Analyzing frequency and distributional pattern of discourse markers in the Malaysian newspaper

Hafizah Hajimia

School of Languages, Civilization and Philosophy Universiti Utara
Malaysia
dr83hafizah@uum.edu.my

Manvender Kaur Sarjit Singh*

School of Languages, Civilization and Philosophy Universiti Utara
Malaysia
manvender@uum.edu.my

Mohan Rathakrishnan

School of Languages, Civilization and Philosophy Universiti Utara
Malaysia
rmohan@uum.edu.my

Abstract

Numerous studies showed a discourse marker is a confusing marker and when it is related to various genre, deems necessary for a thorough study. Thus, frequency and distributional patterns of DMs would be able to clear this confusion. The frequency and the pattern of discourse markers 'and' were examined to see the usage of discourse markers between the sentences and the articles of the news report by applying the corpus tools. The result indicated that different grammatical word classes of discourse markers 'and' exist in the articles and various positions of DMs 'and' was identified based on its distributional patterns.

Keywords: Discourse markers, Corpus-based approach, Malaysian online newspaper articles, Fraser, CACA.

Análisis de la frecuencia y el patrón de distribución de los marcadores del discurso en el periódico Malayo

Resumen

Numerosos estudios demostraron que un marcador de discurso es un marcador confuso y cuando se relaciona con varios géneros, se considera necesario para un estudio exhaustivo. Por lo tanto, la frecuencia y los patrones de distribución de los DM podrían aclarar esta confusión. La frecuencia y el patrón de los marcadores del discurso 'y' se examinaron para ver el uso de los marcadores del discurso entre las oraciones y los artículos del informe de noticias mediante la aplicación de las herramientas del corpus. El resultado indicó que existen diferentes clases de palabras gramaticales de marcadores del discurso "y" en los artículos y se identificaron varias posiciones de los DM "y" según sus patrones de distribución.

Palabras clave: Marcadores del discurso, enfoque basado en corpus, artículos de periódicos en línea de Malasia, Fraser, CACA.

1.INTRODUCTION

In this era, writing is a technology (Grabe & Kaplan, 1996). It evolves over time. Writing represents a set of skills that are practiced and learned over time (Manvender, 2014). In the present study, writing is perceived as an activity that combines both technology and particular skills required in order to prepare a written communicative event such as a newspaper article. The written communicative event is defined as the written communication occurring in replace of spoken or

verbal communication and represents various events such as newscasting, agreements, and procedures.

In order to prepare and produce a comprehensive written communicative event, structural awareness which often lacks in non-native speakers of a language is required. According to Manvender (2014), for a non-native speaker to be able to understand and produce a written communicative event, it is necessary for the non-native speaker to be aware of the structural rhetorical organization and the linguistic conventions of the written communication to be produced. Further elaborating it, the researcher stressed that non-native speaker of a language "must realize the communicative purposes of the text to be produced" (Manvender, 2014).

Based on a study carried out by regarding the coherence in writing among English Second Language learners (L2) in UTM, Malaysia, they found that students faced problems with coherence in writing whereby they focused more on the sentence and word level rather than the textual coherence. In line with this, Shukor, et al. (2014) carried out an observation on form 4 student essay writing and realized that they have a problem understanding the position and the functions of discourse markers and lacked variety when writing.

Understanding various forms and functions of a new genre is crucial in order to comprehend the exact meaning embedded in the genre. Any form of misuse in the discourse marker may result in the ambiguity of the texts. Misusing these items renders the texts

unnatural, affects their acceptability and increases the chances of communication breakdowns Given the shared purpose of language usage, writing texts in different genres may adopt general patterns of linguistic structure; at the same time, as argued by Shaw et al. (2014), the language in writing discourse could be affected by disciplines or purposes and thus would show variations in actual application. Failing to acknowledge these purposes-specificities may prevent writers from achieving their desired target. In order to produce acceptable, natural, and communicatively effective texts, text-producers are required to use discourse markers in a certain way that is expected and accepted by their text-receivers. Therefore, studying the frequency, syntactical pattern and function of discourse markers as elements that ensure the texts' acceptability, naturalness, and effectiveness is indispensable for studying texts in any language. As a novel study is identifying discourse markers in the data and describing the main characteristics that distinguish them as a group. The analysis of the data indicated that discourse markers come from different grammatical word classes.

