Types and Features of Noun Phrase in Chinese Scholars' Abstracts

Li Wang[1] & Fei Pei[1]

[1] School of Foreign Languages, Nantong University, Nantong, Jiangsu Province, China

Correspondence: Li Wang, School of Foreign Languages, Nantong University, Nantong, 226019, Jiangsu Province, China. E-mail: wanglily22@ntu.edu.cn

Abstract

Noun phrases, as the basic components of sentences, carry large amounts of information. Based on corpus-based research method, this study aims to explore the use of nouns in the journal abstracts written by Chinese scholars. Statistically significant difference was found in the noun *effect* between Chinese scholars' dissertation abstracts (CSDA) and English-speaking scholars' dissertation abstracts (ESDA), so *effect* was chosen as an example word throughout the research. The results show that (1) Chinese scholars tend to use more simple noun phrases while international journal scholars are inclined to use complex noun phrases in their articles. (2) As for the use of the colligation *adjective+effect*, Chinese scholars are likely to use synonyms or to replace the more appropriate adjectives, which cause non-native expressions. (3) As for the colligation of *effect+preposition*, *in* is most frequently used by Chinese scholars, but seem to be untypical to the international journal scholars. The study found that interlingual transfer (mother tongue transfer) and intralingual transfer appear to be the main causes of these discrepancies.

Keywords: types and features, noun phrase, Chinese scholars, *effect*, corpus-based

1. Introduction

A range of studies has indicated that noun is a dominant part of speech, and the semantic content of sentences is borne mostly by nouns (e.g., Algeo, 2006). Therefore, as phrases based on nouns, noun phrases (NPs) are essential components of sentences. The importance of the NPs in English was acknowledged by generations of language researchers, and a large number of them paid extensive attention to noun-phrase types and features. For instance, in the most important two publications of corpus-based grammars, the Comprehensive Grammar of English (Quirk et al., 1985) and Longman Grammar of Spoken and Written English (Biber et al., 2000), numerous surveys on the structures of NPs were carried out and tremendous useful conclusions were drawn which enlightened the latter language researchers profoundly. This paper focuses specifically on the use of types and features of noun phrase (NPs) for Chinese scholars. So far, however, few of the researchers have focused on NP in the aspect of scholars' dissertation abstracts with the corpus linguistic method and few of them have done the research by comparing Chinese scholars with international scholars. Thus, this paper attempts to find types and features of NP in the aspect of Chinese scholars' academic writing abstracts through a corpus-based contrastive approach, in order to probe into the underlying reasons of those distinctions.

Since the classifications of Biber et al. (2000) and Quirk et al. (1985) are widely accepted, this paper would classify NP structures on the basis of Biber's view with sentences extracted from CSDA and ESDA as follows.

Ⅰ (Determiner) + Head (Simple NP)

1. Zero + Noun

…but it does not take context into full consideration… (CSDA)

…languages can influence word recognition in EFL learners… (ESDA)

2. Determiner + Head

…a schema is an abstraction of the similarities of all the members… (CSDA)

…as important as the deletion of the plural marker… (CSDA)

…this finding is interpreted in the context of… (ESDA)

Ⅱ Premodifier + Head (Complex NP)

1. Adjective + Head

…the most grammatically accessible antecedent is… (CSDA)

…a Coh-Metrix-based study of the linguistic features that… (CSDA)

…researching the longitudinal development of second language learners… (ESDA)

2. Noun + Head

…research on language assessment literacy… (CSDA)

…the approach to descriptor calibration… (CSDA)

…second language acquisition research… (ESDA)

Ⅲ Head+Postmodifier (Complex NP)

1. Head + To-infinite Clause

…their contextual adaptability to interpersonal relationships… (CSDA)

…Chinese applied linguists' attention to research… (CSDA)

…evidence of advancement to a higher stage… (ESDA)

2. Head + Preposition or Subordinating Conjunction (of, for, from, with)

… the number of idea units of the story retelling… (CSDA)

… used as evidence for Chinese as a discourse-oriented language… (CSDA)

… task-essential practice with interpreting agent/patient roles… (ESDA)

3. Head + Ing-clause

… impaired memory arising from an instruction… (CSDA)

… we can study and mind using the same method… (CSDA)

… extract statistical regularities from their environment allowing them to extract words… (ESDA)

4. Head + Ed-clause

… the rating scales of the rubric used in this study… (CSDA)

… a tentative intepretation based on the functional roles… (ESDA)

5. Head + Relative Clause (which, that, who, whom, whose, when, where)

… interlanguage development which is more universal than… (CSDA)

… EFL listening comprehension that has been conducted in China… (CSDA)

… low-anxiety learners who produced high levels of modified output… (ESDA)

… CF did not result in simplified writing when structural complexity… (ESDA)

Two research questions arise at this point:

(i). Is there any difference in using NP between Chinese scholars and international scholars?

