321-3108 (online)

LEXICAL COHESION IN ESL WRITING

Dr. MADHAVI GAYATHRI RAMAN[1*], SHEFALI ELIZABETH MATHEW[2]

[1]Assistant Professor, The English & Foreign Languages University, India
*Corresponding Author: Email: gayathriraman@yahoo.com
[2]The English & Foreign Languages University, India

Abstract

This paper investigates the role of lexical cohesion in the narrative writing of ESL learners using the Halliday and Hasan (1976) framework. Lexical cohesion refers to the manner in which lexical items combine to create continuity in discourse. Cohesion has been found to play a significant role in language production and comprehension. Research on cohesive devices in writing focuses on the number and classification of devices and their interaction with coherence and genre. The study discussed here employs five different types of lexical cohesion - repetition, synonymy, super-ordinate, general item and collocation, and four distance types - immediate, remote, and mediated, to examine how their use contributes to the overall writing quality of narratives produced by thirty Indian ESL learners from diverse SES and linguistic backgrounds. The narratives were composed under test conditions in response to a prompt as a part of a course on writing. The written scripts were also rated using the IELTS rating scale. Repetition has been found to be indicative of writing quality. Our findings suggest that students with low proficiency and lower ratings tend to employ repetitions more often than other types of lexical cohesion. Students with Intermediate level proficiency obtain higher IELTS ratings and use a greater number of mediated ties as well as a higher number of collocations and synonymy.

Keywords: lexical cohesion, cohesive devices, narrative writing, ESL learners

INTRODUCTION

This paper examines the relationship between lexical cohesion and writing quality with a view to determining how specific types of lexical cohesion impact writing quality, i.e., does the use of certain types of lexical cohesion improve or weaken the quality of a written text? In this paper we draw upon the classification of lexical cohesion used by Halliday and Hasan (1976) and the taxonomy of cohesive ties proposed by Witte and Faigley (1981) which in itself is based on that of Halliday and Hasan (1976).

COHESION

A text may be defined as something that is spoken or written which forms a unified whole. According to Halliday and Hasan (1976), it is a semantic unit…not of form but of meaning. Thus, the various parts of a text would tend to talk about the same things, i.e., all texts possess the quality of unity. This quality of unity or cohesion refers to "relations of meaning that exist within the text and that define it as a text" (Halliday & Hasan, 1976). According to Hoey (1991:3), cohesion is "the way certain words or grammatical features of a sentence can connect that sentence to its

KY Publications
www.kypublications.com

predecessors (and successors) in a text". The parts of a text are held together by explicit cohesive ties - the semantic relation between one element in a text and another which is critical to its interpretation. These ties establish continuity in a text. Cohesive ties are broadly classified into two: (i) grammatical relations (reference, ellipsis etc.) and (ii) lexical relations (synonymy, collocations), and allow us to understand sentences as connected discourse rather than as independent sentences.

In the Hallidayan (1976) framework, lexical cohesion refers to the non-structural, text-forming relations. It is concerned with the way in which words relate to each other so that the continuity of the text is not only created but also maintained. It also provides vital information about the manner in which words are organized in discourse (Halliday & Hasan, 1976). The two main types of lexical cohesion that we examine in this paper are: reiteration and collocation. *Reiteration* refers not only to an instance of repetition of the same lexical item but also to the use of a related word. It covers a range of categories from synonymy or near synonymy to a general word that stands for the entire class. Thus, under reiteration are included repetition (same word), synonym or near synonym, super-ordinate item, and general word. For example:

1. There's a temple at the top of a hill.

    a. The ascent to the top wasn't easy.

    b. The climb was difficult because of the heat.

    c. As we got there, we could hear the bell ring.

    d. That's one of the things I like about places of worship.

In (a) the use of *top* provides an example of repetition of the same item. In (b), the word *ascent* is used, while in (c), its synonym *climb* is used to achieve cohesion. In (d), the use of *places of worship* provides an example of superordinate category of which *temple* is a member. *Things* in (d), which is an example of the category general word, also ties up with *bell*.