2.LITERATURE REVIEW

The most frequently studied markers signalling coherence relations are discourse markers. There is a problem in defining the meaning, the functions and the position of discourse markers in a text. This is supported by (Bravo Cladera, 2001).

Fraser (1998) defines discourse markers as connectors which, signals the relationship between the segment introduced known as S2 [segment 2] and relationship with prior discourse segment, known as S1 [segment 1]. In can also be considered as showing a "two-place relation", one sentence lies in the segment that is introduced and the other in the prior segment. The segment can be the sentence or a paragraph. Fraser (1988) defines discourse markers as a class of lexical expressions drawn primarily from the syntactic classes of conjunctions, adverbs, and prepositional phrases.

S2= Segment introduced ⎫signals relationship/functions like a ⎭
two-place relation

S1 = prior segment

represent the canonical form as a sequence <S1- DM+S2>

DM = canonical form, propositional phrase (introduce a separate message), relates two separate messages (a sine quo non of DMs)

(Fraser, 1998)

According to Fraser (1998), there are three conditions of DM;

Condition 1: A DM is a lexical expression (LE), for example, but, so, in addition…

Condition 2: In a sequence of discourse segments S1-S2, a DM must occur as a part of the second discourse segment, S2(initial, medial, final position)

Condition 3: A DM does not contribute to the semantic meaning of the segment but signals a specific semantic relationship which holds between the interpretation of the two Illocutionary Force segments, S1 and S2.

When addressing the problem with positions of DM, according to Fraser (1990), a DM does not necessarily 'introduce' S2 (segment introduced), but it may occur in the medial or final position as well. DMs can also appear sentence medially and sentence finally with functions basically the same as the initially (Schourup, 1999). Fraser (1998) has elaborated in detail on discourse markers especially when using it in the written discourse. Fraser's model on discourse markers is more up to date and practical to be used in this study compared to other scholars.

However, Fraser (1998) also stated that, with the exception of 'though', coordinating and subordinating conjunctions will occupy only the initial position whereas, DMs that are prepositions, prepositional phrases, and adverbials which has a greater latitude syntactically can occur in both final and medial positions. Nevertheless, Kohlani (2010), insist that the initial position can influence the reader's interpretation of the preceding sentence and allows the DMs to have a wider scope over the whole sentence or

paragraph. However, other positions are only responsible for subtle changes in the meaning or function.

3.METHODOLOGY

This research focused on a particular written genre in the form of online newspaper articles. The articles were selected from one particular online newspaper in Malaysia, namely the STAR Online. The purpose of this study was to identify and highlight the discourse markers used in the articles. As a descriptive study, it aimed to provide an in-depth understanding of the use of the discourse markers in online newspaper articles. The present study was placed within the descriptive perspective of the investigation, allowing eloquent and expressive analysis of the data gathered. The research design that was employed in the present study is highly descriptive in nature as it allowed a more precise interpretation of the problem being investigated (Manvender, 2014).

The research design of the present study used a quantitative approach where the discourse markers in the texts of the newspaper articles were quantitatively identified and tabulated, followed by the frequency analysis of the discourse markers used. The frequency analysis was interpreted quantitatively and supported with evidence from the texts. For this research, a systematic probability random sampling method was applied in order to determine the selection of newspaper articles to be analyzed. Random sampling technique

eliminates the possibility of biased examples (Bordens & Abbott, 2008; Din Mohammad, Niami and Fattahipoor, 2018).

The corpus used in this study consisted of 30 news reports (n=30) (Bordens & Abbott, 2008) (from the nation section focusing on Malaysian social and environmental issues downloaded from The Star online newspaper website). The news reports were collected over eight months (March 2016- October 2016). Total word count of the news report was 10, 500 with an average length of 350 words, which was taken as the basis for calculating the frequency of each DM.