(ii). What are the major cause of the above differences?

2. Language Transfer Influence in Studies on Second Language Acquisition

The role of linguistic transfer, or cross-linguistic influence in second language acquisition, has been a field of extensive research in the past few decades (Ellis, 1994; Gas & Selinker, 1994; Kellerman, 1995; Larsen-Freeman & Long, 1991; Odlin, 1989; Selinker, 1992). The language transfer perspective on language has shed light on a wide range of areas. According to Odlin (1989), "transfer is the influence resulting from similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired". It is certain that language transfer does exist indeed in almost all aspects of second language learning, including phonology, morphology, lexis, semantics, syntax, pragmatics, etc. The motivation is to explore Odlin's claim, among others, that there is a link between mother tongue and EFL learners' use of NP. Specifically, the author is interested in testing this claim in the context of writers in different educational and cultural settings. This motivation is additionally based on previous research that has found that "There is now overwhelming evidence that language transfer is indeed a real and central phenomenon that must be considered

in any full account of the second language acquisition process", as Gass & Selinker (1992, p. 6) state. This paper, based on the corpus contrastive approach, focuses on transfer in the subsystem of types and features of NP with the word *effect* as an example, attempting to find out what influences their distribution features in the whole corpus. The knowledge of the native language in acquisition of a foreign language can indeed have a promotion or inhibition effect on the learner's progress in second language acquisition. Traditionally, facilitation effect is considered as positive transfer, while inhibition is regarded as negative transfer. The latter seems to be of concern among scientists working on second language acquisition. The question of what is more likely to be transferred from native language to second language and how the mechanism of transfer works has given rise to different linguistic models and hypotheses over the last two decades. One of the earlier hypotheses on language transfer, the Contrastive Analysis Hypothesis (Lado, 1957; Stockwell, Bowen, & Martin, 1965) tried to predict the likelihood of linguistic transfer in second language acquisition based on the similarities as well as differences between various aspects of native language and second language. That is, similarities in linguistic structures in two languages will result in positive transfer, while differences will create an interference which is known as negative transfer. However, the survey of the recent research on language transfer demonstrates that the Contrastive Analysis Hypothesis fails to find significant support and its validity has been questioned by many scholars (Gas & Selinker, 1983; McLaughlin, 1984). Another theory underlying language transfer is a theory of markedness (Eckman, Moravcsik, & Wirth, 1986; Seliger, 1991). The core hypothesis of markedness theory concerns correlations, i.e. pairs of "marked" (least distributed) and "unmarked" (more distributed) structural entities in the language. According to this theory, those linguistic phenomena in the target language which are more marked than the corresponding phenomena in the native language will be more difficult to learn. Nevertheless, there is a problem to apply the markedness principle to cross-linguistic analyses, which makes it problematic to predict which structures in second language would be more likely substituted with the corresponding structures in native language. Except for the above purely linguistic approaches to language transfer, there is an array of theories pertaining to a psycholinguistic view on language acquisition. This view has been always shaped by dominant psychological frameworks, i.e. behaviorist or cognitive. Within a behaviorist framework, which was particularly popular in 40s-50s (Fries, 1945; Lado, 1957), transfer was seen as a direct result of the influence imposed by native language structures on corresponding structures in second language. Over the last twenty years a cognitive approach to language transfer, as well as to other psycholinguistic phenomena, has prevailed in the field of SLA. One of the most important findings of the time was that native language directly and indirectly influences second language acquisition. Indirect influence, in turn, reflects underlying organization principles of the language and the learner's metalinguistic awareness of that knowledge. The most revolutionary linguistic theory of the past few decades within the cognitive framework was that of universal grammar proposed by Chomsky (1965). In the light of this theory, cross-linguistic influence must be predetermined by certain innate constraints existing in any natural language acquisition. According to Chomsky, the learner must take a very limited input in second language and construct a clean grammar of the language being learned. The universal grammar theory and its application to the major linguistic fields, including second language acquisition, have attracted much scientific attention over the last three decades (e.g., van Buren & Sharwood Smith, 1985; Corder, 1992; Flynn, 1986; White, 1992). However, it was also discussed and opposed by the connectionism theory (Gasser, 1990). Rather than focusing on innate constraints, connectionists try to look at the ways in which the learner extracts regularities from the second language input. In addition to models and theories which were briefly discussed above, different psycholinguistic factors, like metalinguistic awareness, processing demands, language proficiency, etc., have been reflected in studies on language transfer. Studies on these aspects of language transfer, although not frequent in the field, were mostly concerned with compensatory strategies that second language learners use to perform different linguistic tasks (Bialystok, 1990; Faerch & Kasper, 1983).

