A Study on the Features of Adversative Conjunctions in English Theses of Chinese EFL Postgraduates

A Corpus-Based Analysis and its Implications

Shihan Wang

School of Foreign Languages & Literature, Shandong University, Jinan, Shandong, China

wangshihan0314@outlook.com

Abstract. English as a Foreign Language (EFL) learners frequently encounter issues concerning the effective acquisition and deployment of parts of speech during their second language learning process. This study aims to investigate the frequency, distribution, and semantic features of adversative conjunction (AC) usage among Chinese EFL postgraduates (Master's and Doctoral candidates) through a combination of informational and quantitative analyses of self-established corpora. It further delves into the fundamental causes that might contribute to the prevalence of errors in the use of such conjunctions by these learners and explores the association between their proper deployment and the quality of students' writing. Ultimately, the study offers actionable insights that can inform pedagogical practices in English language instruction.

Keywords: adversative conjunctions; thesis; academic discourse; corpus-based analysis; EFL; cohesion

1 Introduction

In the process of second language acquisition, non-native English learners often encounter the challenges about parts of speech. Conjunctions classified among the eight major parts of speech in traditional grammar have received relatively less scrutiny compared to other. While the significance of conjunctions cannot be overstated, their improper usage has become conspicuous in the context of Chinese EFL learning. This research focuses on conjunctions.

In academic English writing, cohesion referring to the interrelation within a text is vital for establishing coherence, comprehensibility, and persuasiveness. It also mirrors the writer's capacity to articulate complex arguments in a lucid and structured way. The cultivation of cohesion skills in English writing is imperative for conveying thoughts effectively. And the proper usage of conjunctions directly impacts the achievement of cohesion.

For EFL learners, essay writing remains a daunting task, with academic writing presenting an even greater challenge. Although numerous studies have reported that Chinese EFL learners commit fewer grammatical errors, their written work does not

always demonstrate high quality. This study contemplates whether the usage of conjunctions, impacts the overall quality of writing. To this end, it meticulously examines the distinctive features of conjunction use in the English Theses of Chinese EFL postgraduates, encompassing aspects such as frequency, distribution, and semantic attributes. Furthermore, the research findings will provide instructive perspectives for English (writing) pedagogy.

2   Literature Review

Every discerning text inclines towards conveying a consistent theme or argument; this implies that a text exhibits a quality of unity and consistency throughout its discourse [6]. It is imperative that the components within a sentence or linking separate sentences, are bound together by specific linguistic ties. Cohesion has consistently remained a focal point of extensive scholarly attention within this particular domain. The notion of cohesion embodies a semantic dimension, referring to the meaningful connections that reside within the fabric of a text [5]. In the framework posited by Halliday and Hasan (1976), cohesion has been categorized into reference, substitution, ellipsis, conjunction, and lexical cohesion.

Connection are regarded as a key mechanism in achieving textual cohesion, as they serve to concatenate sentences and paragraphs within different textual constructs [3].With the increasing scholarly attention towards conjunctions, they have been variously classified, including scholars but not limited to Halliday & Hasan (1976), Quirk et al. (1985), Biber et al. (2000), and Martin and Rose (2003, p. 119). Each of these scholars has contributed influential categorizations. However, this research primarily employs the framework provided by Halliday and Hasan (1976) due to its perceived comprehensiveness. They classify conjunction into additive conjunctions, adversative conjunctions, causal conjunctions and temporal conjunctions. The core of this research lies in ACs.

Halliday and Hasan posited that, ACs serve to denote contrast and comparison signifying element that is "contrary to expectation" [5]. In conducting the semantic analysis of ACs, reference was made to Yuan Shan's categorization of the semantic dimensions of ACs into adversative relations and concession relations [10]. This classification schema was chosen due to the fact that this research primarily focuses on overarching features rather than delving deeply into the usage patterns of individual lexical items.

Upon meticulous scrutiny of both domestic and international scholarly literature, the majority of the studies rely on written compositions such as informal writings, or preexisting corpora as empirical bases. Despite the abundance of research contributed by scholars globally, investigations situated specifically within academic discourse contexts are relatively scarce. In recent years, a multitude of scholars have directed their attention to the study of academic discourse. And logical connectives, including ACs, have become integral components of writing techniques because of scholars' expression needs. Against this backdrop, the present study endeavors to examine the use of ACs in English Theses of Chinese EFL postgraduates.