Collocation refers to cohesion that is "achieved through the association of lexical items that regularly co-occur" (Halliday and Hasan, 1976). This type of cohesion is considered to be the most difficult to analyse since the items that collocate involve none of the categories mentioned under reiteration. What is most important for collocation to be achieved is that the items share the same lexical environment. For example:

2. They went to Goa all set to spend their days on the beach.

3. They packed swimsuits, goggles, hats, sunscreen, beach towels and sunglasses.

The words swimsuits, goggles, hats, sunscreen, beach towels and sunglasses occur in the same lexical environment as beach and therefore are considered as examples of collocations.

When items like these appear in sentences that are adjacent, they generate a cohesive force. It is possible for long cohesive chains to be built out of collocations like *hair…brush…smooth, garden….fertilizer…plants, mountain…summit…base camp* and so on.

In addition to the classifying cohesive ties into the categories mentioned above on the basis of the function, Halliday and Hasan also employ a second taxonomy. On this taxonomy, cohesive ties are classified according to "the amount of text spanned by the presupposed and presupposing elements of a given tie" (Witte & Faigley, 1981). They propose four "text-span" categories. Membership in a category is determined by the number of T units or minimally terminable units (which are the shortest grammatically allowable sentences into which writing can be analysed. Most often, a T-unit is a sentence) that are spanned by a particular cohesive tie. Combined, the two taxonomies allow a particular tie to be classified in two different ways: one on the basis of function and the other on the basis of distance. The four "text-span" categories are: (i) Immediate, (ii) Mediated, (iii) Remote, and (iv) Mediated-Remote. Let us examine the following sentences from a student's writing:

4. I visited Golkonda fort which is in Hyderabad last week.

5. Actually, before coming to Hyderabad I didn't know anything about it.

6. The Golkonda is a very big and old fort in the city.

7. When I joined EFL University, I asked my seniors to suggest places in the city that I must see.

8. They told me and my friends that one of the best places to visit in the city was Golkonda.

9. But we never went there unit last week.

10. And that was because I saw a picture of a friend on Facebook in front of Golkonda.

Immediate cohesive ties link T-units that are adjacent to one another. Thus, the repetition of *Hyderabad* in sentences (5) creates an *Immediate* tie between (4) and (5). In contrast, the repetition of *city* in sentences (6), (7) and (8) forms a *Mediated* tie. The semantic bridge that is established by the occurrence of *city* in (6) and (8) is mediated by its repetition in (7). While it might seem that the repeated occurrence of the word *city* could be a series of immediate ties, this is not the case because once a lexical item has been introduced, all subsequent appearances or uses of that item presuppose the first one. Mediated ties like Immediate ties link items in adjacent T-units. These kinds of ties allow writers to introduce a thought or idea in one T-unit and to expand or modify it in subsequent T-units.

When the two elements of a cohesive tie are separated by one or more intervening T-units, the tie is said to be a Remote one. For instance, the tie between *friends* in (8) and (10) is *Remote*. Mediated-remote ties are those ties that are both mediated and remote. E.g., consider the repetition of *Golkonda fort* in sentences (4), (8), and (10).The presupposing *Golkonda fort* in (10) is separated from the presupposed *Golkonda fort* in (4) by intervening T-units (5), (7) and (9) which do not contain this element and so the tie is remote. However, the presupposing *Golkonda fort* is also

mediated through its repetition in (6) and (8), making it a mediated-remote one. Mediated-remote ties are used by writers to weave the main themes into a text.

Lexical cohesion and writing quality

Teachers of English often find that the writing skill of undergraduate students is quite poor despite the fact that they have been learning the language as well as learning how to write - paragraphs and essays in particular - for a number of years. An analysis of the written scripts of students has shown that their writing is characterized by a paucity of and poor organization of ideas, incomplete sentences, direct translation from the mother tongue and poor grammar. This could be due to the fact that the teaching of writing does not receive the kind of focussed attention that it deserves since teachers are more involved either in the teaching of grammar or the unit of study per se. Given that the examination system is largely memory-based, the students' writing ability is not really tested. It is at institutions of higher education (and especially in those that deal with the Humanities), that students find themselves at a loss when it comes to written communication. Since academic performance at this level is gauged largely on the basis of written performance, it is essential that students are able to produce coherent and cohesive pieces of writing.