The present study drew upon a quantitative interpretation of the data analyzed, by using the CACA approach. Data that was gathered in terms of quantitative identification of the discourse markers was tagged by using the Claws tagger. After that, a frequency analysis of the occurrences was carried out using the AntConc software. This was then tabulated and used to answer the research questions of the study. In-depth exploration of the use of DM was the main aim of the study. For this study, concordance, concordance plot, file view, and clusters/N-grams were used. Frequency analysis of DMs was obtained from the concordance, distributional patterns and positions of-of DMs were retrieved and analyzed by using the concordance plot and the clusters/N-grams.

In this research, the text is treated as a raw data (Bird, Klein, & Loper, 2009) to be used into tools such as AntConc and Claws Tagger that will manipulate and analyze the raw data in a variety of interesting

ways. The following gives a simple explanation of the concordance tools;

a) Concordance – permits us to see words in context. (Bird et al., 2009)

For example, we saw that AND occurred in contexts such as the "___ astonished" the "___ anxiety**.**

b) Lexical dispersion plot(Bird et al., 2009), concordance plot tool (Anthony, 2018)

– This tool shows search results plotted as a 'barcode' format. This allows you to see the position where search results appear in target texts (Anthony, 2018)

- Reveals the positional information about words by showing the *location* of a word in the text: how many words from the beginning it appears. Each stripe represents an instance of a word, and each row represents the entire text. It can also be used to investigate changes in the language use over time (Bird et al., 2009)

c) Frequency distribution

– it shows the frequency of each vocabulary item in the text which is most informative about the topic and genre of the text.

-Just like scientific experiments where a scientist needs to track the changes of animal behaviour, (Bird S. et. al., (2009) frequency distribution helps to find the most frequently used words which would help the researcher to keep track of each vocabulary item (Bird et al., 2009).

-Keeping track of words manually can be a very strenuous task. Thus, AntConc is used to give a frequency distribution analysis. The tabs clusters and word list in AntConc provide a list of all the distinct types of the word used in the text.

e) Clusters

- Scans the entire corpus for 'N' (e.g. 1 word, 2 words, …) length clusters. This allows you to find common expressions in a corpus (Anthony, 2018)

Once the frequency distribution of words is identified, a cumulative frequency plot for these words is produced to shows clearly the occurrences of the word.

4.RESULTS AND DISCUSSION

Corpus tools were used to analyze the discourse markers (DM) used in the newspaper articles. This section will show some of the results that were retrieved and analyzed using the corpus tools.

Frequency analysis

Research question 1: What discourse markers were used in the online versions of newspaper articles published in the STAR newspaper in Malaysia?

This research question shows the types of DMs used in newspaper articles and the frequency of its usage. In the CorONAS corpus, a total of 9 categories of DMs were identified with a 719-total frequency of hits. These discourse markers (DM) were identified based on Fraser (1998) definition which said that discourse markers are conjunctions, adverbs, and prepositional phrases.

Table 1 Concordance Hits of classes of Discourse Markers

CODE	DEC. OF CODE	FREQUENCY OF HITS	PERCENTAGE
CSN	than (conjunction)	5	1%
CSW	whether (conjunction)	5	1%
CSA	as (conjunction)	19	3%
CCB	adversative conjunction	20	3%
RT	quasi-nominal adverb of time	58	8%
CS	subordinating conjunction	81	11%
CST	that (conjunction)	91	13%
RR	general adverb	195	27%
CC	coordinating conjunction	245	34%
		719	100%

Table 4.1 shows the concordance hits, which highlighted the frequency of discourse markers used in the corpus. Based on the analysis, the DM in CorONAS comprised of coordinating conjunction, adversative coordinating conjunction, subordinating conjunction, as (conjunction), that (conjunction), whether (conjunction), than (conjunction), general adverb and quasi-nominal adverb of time. Coordinating conjunction which hits 245 concordances or 34%, tops the list, followed by general adverb which hit 195 concordance or 27%, that –conjunction (CST) with 91 hit or 13% and subordinating conjunction (CS) with 81 hits or 11%. Meanwhile, 'whether' (conjunction) and 'than' (conjunction) had the least concordance hits with 5 times or 1% each. This drew to a conclusion that coordinating conjunction which binds together a piece of writing is highly useful in the CorONAS. Coordinating conjunction (CC) is the most occurring forms which were highlighted in the frequency analysis.