3   Research Questions

(1)   What are the frequency features and distributional patterns of adversative conjunctions (ACs) in English Theses of Chinese EFL Postgraduates?

(2)   Among the 12 high-frequency ACs, what are the features of their semantic application by Chinese EFL Postgraduates?

4   Methodology

This study constitutes an investigation into the distinctive features of ACs usage among Chinese EFL postgraduates within the context of their academic theses. The research is underpinned by self-established corpus and employs both informational analysis and quantitative analysis as dual methods of inquiry. The following provides a concise overview of the corpora, the tools employed, and the procedures throughout the research process.

4.1   Corpora

This study is grounded in two self-established corpora: the first comprises 100 MA theses authored by Chinese EFL postgraduates from 12 different universities, whereas the second corpus consists of 22 PhD theses sourced from 6 distinct universities. The former corpus was quantified to contain 401,776 running words, while the latter corpus comprised 398,174 running words. The principal motivation for selecting theses from two distinct academic stages is to illuminate the differential patterns of cognitive processing and linguistic expression between EFL learners at different academic stages.

4.2   Tools

In the course of data extraction and analysis for this study, AntConc version 4.0.4 will serve as the primary software tool to facilitate corpus interrogation. The process entails importing plain text format into AntConc; accordingly, the plain text format of the 122 articles included in this research will be ingested into the application to perform exhaustive searches pertinent to the research questions. Concurrently, Microsoft Excel software will be utilized for statistical data tabulation and computation. Moreover, chi-square goodness of fit tests will be conducted to undertake rigorous calculations aimed at examining the associations among categorical variables.

4.3   Research Procedures

Fundamentally, this study adopts the theoretical framework posited by Halliday and Hasan to systematically categorize conjunctions, listing 12 high-frequency ACs as the focal point of inquiry. The subsequent research procedures through the following stages: Initially, it involves two self-established corpora composed of academic theses

written by Chinese EFL postgraduates. Subsequently, AntConc version 4.0.4 is employed as the principal corpus analysis software, and the retrieved data are subjected to statistical compilation and computation using Microsoft Excel, alongside the application of chi-square goodness of fit tests to ascertain whether there exist statistically significant differences among the preliminary findings. Through this progressive methodology, the research questions will be explicated and addressed.

5 Results and Discussion

This segment of the study primarily utilizes informational analysis and quantitative analysis to attain research objectives, and discusses pertinent findings relative to the results.

5.1 Data Analysis

The initial stage of the study necessitates addressing the frequency patterns of ACs across both corpora. Below, a table presents the 12 high-frequency ACs that receive focused scrutiny in this study. In the course of search using AntConc, this study employs wildcard characters, specifically the asterisk '*', to retrieve text segments containing the targeted string. The search terms utilized forms like '*but*' as depicted in Table 1. This strategy ensures that instances of these ACs are captured regardless of any irregularities in their immediate contextual spacing. In order to render the statistical data more comparable, the table showcases dual metrics—raw counts and the percentages of these frequencies relative to the total running words.

Table 1. Frequency of ACs in MA Thesis corpus and PhD Thesis corpus

	MA Thesis corpus	PhD Thesis corpus
	Raw counts (Percentage)	*Raw counts (Percentage)*
but	5907 (1.47%)	3995 (1%)
while	1836 (0.46%)	1299 (0.33%)
however	1352 (0.34%)	1186 (0.3%)
yet	588 (0.15%)	469 (0.12%)
although	540 (0.13%)	493 (0.12%)
instead	508 (0.13%)	384 (0.1%)
rather	496 (0.12%)	577 (0.14%)
though	395 (0.10%)	1730 (0.43%)
whereas	144 (0.04%)	177 (0.04%)
nevertheless	92 (0.02%)	137 (0.03%)
anyhow	32 (0.008%)	6 (0.002%)
nonetheless	20 (0.005%)	16 (0.004%)
Total	11910 (2.96%)	10469 (2.63%)

As Table 1 indicates, there is a discernible disparity in the frequency of ACs across the two corpora. The overall rate of occurrence in the MA thesis corpus stands at approximately 2.96%, while in the PhD thesis corpus, it amounts to about 2.63%. This constitutes a difference of 0.33 percentage points, which aligns with similar findings from prior research, suggesting that MA students tend to use ACs more frequently than their PhD counterparts in China.