In this context, a number of researchers have used Halliday & Hasan's (1976) taxonomy to examine cohesion in the essays of learners with varying language proficiency levels. Cherry and Cooper (1980) used text span (the distance between cohesive ties) to study the essays of students in grades 4, 8, 12, and college level. According to them, writers tended to depend more on lexis rather than on reference and conjunctions as they matured. Neuner (1987) investigated the essays of freshman writers and found that cohesive chains rather than cohesive ties helped to discriminate good essays from poor ones. According to him, in a good essay, the chains are spread over greater distances and the writing is characterized by "longer chains, greater variety of words, and greater maturity of word choice"

(Neuner, 1987). In another study, Witte & Faigley (1981) analyzed five good and five poor freshman essays with a view to examining the relationship between cohesion, coherence, and writing quality. They found that a large percentage of the cohesive ties used were lexical in nature. The good essays were characterised by a greater number and variety of cohesive ties, more conjunctive and reference ties, and collocations (lexical cohesion ties) than the poor essays. The latter contained more lexical repetition (65%) than the good essays. This could be attributed to the students' impoverished vocabularies and their limited ability to elaborate and expand ideas. According to Witte & Faigley, lexical cohesion is an indicator of writing quality and it is lexical collocation that serves to distinguish between good and poor writing.

In a study that examined the cohesion features of ESL students' written texts, Connor (1984) found that ESL writers tend to rely on repetition, frequent use of synonymy and collocation. This is in contrast to native speakers who use repetition rather infrequently. Based on her findings, she suggests that the use of lexical cohesion may follow a developmental pattern with improved proficiency leading to better writing. The implication here is that this is related to vocabulary development and this finds echoes in Carrell (1983). In another study, Crowhurst (1987) studied the types of cohesive ties used at three grades. Repetition was the most frequently used kind of cohesion along with demonstratives, pronominals and the definite article. In higher grades there is increased use of synonyms and collocation and repetition decreases. While in lower grades, repetition is indicative of immature lexical choices, in the higher grades it reflects students' tendency to elaborate and summarize their arguments. The increased use of synonyms and collocation in the higher grades is indicative of a larger vocabulary and the ability to elaborate arguments. This result finds echoes in those of Witte and Faigley (1981) and Ferris (1994) who found that advanced ESL students use a greater number of lexical and syntactic devices in their compositions in addition to a wider range of cohesive devices such as synonymy, antonymy, referencing and the definite

article. Poor writers on the other hand tend to employ lexical repetition more often in their writings to promote textual cohesion.

Given the findings in the literature and the overall poor writing ability of undergraduate students enrolled in our University, we decided it would be worthwhile to examine their writing to understand students' ability to use lexical cohesion devices and at the same time to examine the relationship between overall writing ability (as revealed by a composite score like the IELTS score) and lexical cohesion. The questions we address are the following:

1. What are the kinds of lexical cohesive devices – repetition, synonymy, superordinate item and general item – that are most frequently used by our students?

2. What kinds of ties – immediate, mediated, remote and mediated-remote ties – are used most frequently?

3. What patterns of use differentiate good writers from poor writers?

Sample

The sample of this study initially consisted of 30 students enrolled in the first year of the BA Foreign Languages programme at the English and Foreign Languages University, Hyderabad. However, one student's script had to be left out of all analyses as it was incomplete in nature.