As these DMs are broken down to its respective lexical items, finding shows that the DMs used in the CorONAS most frequently were DMs 'and', 'that', 'also', 'when', and 'but' topped all other DMs followed by 'as', 'then', while', 'or', and 'only' with the lowest number of occurrences. However, the number of occurrences of each of the DMs was considerably varied, especially the DMs 'and', 'that' and 'also' with the rest of the DMs. Based on the findings, DMs 'and' shows the most number of occurrences in CorONAS. This could be due to the common exposure of only certain discourse markers compared to the others among the journalist.

Analyzing frequency and distributional pattern of
discourse markers in the Malaysian newspaper
2414

The figure below shows the frequency of DMs 'and' which was retrieved from the corpus tool AntConc concordance.
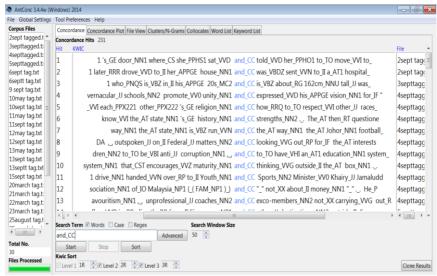

Figure 1. Concordance of DM 'AND'

Based on the data, DM coordinating conjunction (CC) 'and' had the highest frequency occurring 231 times and having 49% of the total percentage of occurrences. The figure below shows the frequency of 'and'.

Distributional patterns of DMs in CorONAS

Research question 2: How is the discourse markers 'and' used in the online versions of newspaper articles published in the online STAR newspaper in Malaysia?

In order to answer this research question, the concordance plot and the cluster/N-grams of the AntConc were used. Here, all the hits for each file were plotted in the form of a 'barcode' indicating the position in the file where the hit occurred. It showed the most favoured positions of DM in the paragraph. The plot provided an easy way to see which files include the DM, and can also be used to identify where the search term hits cluster together. An example of the use of the plot is in determining where specific content words appear in CorONAS.
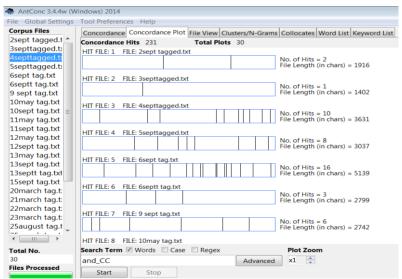
Concordance Plot for DM 'And'

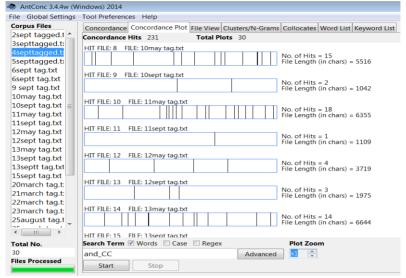Figure 2. Concordance plot for DM 'AND'

Figure 3. Concordance plot for DM 'AND'

Figure 3 shows the distribution of DM 'and' in the files. The findings show that there are 3 positions in the plot, which is at the front, middle and end of the paragraph. Based on the 3 positioning of DM 'and' most of the concordance plot shows that DM 'and' appears to be clustered from the middle to the end of the files for example file 3, 6, 8 and 10.

Besides that, based on the analysis, files 1, 2, 6, 9, 11, 12 and 13 has the least number of occurrences of DM 'and'. The length of file 12 is 3719 characters but with only 4 occurrences of DM, whereas file 3 and 4 has a similar amount of length of characters but more usage of DMs 'and'. Journalist of file 1, 2, 6, 9, 11, 12 and 13 either underused or misused the DMs 'and' in the writing or was able to replace it with

another form of DMs to relate the newspaper report. This shows that the journalist of these newspaper articles did not have a similar style of writing or format when using DMs 'and'.

Clusters/N-grams

Next analysis of the distributional pattern was observed by looking at the clusters of DMs in the news report. The analysis of the distributional patterns of DM was conducted using three placements which is the initial, medial and final.