Through corpus analysis, it is evident that in the MA thesis corpus, the three most frequently used ACs are 'but', 'while', and 'however', with respective frequency percentages of 1.47%, 0.46%, and 0.34%, cumulatively accounting for approximately 76.69% of the overall usage frequency of such conjunctions, which totals to 2.96%. In contrast, within the PhD thesis corpus, these same conjunctions register at frequency percentages of 1%, 0.33%, and 0.3%, adding up to a combined total of 1.63%, which constitutes around 61.98% of the overall frequency of AC use in this corpus. The usage frequencies of 'but', 'while', and 'however' is notably considerable within both corpora, thus indicating a shared feature in their distributional patterns. However, in conducting the corpus search, results revealed that MA thesis corpus infrequently employed alternative ACs such as 'on the contrary', 'on the other hand', and 'by contrast', whereas the PhD thesis corpus displayed a more appropriate usage of these particular conjunctions. Notably, Table 1 illustrates that in the PhD thesis corpus, the conjunction 'though' is relatively high in frequency, ranking second among the ACs used in this data-set. These data suggest that the MA thesis corpus exhibits a relative paucity in the variety of ACs employed, indicating that Chinese MA students tend to concentrate their usage on a subset of highly frequent conjunctions. Conversely, PhD students exhibit a broader and more diverse selection of ACs in their academic writing.

Furthermore, this study delves into the semantic features of the 12 typical ACs across both corpora. In Yuan Shan's research, it is highlighted that the semantic categories of ACs primarily encompass adversative relations and concession relations [10]. This study randomly selects three MA theses and three PhD theses from each corpus, respectively, and analyses the semantic types of 12 high-frequency ACs. The chi-square goodness-of-fit test is then employed to assess whether there exists a statistically significant difference in the distribution of these semantic types between the MA and PhD thesis corpora.

Table 2. Semantic distribution of ACs in MA Thesis corpus and PhD Thesis corpus

Semantic Typology	MA Thesis corpus	PhD Thesis corpus	P (Compared with 0.05)
	Raw counts (Percentage)	Raw counts (Percentage)	
adversative relation	289 (66.13%)	946 (59.95%)	0.018852167 (< 0.05)
concessive relation	148 (33.87%)	632 (40.05%)	0.021850752 (< 0.05)

As illustrated in Table 2, within the randomly sampled graduate papers, the proportion of adversative relation usage for ACs in the MA thesis corpus stands at 66.13%,

while that of concessive relation is at 33.87%. Conversely, in the PhD thesis corpus, the ratio for adversative relation is 59.95%, and the concessive relation constitutes 40.05% of the total usage. A striking observation here is that Chinese EFL postgraduates exhibit a preference for employing the adversative relation of such conjunctions, exemplified by but and however, rather than their concessive relation, as seen in conjunctions like though and anyhow. This tendency indirectly suggests an uneven distribution of the types of ACs employed in the academic writings of Chinese EFL postgraduates. This section is intimately related to the outcomes presented in Table 1, collectively underscoring the differential patterns of usage in terms of the semantic functions of ACs among Chinese EFL postgraduates' theses.

Subsequently, the study employs the chi-square goodness of fit test within the framework of the chi-square (χ^2) statistical test to ascertain whether there exists a significant discrepancy between the two corpora. According to standard practice, when the calculated p-value is less than 0.05, it is generally considered that a substantial difference exists. Upon conducting this analysis, the research yields p-values of 0.018852167 and 0.018852167 for the respective MA and PhD thesis corpora. The p-values indicate that there exist statistically significant differences in the semantic usage patterns of ACs between Chinese master's and doctoral graduate students in their academic writing. Specifically, doctoral students demonstrate a more balanced distribution of usage across both concessive relation and adversative relation than master's students. In other words, doctoral candidates do not incline as strongly towards the employment of the adversative relation of these conjunctions as master's candidates do.