These students belong to mixed SES backgrounds and come from the urban as well as rural areas of India. They speak a variety of Indian languages such as Hindi, Telugu, Tamil, Malayalam, Bengali and Punjabi as their mother tongues. They have all learnt English at school where it has been either been taught as a subject or was the medium of instruction throughout the school years. Thus the sample is extremely heterogenous in nature. While pursuing an undergraduate degree in a variety of foreign languages such as Spanish, French, Japanese, German, Arabic at the University, these students are also required to study English as a compulsory subject for the first four semesters.

Most of the courses offered in English across the semesters are designed to help them improve their communications skills, academic reading and writing abilities. On the basis of a diagnostic written test that assessed their general language proficiency administered by the University at the beginning of the year, the 29 students were roughly identified as belonging to one of two proficiency levels: (i) low proficiency level group consisting of 15 students (which corresponds roughly to A2-B1 levels of the CEFR scale) and (ii) intermediate level proficiency group comprising 14 students (which corresponds roughly to B1-B2 levels of the CEFR scale).

Methodology

A written test consisting of one expository task and one narrative task were administered after regular class hours on a working day in one of the classrooms of the University. Students were provided the details of the task on a task sheet and sufficient space was provided below each of the tasks for them to write their responses. Before they attempted the task, the instructions were made clear and clarifications were provided when necessary. Students were encouraged to complete one task, hand in the papers and then move on to the next one. The task was constructed on the lines of the IELTS writing task in which students have to write a short paragraph based on a question for which prompts are provided. These prompts are provided to ensure that all students respond adequately to the question and also to ensure that the conceptual load placed on them by task demands is minimised. Familiar topics were chosen and three prompts were provided with a view to helping them to structure their writing. The IELTS

rating scale was used to assess the scripts for overall writing ability. This was done by two trained IELTS examiners. The Halliday and Hasan (1976) framework for analysing lexical cohesion at two levels, one of function and the other of text span, were then used to examine the relationship between lexical cohesion and writing quality. We felt that using these two measures would give us a relatively comprehensive picture of the writing ability of the students. This in turn would help structure an effective writing course for these students. This paper presents an analysis and discussion of the data from the narrative writing samples.

Analysis and discussion

As mentioned earlier, the scripts were rated by two trained IELTS raters to get a wholistic picture of the students' writing ability. (The scripts were masked so that the raters did not have any information about the students or their proficiency level as determined by the diagnostic test.) We found that the 15 low proficiency level (LPL) students scored between 3.5 and 4.5 on the IELTS rating scale. The average length of these scripts was 139.6 words per script with 69 being the least number of words and 224 being the highest. The 14 intermediate level (IPL) students scored between 5.0 and 7.5 on the IELTS rating scale with 11 students scoring 5.5 or more. This score clearly indicates that they possess an intermediate level of language proficiency. The average length of their scripts was 236.9 words per script with 179 being the least number of words and 374 being the highest number of words in a script. The table and graphs below present an analysis of the kinds of lexical cohesion ties used by the two groups.

Table 1. LPL: Types of cohesion

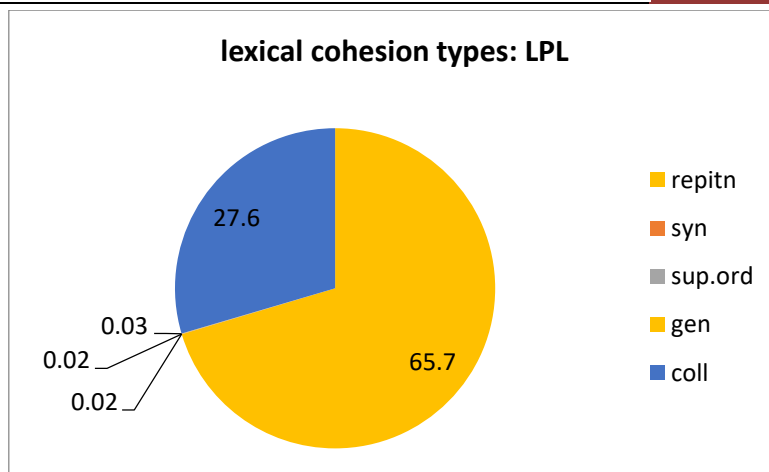|	Types of cohesion					
	Rep.	**Syn.**	**Sup.ord.**	**Gen. item**	**Coll.**	**Total**
LPL students (15)	150	5	4	6	63	228
IPL students (14)	216	18	18	6	113	371