3. Methodology

In the following paragraphs the author is going to describe the data and software that took part in the research as well as the rationale behind our paper.

Our research explores two sub-sets of corpora, specifically the Chinese scholars' dissertation abstracts (CSDA) and English-speaking scholars' dissertation abstracts (ESDA), on the basis of one international peer reviewed journal L*anguage Learning* and one Chinese journal M*odern Foreign Language* respectively. Abstracts from Chinese scholars and international scholars produced the corpora from 2005 to 2015, and each of them is made up of over 50,000 words.

In order to retrieve the data for analysis, five tools, AntConc, Log-likelihood Value Calculator, Normalization, Treetagger and BFSU Colligator were employed.

The most obvious advantage of AntConc is its ability to compare the frequencies of the same word between two corpora. With the help of wordlist function, a list of words which is based on a comparison of the frequencies were obtained respectively, from which the top 20 frequency nouns in CSDA and ESDA were achieved and the word *effect* was chosen as an example for following analysis. Normalization is used together with Antconc to get standardized calculation of each word so that more objective results would be obtained. The log-likelihood value (LL value) is calculated for measuring the extent of difference in the relative frequency of a particular item between the two corpora. The larger the absolute value of the LL parameter is, the more different the two corpora are in terms of the relative frequency of the target item. The beauty of the Treetagger is in that it can get each word in the corpus tagged with the appropriate semantic category, so that the frequencies of different types of NP between the two corpora can be compared. BFSU Colligator is the main tool used to investigate features of colligation. It is developed by researchers from Beijing Foreign Studies University (BFSU) for analyzing characteristics of English syntactic pattern. By running this tool, sentences of different NP types are retrieved in each corpus so that frequencies and proportions of each NP type could be analyzed. COCA (Corpus of Contemporary American English) is founded by Mark Davies of Brigham Young University. It is the largest freely-available corpus of up-to-date English data of America compared with other corpora abroad and the only large and steady-going corpus of American English, which is used in this paper to check idiomatic English expressions of collocations.

4. Results and Discussion

In the following paragraphs, the author will report the occurrences of specific features within the appropriate classifications of NPs. Results pertaining to each of the major classifications of NP form the basis of the following sections. Examples from the two corpora are provided to exemplify the descriptions where necessary.

4.1 Classification of NP Types in CSDA and ESDA

According to the simplified classification of noun phrases described in Introduction, noun phrases included in the present study are classified into three major types: *(Determiner)+Head* (determiner is optional), *Premodifier+Head*, and *Head+Postmodifier*. Their features of use are investigated in terms of the frequencies as well as distributions across CSDA and ESDA.

4.1.1 Overall Frequencies and Distributions of NP in CSDA and ESDA

Table 1. Overall frequencies and proportions of NP

	CSDA		ESDA		
Item	Frequency	Proportion	Frequency	Proportion	LL
NP	12173	22.05%	11128	19.26%	0.000***+

Note. * P<.05; ** P<.01; *** P<.001; + indicates overuse in CSDA relative to ESDA.

As Table 1 shows, the overall proportion of NPs in these two corpora is above 20%, which leads support on the significance of NPs in sentences. To further explicate the usage of NP, the distributions of two types of NP across CSDA and ESDA were investigated.

Table 2. Frequencies and proportions of simple NP and complex NP

	CSDA		ESDA		
Item	Frequency	Proportion	Frequency	Proportion	LL
Simple NP	2876	23.63%	2413	21.68%	0.002**+
Complex NP	9297	76.37%	8715	78.32%	0.092-

Note. - indicates underuse in CSDA relative to ESDA.