Discourse markers 'And'

Figure 4 below shows the frequency and POS cluster types preceding DM 'and'.
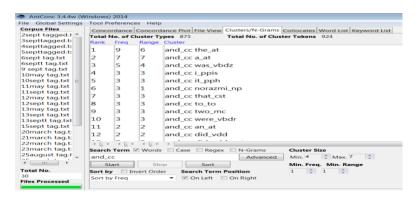

Figure 4. The frequency of POS cluster types preceding DM 'AND'

Figure 4 shows the frequency of DM 'and' cluster types in the left position which precedes the DM 'and'. DM 'and' clusters most with "the_at (article)", "a_at (article)", "was_vbdz" and "I_ppis (1st person sing. subjective personal pronoun)". Examples of the distributional pattern of DM 'and' within these types of clusters can be seen in the pattern 1.

Pattern 1:

1) " The_AT way_NN1 the_AT state_NN1 is_VBZ run_VVN **and_CC** the_AT way_NN1 the_AT Johor_NN1 football_NN1 team_NN1 is_VBZ run_VVN is_VBZ the_AT same_DA….

(4septtagged.txt)

Distributional pattern of DM: and_CC

VVN   CC   AT

The analysis:

In this pattern, the prior word to 'and_CC (coordinating conjunction)' is 'run_VVN(past participle of lexical verb)' and proceed with 'the_AT(article). 'And' is in the medial position of the sentence whereby 'And' (coordinating conjunction) links past participle of the lexical verb(VVN) and article(AT).

Pattern 2:

> 2)(a) Lee_NP1 **and_CC** Ken_NP1 sought_VVD
> department_NN1 head_NN1 Datuk_NP1 Seri_NP1
> Michael_NP1 Chongs_NP1 help_NN1….
>
> (b)… on_II the_AT issue_NN1….**and_CC** the_AT 28-year-
> old_JJ son_NN1 ,_, once_RR21 again_RR22 begged_VVD…
> (6sept tag.txt)
>
> **Distributional pattern of DM: and_CC**
>
>   NP1   CC   NP1
>
>
>   **…**Lee **and** Ken…
>
>   NN1   CC   AT

The analysis:

In 2(a) the prior word to 'and' is Lee_NP1 (singular proper noun) and proceed with Ken_NP1 (singular proper noun). 'And' is in the medial position of the sentence whereby 'and (coordinating conjunction)' links NP1 (singular proper noun) and NP1 (singular proper noun). In this sentence, it can be clearly seen that the coordinating conjunction 'and' links words from the same word class which is the singular proper noun).

In 2(b) the prior word to 'and_CC (coordinating conjunction)' is 'issue_NN1(singular common noun)' and proceed with 'the_AT(article). 'And' is in the medial position of the sentence whereby 'and (coordinating conjunction)' links NN1 (singular common noun) and AT(article).

Pattern 3:

<div style="border:1px solid">

3)…budget_NN1 to_TO counter_VVI flash_JJ floods_NN2 had_VHD been_VBN increased_VVN to_II RM80mil_FO **and_CC** the_AT relevant_JJ agencies_NN2 were_VBDR fully_RR prepared_VVN…

(11may
tag.txt)

Distributional pattern of DM: and_CC

FO   CC   AT

</div>

The analysis:

In this pattern, the prior word to 'and_CC (coordinating conjunction)' is 'RM80mil_FO (formula)' and proceed with 'the_AT(article)'. 'And' is in the medial position of the sentence whereby, 'and' links FO (formula) and AT(article).

Pattern 4:

> 4)…the_AT victim_NN1 later_RRR drove_VVD to_II her_APPGE house_NN1 **and_CC** was_VBDZ sent_VVN to_II a_AT1 hospital_NN1 here_RL by_II her_APPGE (2sept tagged.txt)   housemate_NN1 ._.
>
> **Distributional pattern of DM: and_CC**
>
> NN1   CC   VBDZ

The analysis:

In this pattern, the prior word to 'and_CC (coordinating conjunction)' is 'house_NN1(singular common noun)' and proceed with 'was_VBDZ'. 'And' is in the medial position of the sentence whereby, 'and' links NN1(singular common noun) and first and third person singular of the past tense (VBDZ).