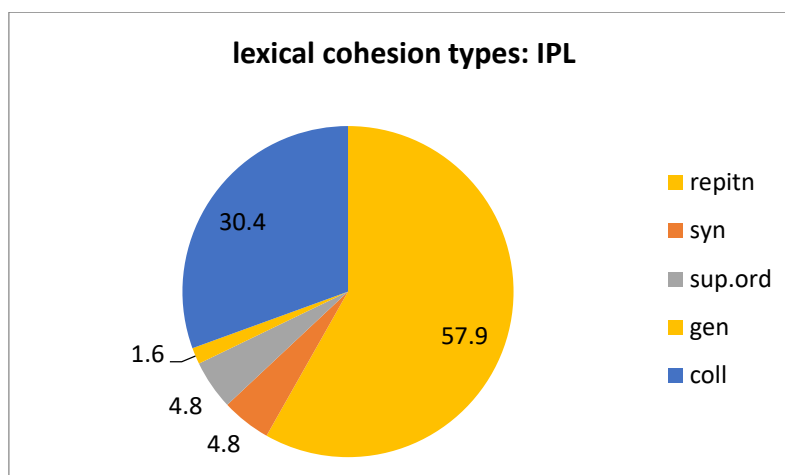Figure 1: LPL: Types of cohesion

Figure 2: IPL Types of cohesion

As we can see from fig.1, in the LPL group, repetitions of the same item have been produced by all the students and make up for the largest category of lexical cohesive ties (150/228) accounting for 65.7%. This mirrors the findings in other studies like Witte and Faigley (1981), Kadiri et. al. (2016) in which the authors found that the low ability students produced high percentages of repetitions. Unlike in these studies however, our students produced a large relatively number of collocations, which account for 63/228 ties or 27.6% of all ties.   A close analysis of the scripts shows that students belonging to this level were able to use words related to one another semantically such as *teacher, classroom, principal, assignment, grade, exam.* What is interesting to note is that the topic required them to write about

an occasion when they had to make an excuse and with the exception of a few, almost all of them wrote about a time when they made excuses to avoid an exam, writing an assignment or attending class. Since this is something they're extremely familiar with, it could've been easy for them to produce and use words that belong to the same semantic field. However, we find that these words were used by them repeatedly thereby leading to a repetition of ideas instead of elaborating upon them as we would expect. The constant repetition of words also makes for tiring reading. While we do find that they have used referential pronouns and conjunctions, these are not discussed here as they do not form the focus of this paper. The other three categories of synonyms, superordinate category and general item accounted for merely 0.02%-

Dr. MADHAVI GAYATHRI RAMAN & SHEFALI ELIZABETH MATHEW

0.03% each showing that students have limited abilities in using their vocabulary and do not go beyond the familiar domains. Here is an extract from a student's script:

I made an excuse to avoid to write an examination when I was in my intermediate. First I made a call to my sister and I told her that I am not going to write the exam and she asked me why and i told her that I am not prepared for the exam. She told me for so many times to write the exam but I didn't agree with her. The excuse I made to avoid writing exam is ….

As is evident from this piece of writing, the repetitions of *write*, *exam*, *avoid* and *excuse* contribute to the achievement of cohesion, but the redundancy makes it flawed and reveals the writer's inability to rephrase the ideas or build upon it by introducing new information. This is somewhat in contrast to the pieces of narrative writing produced by the students belonging to the IPL group.