The results in Table 2 can be summarized as follows: the overall frequencies and proportions of complex NPs are much higher than those of simple NPs, which indicate that complex NPs predominate NP types for both Chinese scholars and international scholars. Moreover, simple NPs are slightly overused in CSDA than in ESDA, while complex NPs are less frequently used in CSDA.
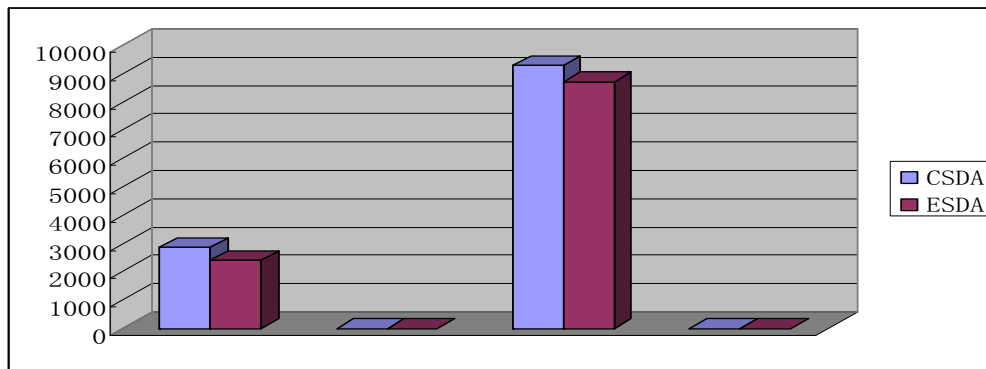
Figure 1. Distributions of two NP types in CSDA and ESDA

As the noun-phrase type comprising determiner and head is usually called simple NP whereas the other two types are called complex NP, the difference in the use of NPs can be noted more clearly as follows.

Table 3. Frequencies and proportions of three types of NP

Item	CSDA		ESDA		
	Frequency	Proportion	Frequency	Proportion	LL
1 (Determiner)+Head	2876	23.63%	2413	21.68%	0.002**+
2 Premodifier+Head	5465	44.89%	4977	44.73%	0.847+
3 Head+Postmodifier	3832	31.47%	3738	33.59%	0.005**-

As shown in Table 3, simple NPs were presented by the item *(determiner) +head*, while complex NPs were further divided into *premodifier+head* and *head+postmodifier*. It is clear to see that, in terms of the total proportions, Type 2 (*premodifier+head*) is the most frequently used one among the three types, while Type 1(*(determiner) +head*) is the least frequent one. Type 3(*head+postmodifier*) stands between Type 1 and Type 2, making up for more than 30%. Thus it can be concluded that the majority of noun phrases in both corpora is Type 2 and Type 3, which represent for complex NPs.

Furthermore, similarities can be noted between the frequencies of occurrence in CSDA and ESDA, and the discrepancies are easily found from the result of LL value. It is clear that Type 1 and Type 2 are more welcomed by Chinese scholars, whereas Type 3 is more frequently used by international scholars. There was a statistically significant difference in Type 1 between CSDA and ESDA, which can thus imply that Chinese scholars are more inclined to use simple NPs in academic writing abstracts.

4.1.2 Exampled Word "Effect"

Table 4. Word *effect* in Log-likelihood Ratio Calculator

Word	Freq. in corpus 1	Freq. in corpus 2	LL	Sig.
effect	135	197	9.00	0.003**-

* Corpus Size 1:55200; Corpus Size 2: 57770

As nine words occurred both in CSDA and ESDA among the top 20 frequency nouns, the LL demonstrates a plus or minus symbol before the log-likelihood value to indicate overuse or underuse respectively in CSDA relative to ESDA. The log-likelihood value itself is always a positive number. However, my script compares relative frequencies between the two corpora in order to insert an indicator for "+"overuse and "-"underuse of CSDA relative to ESDA. A plus sign is on behalf of the excessive use, while minus sign is on behalf of the lack of use. It is easy to find that the word *effect* shows a remarkable discrepancy between CSDA and ESDA by running Log-likelihood Value Calculator.

4.2 Distributions of Major NP Types in CSDA and ESDA

As advanced in section 4.1, it is stated that complex noun phrases account for an overwhelming majority in

CSDA and ESDA. Hence, in this section, the comparison of NP types between the two corpora will be carried out in terms of *premodifier+head* and *head+postmodifier* by the example word *effect*.

4.2.1 Distributions of *Premodifier+Head*

To further investigate the usage of the type *premodifier+head* in abstract writing, two major patterns, *adjective +head (JJ+NN)* and *noun+head (NN+NN)* are presented. The frequencies and proportions of these two types in CSDA are compared with those in ESDA, and the results are listed below.

Table 5. Frequencies and proportions of *premodifier+head* types

	CSDA		ESDA		
Item	Frequency	Proportion	Frequency	Proportion	LL
1 Adjective+Noun	3423	62.63%	3198	64.26%	0.299-
2 Noun+Noun	2042	37.37%	1779	35.74%	0.171+

The Table above illustrates that the total number of Type 1 is much larger than those of Type 2. It is obvious that Type 1 predominates the uses of *premodifier+head* in both CSDA and ESDA, as frequency of Type 1 are nearly double the sum of frequency of Type 2.