Pattern 5:

> 5)… the_AT man_NN1 paid_VVD no_AT heed_NN1 to_II them_PPHO2 **and_CC** was_VBDZ quite_RG content_JJ on_II continuing_VVG his_APPGE walk_NN1 along_II the_AT (11sept tag.txt)   road_NN1 ._.
>
> **Distributional pattern of DM: and_CC**
> PPHO2   CC   VBDZ

The analysis:

In this pattern the prior word to 'and_CC (coordinating conjunction)' is 'them_PPHO2)' and proceed with 'was_VBDZ'. 'And' is in the medial position of the sentence whereby, 'and' links 3rd person plural objective personal pronoun(PPHO2) and the first and third person singular of the past tense(VBDZ).

5. CONCLUSION

One important contribution of this study is identifying discourse markers in the data and describing the main characteristics that distinguish them as a group. The analysis of the data indicated that discourse markers come from different grammatical word classes. These word classes include: And, Or (Coordinating Conjunction- CC); But (adversative coordinating conjunction - CCB); When, While (subordinating conjunction -CS); That (Conjunction - CST); As (Conjunction - CSA); whether (conjunction - CSW); than (conjunction - CSN);Then (quasi-nominal adverb of time - RT); Also, Only (general adverb -RR). However, they can be distinguished based on their preference to occur in various positions in the sentence and the whole file, whether structurally or semantically, to the sentences they introduce and their connections with the pragmatic force (functional purpose) of the text rather than the grammatical structure of the sentence. This contributes to the various terminology used to identify DMs, as the main issue of this research. The numerous word classes showed various interpretations of the text, and that their function went well beyond the mere signalling of structural relations.

The frequency and types of DMs revealed the number of occurrences of DM in the online newspaper genre. Overuse or underuse of these DMs may render the text unnatural. Different genre, text-type, and individual style of the authors are factors that may affect the choice and frequency of discourse markers in any text (Dooley and Levinsohn 2001). If they are not controlled, these factors may pose problems for the findings of the analysis of any study of discourse markers. Thus, this result can contribute to the frequency of DMs that should be applied by authors, teachers, and learners.

The second contribution is the distributional patterns of the sentences. The clustering of DM in each file (front, middle, end), the positions of DM (initial, medial, final) and the way it linked with other POS (part-of-speech) are derived from looking at the distributional patterns. It is important to look at the concordance plot and the distributional pattern of the sentences. When there is a logical arrangement in the POS it indicated a well-organized sentence structure. If the author used it with another word, it may affect the coherence or cause ambiguity to the meaning of the text.

The position of the discourse markers revealed the type of sentence that the authors are prone to produce in the articles. The authors had used mainly simple and compound sentences in their writing. This showed their cognitive level when it comes to writing articles in the English language. Children mostly use simple sentences to relate their thoughts. In these articles, the authors have merely linked two phrases and simple sentence clauses making it a compound

sentence, but with the level which sounds like the ability of a child. Being a non-native speaker (NNS), the influence of the author's first language (L1) can be seen in writing these types of sentences with discourse markers linking the sentences. Even though it is a report writing, readers need to feel the impact of the story with various types of writing styles. Besides that, the position of the discourse markers influences the reader's interpretation of the meaning (Suranto & Soetjipto, 2019).

One of the most noticeable features that distinguished discourse markers in CorONAS are their preference to occur sentence-medially. The frequent employment of DMs in sentence-medial position by the authors contradicts Schourup (1999) and Kohlani (2010) who argue that DMs tend typically to occur in sentence-initial position. Some of these expressions, the analysis indicated, may occasionally appear in other segments in the sentence. However, the medial position of the sentence is the place where these items are most likely to be found. In this study, all 3 positions of discourse markers that occurred sentence-initially, medially and end were taken into consideration. By identifying the items that are included in this group, future research can be carried out to examine how the occurrence of these items' functions in other genres.