An examination of the pattern of production of lexical ties by the IPL group reveals certain similarities with and differences from the LPL group. Fig. 2 shows that though repetitions accounted for the highest percentage of cohesive ties produced (57.9%), this percentage is relatively smaller than that produced by the LPL group. The percentage of collocations produced by the IPL group is higher at 30.4%. In this group we see some use of synonyms (4.8%), superordinate (4.8%) and general item (1.6%). This shows that the writing ability of this group is better. An analysis of the scripts reveals that the sentences are longer, better formed and connected by the use of a greater variety of words that occur in the same lexical environment. Ideas are introduced and elaborated upon in a better manner than in the LPL scripts. There is also a greater range of vocabulary in use. Let us look at this extract from a student's script:

So quickly I thought of an excuse – I would pretend to be ill. I started executing my plan hoping that my parents won't send me to tuition that day. I started going to the washroom every 20 minutes and made it a point not to come out before 10 minutes. Yes, I was pretending that I had the loosies. My parents gave me medicines and I carefully threw them away behind their backs, after all who wants to take those bitter medicines without any cause… were really worried about my condition ….

As we can see, there is a certain amount of lexical sophistication in this script. The writer introduces ideas and builds upon them using a variety of phrases. Ties are established across sentences through the use of not merely repetitions but also collocations and these can be found within and across sentences. We find that students belonging to this group tend to use a greater range of vocabulary which also contributes to their higher IELTS score. As Witte and Faigley (1981) point out, "skill in invention, in discovering what to say about a particular topic, may depend … on the prior development of adequate working vocabularies" and this better developed vocabulary appears to distinguish the better writing samples of the IPL group from those of the LPL group.

In addition to the types of lexical cohesive ties, we also analysed the distance between these ties, i.e., the text span. According to Halliday and Hasan (1976) and Witte and Faigley (1981), good pieces of writing are characterised by immediate and mediated ties that introduce and integrate information into the text while mediated-remote and remote ties reach across units, suggesting that the writers of poor quality essays are unable to elaborate ideas. However, our results provide a slightly different picture as we shall see from the table and figures below.

Dr. MADHAVI GAYATHRI RAMAN & SHEFALI ELIZABETH MATHEW

Table 2. Distance of ties

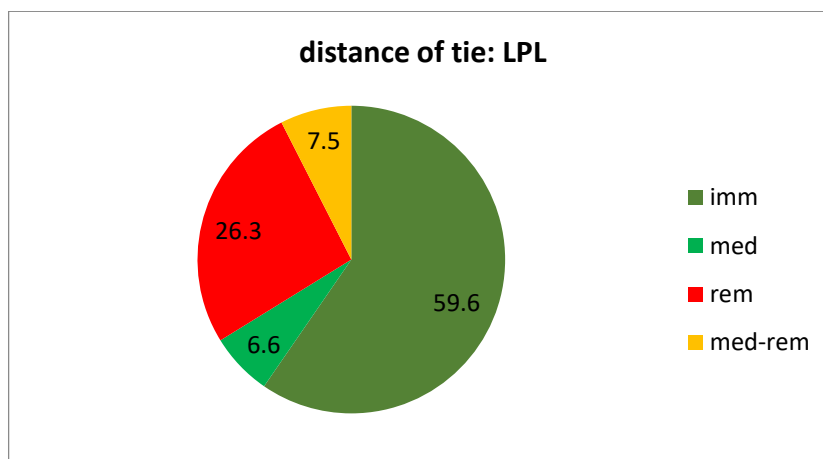|	Distance of ties				
	Imm	**Med**	**Rem**	**Med-Rem**	**Total**
LPL (15 students)	136	15	60	17	228
IPL (14 students)	215	102	46	8	371