The differences between CSDA and ESDA are also presented in Table 5. Though the frequency of Type 1 in CSDA are more than those in ESDA, the proportion of Type 1 in CSDA are less than those in ESDA, while the proportion of Type 2 in CSDA is larger than that in ESDA.

Table 5 shows that in different patterns of *premodifier+head*, *adj+noun* is the typical collocation type, accounting for more than half of all the occurrences. So the next section will take word *effect* as an example to further compare and analyze features of NP in CSDA and ESDA.

4.2.2 Distributions of *Head+Postmodifier*

Head+postmodifier can be subdivided into several types, among which five major patterns are investigated in this study. They are *head+to-infinite clause (NN+TO), head+preposition or subordinating conjunction (NN+IN), head+ing-clause (NN+VVG), head+ed-clause (NN+VVN)* and *head+relative clause*.

Table 6. Frequencies and proportions of *head+postmodifier* types

	CSDA		ESDA		
Item	Freq	Proportion	Freq	Proportion	LL
1 noun+to-infinite clause	244	6.37%	320	8.56%	0.000***-
2 noun+preposition/ subordinating conjunction	3137	81.86%	2953	78.20%	0.165+
3 noun+ing-clause	155	4.04%	154	4.11%	0.872-
4 noun+ed-clause	134	3.50%	144	3.85%	0.420-
5 noun+relative clause	162	4.23%	167	4.47%	0.616-

As Table 6 shows, on the whole, Type 2 (NN+IN) is the vast majority among the five types, making up for 80.17%, while the sum of other types accounting for barely 19.55%. The most frequent NP types with postmodifiers in total are, in descending order of frequency, Type2 (3137 occurrences), Type1 (244 occurrences), Type5 (162 occurrences), Type3 (155 occurrences), Type4 (134 occurrences). In a similar way, the differences can be observed between CSDA and ESDA as well. Only the proportion of Type 2 is larger in CSDA than those in ESDA, while the proportions of other four types, Type 1, Type 3, Type 4 and Type 5 are smaller in CSDA than those in ESDA.

The above chart shows that in different patterns of *head+postmodifier*, *noun+preposition* are the typical collocation type, accounting for more than 80%. So the following section will focus on this type of collocation and take word *effect* as an example to further compare and analyze features of NP in CSDA and ESDA.

4.3 Features of Major NP Exampled by Word "Effect"

In this Section, two major NP types, *Adj+Noun* and *Noun+Preposition*, will be discussed by the example word *effect* in terms of collocation and colligation features.

4.3.1 Features of "adjective+effect"

To observe the distribution tendency NP, the most frequently used adjectives in CSDA and ESDA are listed respectively in Table 7 with their corresponding frequency ranks.

Table 7. Significant collocates of *adj.+effect* in CSDA and ESDA

Rank	CSDA		ESDA	
	Collocates	freq.	Collocates	freq.
1	significant	6	larger	6
2	priming	6	positive	5
3	main	6	priming	4
4	dependent	3	learning	4
5	stylistic	2	significant	3
6	strong	2	overall	3
7	pragmatic	2	greater	3
8	positive	2	beneficial	2
9	differential	2	reactive	2
10	cognitive	2	possible	2
11	weak	1	negative	2
12	same	1	mixed	2
13	rhetorical	1	genetic	2
14	predicting	1	differential	2
15	morphological	1	combined	2
16	linguistic	1	unfavorable	1
17	less	1	supporting	1
18	large	1	subsequent	1
19	interactive	1	stronger	1
20	good	1	similar	1
21	forgetting	1	reading	1
22	different	1	main	1
23	contextual	1	large	1
24	ambiguous	1	investigated	1
25			interpretating	1
26			independent	1
27			facilitative	1
28			facilitating	1
29			direct	1
30			different	1
31			detrimental	1
32			cumulative	1
33			consistent	1
34			comparative	1
35			clear	1
36			argued	1
37			accumulated	1

*Total number of collocate types in CSDA: 24

*Total number of collocate types in ESDA: 37

By the statistics of frequency in Antconc, the Table 7 listed top 24 and top 37 most frequent adjectives bonded with *effect* in CSDA and ESDA. It is found that among the twenty nouns, six of them occurred in both sides. They are *significant, priming, positive, main, differential, different,* which reflect some common features of the Chinese scholars and native speakers in academic writing abstracts. In the following sections, the author chose the word *effect* as a typical case from the aspect of collocation in order to have a detailed understanding of different types of NP features.