5.2   Discussion

Numerous previous scholars have posited that cohesion poses a significant challenge for EFL learners. Shirazi and Nadoushani specifically contended that cohesion is highly demanding due to its reliance on advanced knowledge in grammar, semantics, and pragmatics of the target language [8]. This study's findings demonstrate that even advanced-level Chinese EFL learners exhibit misunderstandings and errors in both the comprehension and application of cohesive devices, particularly with regard to the use of adversative conjunctions (ACs). The fundamental reasons underpinning such phenomena can be distilled into two key aspects:

The qualitative shift in learners' capacities and linguistic repertoire when transitioning from first language (L1) acquisition to second language (L2) learning. While L2 learners already possess proficiency in at least one language, this background may also engender negative transfer effects. Notably, L2 learners leverage their existing linguistic cognition and metalinguistic awareness about how languages function, leading them to employ ACs in their L2 based on the syntactic and semantic patterns of their L1. Consequently, Chinese EFL postgraduates often resort to direct expressions of logical relations through these conjunctions rather than adopting alternative forms of cohesive linkage in their writings. Amayreh and Abdullah have previously highlighted that English as a Second Language (ESL) learners appear to manifest deficits insofar as their compositions exhibit reduced lexical diversity or heightened concep-

tual redundancy, contrasting with the richer vocabulary repertoire displayed by native speakers [2].

The stark contrast between the learning environments for L1 and L2 acquisition. L2 learning effect are significantly influenced by the prevailing conditions and environments in which learning takes place. Chinese EFL postgraduates, operating in different cultural contexts and high-pressure academic environments, display a more mechanical employment of lexical items to meet their writing needs.

Despite being classified as advanced EFL learners, Master's and Doctoral students alike demonstrate persistent errors in using ACs due to these influences. However, results indicate that doctoral students exhibit relatively stronger abilities compared to master's students in the appropriate use of such conjunctions. Xu et al. interpret this as lower-level L2 learners predominantly allocating resources and effort towards processing lexical and syntactic information [9].

An increasing body of scholarly attention is directed towards the potential correlation or intrinsic link between the use of conjunctions and writing quality. The absence of conjunctions in text leads to longer processing times and poorer text comprehension [1]. And in academic discourse, conjunctions are instrumental in facilitating textual parsimony and intelligibility [4]. For this study, the strategic use of ACs facilitates easier interpretation of authors' viewpoints and stances. Nevertheless, the challenge of deploying these conjunctions appropriately to ensure coherent text progression remains a topic of ongoing scholarly interest. Drawing on Zone of Proximal Development (ZPD) theory of Lev Semenovich Vygotsky's, social interaction stimulates cognitive growth [7]. Educators must recognize that fostering interaction and guided collaboration is crucial for propelling EFL learners to the next level of competence in mastering the nuanced use of cohesive devices.

6   Conclusion

This study employs self-established corpora and adopts informational analysis and quantitative methods to examine the features of AC usage in theses authored by Chinese EFL postgraduates. This section presents a concise summary of the research findings, alongside the study's limitations and implications for further inquiry.

6.1   Main Findings

Firstly, the statistical analyses reveal that Master's students employ ACs at a higher frequency in their theses compared to Doctoral students. And Master's students exhibit a narrower distribution in their choice of ACs, indicating a greater reliance on a smaller subset of these connectors. For instance, the data suggests overuse of conjunctions such as 'but', 'while', and 'however' by Master's students. This study tentatively concludes that as Chinese EFL learners progress to higher academic levels, their utilization of ACs tends to become more judicious, showing reduced incidence of misuse and overuse, along with a better command and deployment of a broader array of such conjunctions.

Secondly, the study investigates the semantic attributes of ACs across the two corpora. Chi-square test results manifest clear disparities between Master's and Doctoral students in their semantic use. Although both groups predominantly exploit their adversative relation, closer scrutiny of the data discloses that Doctoral students exhibit a superior capacity to utilize the concessive relation of these conjunctions, indicative of slightly heightened semantic mastery compared to the Master's counterparts.

Additionally, by examining the differences between L1 learning and L2 learning, the study explores learner characteristics and learning environments as potential sources of misuse of ACs by EFL learners. Simultaneously, the discussion underscores the intimate connection between the use of ACs and the quality of students' writing. The critical pedagogical issue of guiding EFL learners to appropriately use a diverse range of ACs to enhance students' writing quality continues to merit sustained scholarly attention.