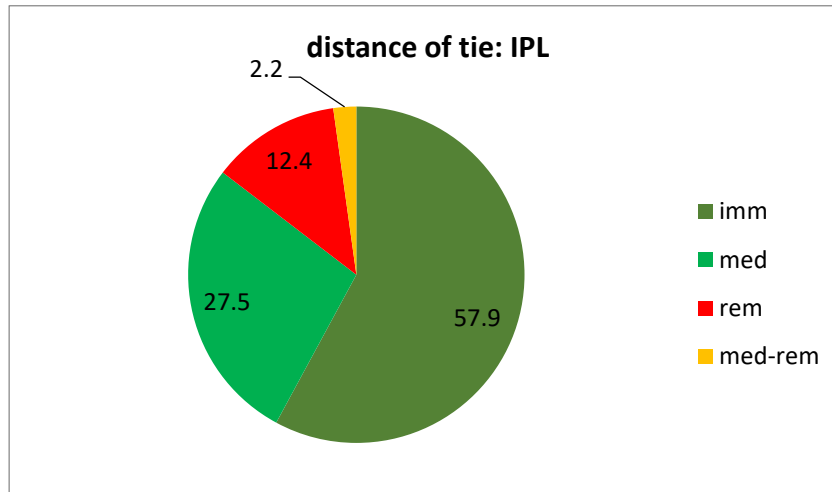Figure 3: LPL Distance of ties

Figure 4: IPL Distance of ties

Figures 3 and 4 reveal that there is very little difference between the two groups in terms of the distance between cohesive ties used. Both groups tend to use more than 50% immediate ties which should contribute to a better quality of writing, However, as we have already discussed, these immediate ties tend to be more in the form of repetitions than collocations, synonyms or superordinate items and so take away from the quality of writing. Though cohesion is maintained in the writing, the quality of writing is affected by the overall limited vocabulary that our students possess. We find that what distinguishes the two groups is the use of mediated ties, with the IPL group using more of such ties in comparison to the LPL group. The IPL group tends to links ideas across sentences by building bridges to link ideas either through the use of collocations or by repeating the

Dr. MADHAVI GAYATHRI RAMAN & SHEFALI ELIZABETH MATHEW

same item. They also use a number of referential pronouns (which as we have mentioned before do not find a place in this analysis).
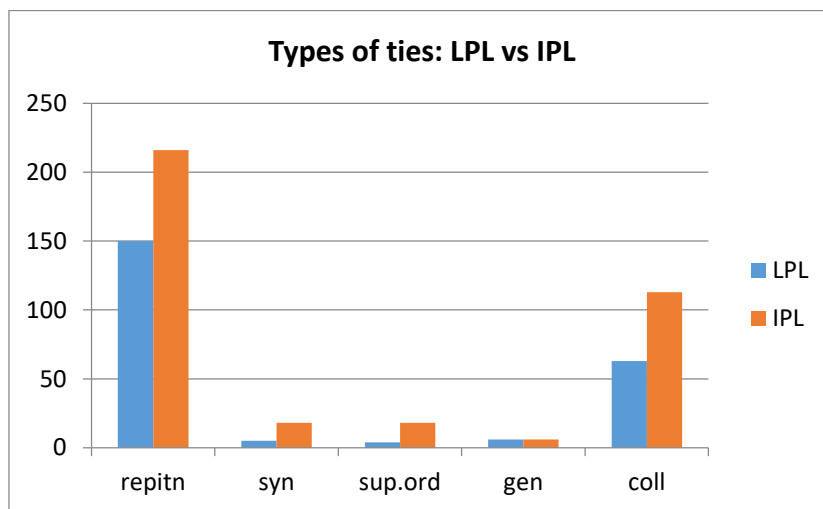
Comparing the groups