In CSDA, it can be seen from the following parts that adjectives can be further divided into three categories in terms of the semantic prosody, namely positive, neutral and negative.

Positive collocates
{
larger, positive, priming, significant
greater, beneficial, reactive, supporting
stronger, main, large, independent
facilitative, facilitating, direct
cumulative, clear, accumulated
}

Neutral collocates
{
learning, possible, mixed, genetic, differential
combined, subsequent, similar, reading
investigated, interpretating, different
consistent, comparative, argued
}

Negative collocates   negative, unfavorable, detrimental

Similarly, in CSDA, the adjectives could also be divided into three categories.

Positive collocates
{
significant, priming, main, strong, good
positive, interactive, large, predicting
}

Neutral collocates
{
stylistic, pragmatic, differential, cognitive
rhetorical, same, morphological, linguistic
dependent, forgetting, less, different
ambiguous, contextual
}

Negative collocates   weak

The concordances of *effect* in CSDA and ESDA are shown below.

…have a differential **effect** on adult second language reading comprehension… (ESDA_TXT)

…..exhibited larger **effect** of the form than of the semantic manipulation…….. (ESDA_TXT)

…The comparative **effect** sizes were used in a subsequent moderator………… (ESDA_TXT)

...CBI had a greater **effect** than PBI when the acquisition was measured……... (ESDA_TXT)

……have different **effects** on the listening performance………………………. (CSDA_TXT)

…..their pragmatic **effects** in the constructive and deconstructive processes..… (CSDA_TXT)

…........the cognitive **effect** of the two discourses shows that this strategy…........ (CSDA_TXT)

…........a significant **effect** on the learners…………………………………. (CSDA_TXT)

………Ambiguous **effects** are found both in the initial stage and at the end…... (CSDA_TXT)

While the collocations of the word in CSDA are closely to international scholars and more or less proficient, the difference between the two corpora still exists. Some easy words, or simple words, can still be seen in CSDA, such as the word *good*. Because of the general collocations, the writing will be led to a dull and imprecise level.

In Table 7 there are altogether 24 significant adjective collocates of *effect*, while in ESDA there are 37 types of adjectives, which indicates that compared with collocations in use in CSDA, the collocations in ESDA has a wider range of complicated and advanced words. These concrete words can help the writer express himself more vividly. So compared with international scholars, Chinese scholars' writing level is relatively low as less using concrete and advanced collocates.

The wide range of collocation is also reflected in the positive collocates of word *effect* in ESDA. International scholars use eighteen different types of positive adjectives to modify the word *effect*, such as *significant, larger, facilitative, cumulative,* etc. Chinese scholars, on the other hand, tend to use only nine different types of adjectives which belongs to more common ones, such as *significant, good, positive, strong,* etc.

Besides, as for the word *effect*, more interlanguage collocations exist in CSDA than typical collocations. For example, native English speakers tend to use *facilitative effect, cumulative effect, unfavorable effect, detrimental effect, consistent effect, differential effect* and so on, while Chinese scholars use *good effect, large effect,* and *strong effect* more often. Albeit Chinese English learners also use the expression "*weak effect*" to express negative semantic prosody, there is still a little bit difference with the expression "*unfavorable effect*" and "*detrimental effect*" used by native English speakers.

All in all, there are similarities as well as differences for the collocations between international scholars and Chinese scholars. Here in CSDA, there are hardly any abnormal collocations as these Chinese scholars are of high level of English competence. Nevertheless, it is still found that English learners can use English more correctly at grammatical level, while they are chinned to choose a small range of simple or common collocations to express their opinions, which makes the Chinese scholars less native and natural despite the fact that it corresponds to the rules of English collocation. Chinese scholars are likely to use synonyms to replace the more appropriate adjectives, which cause non-native expressions. i.e. synonyms (wide perspective instead of broad perspective), hyponyms (*good effect* instead of *facilitative effect* or *beneficial effect*), or antonyms (not happy instead of sad), and the analytic strategy refers to the way in which EFL learners use indirect expressions to convey their ideas.

4.3.2 Features of Colligation "e*ffect+preposition*"

In CSDA, the colligation of *effect+Prep.* appeared 44 times, accounting for 56.41% of the total frequency; In ESDA, on the other hand, *effect+Prep.* was also used 44 times, with the proportion of 55.69% in total, which is practically similar to the frequency in CSDA. There are five types of prepositions appeared *respectively* in CSDA and ESDA, namely, *of, on, in, by, to* in CSDA, and *of, on, for, between, at* in ESDA.