Occurring in the sentence-medial position, therefore, is a strong indication of the discourse marker status of the members of this group. Moreover, in CorONAS, not only do discourse markers favour the

sentence medial position, but they also have a strong tendency to cluster at this position as well as the end position in a text. However, previous studies lack structural analysis such as the clusters analysis. They only focused on the positions of DM (Schourup (1999) and Kohlani (2010). In the present study, the cluster analysis contributes to the understanding of an in-depth analysis of DM whereby by looking at the clusters the types of sentences used by authors can be identified.

In order to produce acceptable, natural, and communicatively effective texts, text-producers are required to use discourse markers in a certain way that is expected and accepted by their text receivers. However, the choice, frequency, and distribution of discourse markers that render texts acceptable and natural differences among genres. This means that discourse markers are governed by the pragmatic norms of the language in which they operate.

REFERENCES

ANTHONY, L. 2018. **AntConc.** Tokyo: Waseda University. In.Japan

AL KOHLANI, F. A. 2010. **The function of discourse markers in Arabic newspaper opinion articles**. Doctoral dissertation, Georgetown University.USA.

BIRD, S., KLEIN, E., & LOPER, E. 2009. **Natural language processing with Python: analyzing text with the natural language toolkit**: " O'Reilly Media, Inc.".USA.

BRAVO CLADERA, N. 2001. **Un estudio de los marcadores del discurso y su distribución en interacciones diádicas de jóvenes bilingües**. Actas del XXIII Congreso Internacional de Lingüística y Filología Románica / coord. Vol. 1, pp. 133-146.Spain.

BHATIA, V. K. 2014. **Analysing genre: Language use in professional settings**. Routledge.UK.

BORDENS, K. S., & ABBOTT, B. B. 2008. **Research methods and design: A process approach**.Boston,MA:McGraw-Hill.USA.

CRESWELL, J. W. 2014. **A concise introduction to mixed methods research**. Sage Publications.USA.

DOOLEY, R. A., & LEVINSOHN, S. H. 2001. **Analyzing discourse: A manual of basic concepts**. SIL International.USA.

FRASER, B.1988. **Types of English discourse markers**. Acta Linguistica Hungarica, Vol. 38 N° 4, pp. 19-33.Hungary.

FRASER, B. 1990. **An approach to discourse markers**. Journal of pragmatics, Vol. 14 N° 3, pp. 383-398.Netherlands.

FRASER, B.1998. **Contrastive discourse markers in English**. Pragmatic and beyond new Series, Vol. 1, pp. 301-326. Netherlands.

GRABE, W., & KAPLAN, R. B. 1996. **Theory and practice of writing: An applied linguistics perspective**. Harlow: Pearson Education.UK.

JUCKER, A. H., & ZIV, Y. 1998. **Discourse markers: Descriptions and theory**. John Benjamins Publishing. Netherlands.

MANVENDER KAUR SARJIT SINGH. 2014. **A Corpus-Based Genre Analysis of the Quality, Health, Safety and Environment Work Procedures in Malaysian Petroleum Industries**. Unpublished PhD Thesis: Universiti Teknologi Malaysia.

MANVENDER, K., YASMIN, H. Z., & SHAMSUDIN, S. 2012. **A computer-assisted corpus analysis (CACA) of professional discourse**. Sino-US English Teaching, Vol.9, N° 6. pp. 1236-1245.USA.

DIN MOHAMMAD, M. NIAMI, M. AND FATTAHIPOOR, M. 2018. **The Impact of Applying Lexical Equivalent Corpus on Iranian**. Iranian Journal of Social Sciences and Humanities Research, Vol. 6, pp. 1-12. Iran.

SCHOURUP, L. 1999. **Discourse markers**. Lingua, Vol. 107. pp. 227-265. Netherlands.

SHAW, V. 2014. **The short story: a critical introduction**. Routledge.UK.

SHUKOR, S. S. 2014. **Effectiveness of Face-to-face and Facebook Collaborative Writing in Malaysian ESL Undergraduates'** Writing Performance (Doctoral dissertation, Universiti Putra Malaysia).

SURANTO, S., & SOETJIPTO, B. 2019. **Conceptual model on entrepreneurial intention in higher education.** Humanities & Social Sciences Reviews. Vol. 7, N° 3: 17-24. India.

UNIVERSIDAD
DEL ZULIA