6.2   Limitations

Regarding the research methodology, this study did not adopt qualitative analytical approaches to address the research inquiries. Nonetheless, qualitatively investigating the patterns and developmental features of AC usage among Chinese EFL postgraduates would constitute a valuable research endeavor. With respect to the research sample, the present study employed a random collection process without stringent control over the institutional affiliations of the Chinese EFL postgraduates involved. Despite this, the inclusion of thesis samples from students across different tiers of institutions potentially enhances the universality of the findings. Concerning the semantic categorization of ACs, this study relied on a relatively straightforward classification scheme established in prior research. While this categorization is suitable for studies that deal with statistical data, employing a more refined taxonomy, such as those developed by Halliday and Hasan or Hyland, would render the research outcomes more precise and rigorous.

Moreover, future research could extend beyond purely analyses of ACs themselves and instead focus on exploring whether Chinese EFL learners experience anxiety in L2 writing tasks that might be intrinsically linked to the difficulty in mastering the use of such conjunctions. Additionally, educators should contemplate strategies to tackle issues related to part of speech usage in L2 writing contexts. Such research endeavors could shed light on pedagogical practices that effectively address these challenges.

6.3   Implications

The findings of this study affirm the pivotal role of ACs in the writings of EFL learners, revealing a strong correlation between their appropriate usage and overall writing quality. Based on these results, the following recommendations are put forth for consideration in English language instruction:

With respect to curriculum content, the judicious employment of ACs in students' compositions is demonstrably tied to the enhancement of text quality. Although English evolves continually, adherence to rules and established patterns in the usage of

parts of speech, including conjunctions, remains essential for achieving optimal oral and written expression. Therefore, mastering the fundamental linguistic definitions associated with ACs and their practical applications is indispensable in teaching contexts. Moreover, discourse competence represents a multifaceted skill that many ESL and EFL learners grapple with, a fact increasingly acknowledged and addressed in specialized ESP (English for Specific Purposes) courses. These considerations necessitate the integration of foundational linguistic courses and grammar instruction into English language curricula to serve the purpose of imparting core theoretical knowledge to students.

Regarding teaching methodologies, the analysis of texts and discourse is indispensable. The adoption of case-based teaching approaches that involve corpus analysis is highly encouraged, as it strengthens students' ability to analyze discourse and provides firsthand insight into the practical application of ACs within texts. This innovative departure from traditional pedagogical methods exerts considerable influence on enlightening learners about the correct usage of such conjunctions. Thus, incorporating corpus-driven exercises into the teaching practice serves as an effective means to foster a deeper understanding and more proficient use of ACs among EFL learners.

References

1. Al Shamalat, R. Y. S., & Ghani, C. A. B. A. (2020). The Effect of Using Conjunction as Cohesive Device on the Undergraduates' Quality of Writing in Argumentative Essays of Jordanian EFL Learners. Arab World English Journal.
2. Amayreh, K. S., & Abdullah, A. T. (2021). Error Analysis of Conjunction in Expository Essay Writing by Jordanian Undergraduate Students Studying English as a Foreign Language (Efl). International Journal Of Education, Psychology And Counselling (IJEPC), 482-491.
3. Amayreh, K. S. A., & Bin Abdullah, A. T. H. (2022). Conjunction in Expository Essay Writing by Jordanian Undergraduate Students Studying English as a Foreign Language (EFL). International Research Journal on Advanced Science Hub, 4(02), 24-30.
4. Daud, A., Ajam, A., & Jusnita, N. (2023). Students' grammatical cohesion in essay writing. Langua: Journal of Linguistics, Literature, and Language Education, 6(1), 23-34.
5. Halliday, M. A. K., & Hasan, R. (1976). Cohesion in English. London, England: Longman.
6. Jassim, A. H. (2023). The role of cohesion in text creation. Language, Discourse & Society, 11(1), 159-170.
7. Sage, K. (2022). Zone of Proximal Development. Routledge. https://doi.org/10.4324/9780367198459-REPRW163-1
8. Shirazi, M. A., & Mousavi Nadoushani, S. M. (2017). The locus of adversative conjunctions in the research articles: have they niched or vanished?. Sage Open, 7(1), 2158244017700946.
9. Xu, X., Pan, M., Dai, H., Zhang, H., & Lu, Y. (2019). How referential uncertainty is modulated by conjunctions: ERP evidence from advanced Chinese–English L2 learners and English L1 speakers. Second Language Research, 35(2), 195-224.
10. Yuan, S. (2015). Developmental features of contrastive connections in Chinese English learner's academic writing. Master's thesis of Henan Normal University.