Table 8. Frequencies of *effect+preposition* in CSDA and ESDA

	effect of	effect on	effect in
CSDA	21	15	5
ESDA	30	9	0

As can be seen in Table 8, for the structure *effect + prep.*, *effect of* and *effect on* most commonly occurred in both corpora. According to the explanation of Oxford Advanced Dictionary, effect on sb/sth, or effect of sb/sth illustrated a change that sb/sth causes in sb/sth else; a result. However, there is no explanation of collocate *effect in* in Oxford Advanced Dictionary albeit the collocations of this match indeed occurred in CSDA. For further exploration, all collocations of *effect in* in CSDA are indexed with following results.

1. This paper reports on two word-naming experiments aimed to explore the language-dependent ***effect in*** Cantonese-Mandarin bilinguals of different proficiency.

2. The results from the two experiments indicate that the language-dependent ***effect in*** bilingual language production is influenced by bilinguals.

3. Most previous studies concerning the ambiguity ***effect in*** language processing have focused more attention on monolingual subjects than on bilingual subjects.

4. The current study explored the psychological mechanism of the cross-language syntactic priming ***effect in*** two experiments.

5. The results show a main ***effect in*** metaphor aptness but not in vehicle conventionality.

In these sentences, *effect* appears together with preposition *in* in form, but the understanding of the meaning is separate, thus it could be say that there is no mistake at grammatical level. The author believes the emergence of this form related to the influence of mother tongue migration. As for the frequency of *effect + prep* is not that much in these two corpora, the author chose COCA Corpus for further retrieval, in order to confirm the NP features of international scholars. As shown in the table below, collocates are ranked by collocation strength, and the higher the MI value (mutual information), the greater the collocation strength is. Here are three collocations which MI values are greater than 3. They are *effect upon, effect on, effect of, effect throughout, effect due*.

Table 9. Collocate of *effect+proposition* in COCA Corpus

Collocate	FREQ	ALL	%	MI
Effect upon	249	63715	0.39	6.00
Effect on	9957	3135774	0.32	5.70
Effect of	13133	11950786	0.11	4.17
Effect throughout	41	47023	0.09	3.84
Effect due	21	39838	0.05	3.11
Effect in	13231	8364885	0.16	0.47

The further analysis of the data showed that the MI value of *effect in* (0.47) is much lower than the collocation *effect on* (5.70) and *effect of* (4.17), which denotes that the collocate of *"effect in"* is not that common for native speakers. As the Table shows, international scholars are more accustomed to using *effect on* and *effect of*, and that is similar to the results reflected in Table 9. The author infers that the occurrence of this kind of match used by Chinese scholars is likely to be influenced by mother tongue transfer.

Learning a language as an EFL learner can be a laborious and time-consuming process, and even talented scholars tend to fall short of native-like proficiency. (Lightbown 2000). The reason why Chinese scholars use collocations that are not so common in English is mother tongue inference which refers to the transference from the usages of the mother tongue to the target language. In a sense, these irregular usages of the target language influence Chinese scholars' acquisition of the target language. Corder (1981, p. 24) considers that many language learners have interlanguage which relates to mother tongue inference, and the irregular collocations usually belong to a kind of interlanguage. In different languages, the collocations are various, and the models of collocations have their certain patterns. But Chinese scholars usually translate Chinese into English word by word, which lead to unnatural collocations. Some researchers consider that mother tongue inference is caused by their lack of mastering the target language (Newmark, 1996, p. 99; James, 2001, p. 175). Therefore, before language learners master enough knowledge about the target language, mother tongue inference has predominance, and the errors are called interlingual transfer errors.

Another main reason why Chinese scholars make unnatural collocations is related with their strategies, which is also called intralingual transfer errors. James (2001) discusses the intralingual transfer errors related with collocations in detail. Based on the errors of the communication strategies, there are two classifications, which are the comprehensive strategy and the analytic strategy. The former means that Chinese scholars are inclined to use a synonym to take place of another word, when they have no a proper expression, i.e. synonyms (wide perspective instead of broad perspective), hyponyms (sports instead of basketball), or antonyms (not happy instead of sad), and the analytic strategy refers to the way in which EFL learners use indirect expressions to convey their ideas. There are always a number of incorrect collocations in EFL learners' English. One of the reasons is that Chinese scholars use synonyms to replace the more appropriate words, which cause non-native expressions.

5. Conclusion

Summing up, the research examines the use of NP in Chinese scholars' abstracts in the context of types and features. In particular, the author's interest lies in documenting whether NPs are used differently by Chinese scholars and international scholars in academic writing abstracts. The author's primary motivation is then to explore there is a link between language transfer influence and the use of noun phrases in the context of very talented Chinese scholars, an under-explored area in learner language research.