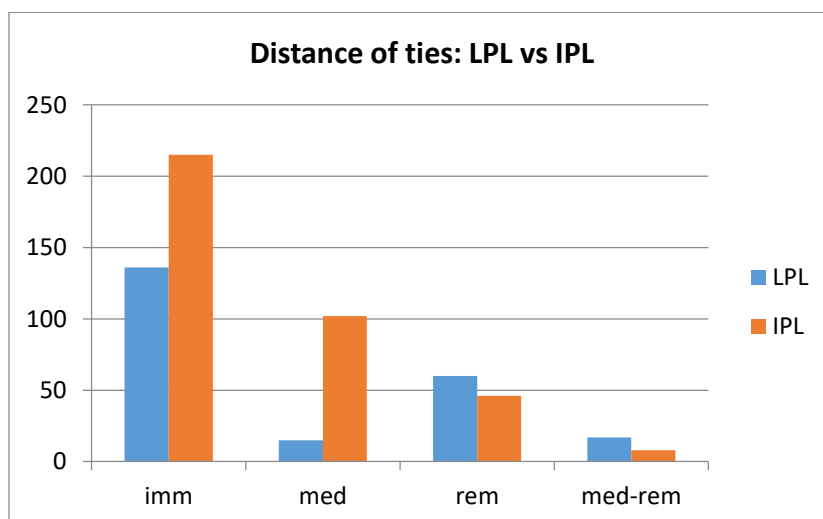Figure 5: Types of ties: LPL vs. IPL

Figure 6: Distance of ties: LPL vs. IPL

These two graphs show the difference between the two groups in terms of types of cohesive ties and the distance of the ties. In fig. 5, the IPL group produces more responses across all five types of cohesive devices than the LPL group. Though there are more repetitions which we would not expect, we note that this could be due to the longer scripts and the fact that even the IPL group has a slightly limited vocabulary though much better than the LPL group. This has also helped them score better on the IELTS scale. As we see from fig. 6, the IPL group records a larger number of immediate ties but these tend to be repetitions.

What serves to distinguish the two groups is the greater number of mediated ties as well as the higher number of collocations and a limited use of synonymy.

Conclusion

Our results suggest that merely having a high number of cohesive ties alone is not enough to enhance writing quality, what matters is the spread across types of ties and the ability to spread these ties across the piece of writing. This means that students should be capable of using not just repetition in their writing but also collocations,

Dr. MADHAVI GAYATHRI RAMAN & SHEFALI ELIZABETH MATHEW

synonyms and the other categories to be able to establish both immediate and mediated ties. This in turn has implications for vocabulary teaching and the teaching of writing. In an ESL context, a great deal of importance is placed on the teaching and learning of vocabulary as the belief is that a better vocabulary will contribute to better language proficiency. While this is true, we fail to teach our students how to translate this into actual language use in production, i.e., speaking and writing. We therefore need to show our students how their receptive vocabulary can translate into a productive repertoire thereby improving the quality of their writing (and speech). This could be achieved by following a process approach where we use word webs, concept maps and the like so that students work and rework on their pieces of writing until they are able to generate better formed written texts.

References

Carrell, P. L. Schema Theory and ESL Reading Pedagogy. *TESOL Quarterly*, vol. 17, no.4, (1983). pp. 553–573.

Cherry, R., & Cooper, C. *Cohesive ties and discourse structure: A study of average and superior texts at four grade levels*. (1980). Unpublished Manuscript. Department of Learning and Instruction. State University of NY at Buffalo.

Connor, U. A study of cohesion and coherence in English as a second language students' writing. *Papers in Linguistics: International Journal of Human Communication*, vol. 17, pp. (1984). 301-316.

Crowhurst, M. Cohesion in argument and narration at three grade levels. *Research in the teaching of English,* 21, (1987). pp. 185-201.

Ferris, D. R. Lexical and syntactic features of ESL writing by students at different levels of L2 proficiency. *TESOL Quarterly,* vol. 28, no. 2, (1994).pp. 414-420.

Halliday, M. A. K., and Hasan, R. *Cohesion in English*. (1976).London: Longman.

Hoey, Michael. *Patterns of lexis in text*. Oxford: Oxford University Press, 1991.

Kadiri, G., C., Igbokwe, U., I., Okebalama, U. N., and Egbe,C. I. The use of lexical cohesion elements in the writing of ESL learners. *Research in Language*, vol. 3, (2016). pp. 222-234.

Neuner, J. N. Cohesive Ties and Chains in Good and Poor Freshman Essays. *Research in the Teaching of English, vol.*21, no.1, (1987). pp. 92-105.

Witte, S. P. and Faigley, L. Coherence, cohesion, and writing quality*. College Composition and Communication,* vol. 32, no.2, (1981). pp. 189-204.