References

Algeo, J. (2006). *British or American English? A Handbook of Word and Grammar Patterns*. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511607240

Baker, P., Hardie, A., & McEnery, A. (2006). *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Bialystok, E. (1990). *Communication strategies: A psychological analysis of second-language use*. Blackwell.

Biber, D., & Reppen, R. (1998). Comparing Native and Learner Perspectives on English Grammar: A Study of Complement Clauses. In S. Granger (Ed.), *Learner English on Computer* (pp. 145-158). London: Longman.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.

Chomsky, N. (1965). *Aspects of the theory of syntax Cambridge*. MIT Press.

Corder, S. P., & Corder, S. P. (1981). *Error analysis and interlanguage* (Vol. 112). Oxford: Oxford University Press.

Ellis, R. (1994). *The study of second language acquisition*. Oxford University Press.

Færch, C., & Kasper, G. (1983). Plans and strategies in foreign language communication. *Strategies in interlanguage communication*, 20-60.

Flynn, L. J., Barry, J. C., Morgan, M. E., Pilbeam, D., Jacobs, L. L., & Lindsay, E. H. (1995). Neogene Siwalik mammalian lineages: species longevities, rates of change, and modes of speciation. *Palaeogeography,*

Palaeoclimatology, Palaeoecology, *115*(1), 249-264. http://dx.doi.org/10.1016/0031-0182(94)00114-N

Fries, C. C. (1945). Teaching and learning English as a Foreign Language.

Gass, S. M., & Selinker, L. (Eds.). (1992). *Language Transfer in Language Learning: Revised edition* (Vol. 5). John Benjamins Publishing. http://dx.doi.org/10.1075/lald.5

Isurin, L. (2005). Cross linguistic transfer in word order: Evidence from L1 forgetting and L2 acquisition. In *Proceedings of the 4th International Symposium on Bilingualism* (pp. 1115-1130).

Kellerman, E. (1995). Crosslinguistic influence: Transfer to nowhere? *Annual review of applied linguistics*, *15*, 125-150. http://dx.doi.org/10.1017/S0267190500002658

Lado, R. (1957). Linguistics Across Cultures: Applied Linguistics for Language Teachers.

Liang, M., Li, W., & Xu, J. (2010). *Using corpora: a practical coursebook.* Beijing: Foreign Langue Teaching and Research Press.

Lightbown, P. M. (2000). Anniversary article. Classroom SLA research and second language teaching. *Applied linguistics*, *21*(4), 431-462. http://dx.doi.org/10.1093/applin/21.4.431

McLaughlin, M. L. (1984). Conversation. *How talk is organized. Beverly Hills (Calif.)*.

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge University Press. http://dx.doi.org/10.1017/cbo9781139524537

Oxford Advanced Learner's English-Chinese Dictionary. (6th ed.). Beijing: The Commercial Press.

Parkinson, J., & Musgrave, J. (2014). Development of Noun Phrase Complexity in the Writing of English for Academic Purpose Students. *Journal of English for Academic Purposes, 14*, 48-59. http://dx.doi.org/10.1016/j.jeap.2013.12.001

Plumb, R. S., Granger, J. H., Stumpf, C. L., Johnson, K. A., Smith, B. W., Gaulitz, S., ... & Castro-Perez, J. (2005). A rapid screening approach to metabonomics using UPLC and oa-TOF mass spectrometry: application to age, gender and diurnal variation in normal/Zucker obese rats and black, white and nude mice. *Analyst*, *130*(6), 844-849. http://dx.doi.org/10.1039/b501767j

Quirk, R. (1985). The English language in a global context. *English in the world: Teaching and learning the language and literatures*, *16*.

Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Stockwell, R. P., Bowen, J. D., & Martin, J. W. (1965). *The grammatical structures of English and Spanish*. University of Chicago Press.

The Corpus of Contemporary American English (COCA). (1990-2012). *Brigham Young University.*

Van Buren, P., & Smith, M. S. (1985). The acquisition of preposition stranding by second language learners and parametric variation. *Second language research*, *1*(1), 18-46. http://dx.doi.org/10.1177/026765838500100103

White, S. R. (1992). Density matrix formulation for quantum renormalization groups. *Physical Review Letters*, *69*(19), 2863. http://dx.doi.org/10.1103/PhysRevLett.69.2863

Yang, S. H. I. (2010). Conceptual Metaphor & Language Transfer: A Corpus-based Study. *Overseas English*, (11).

Zhang, H. P., & Liu, Y. B. (2014). A Corpus Study of Most Frequently Used English Verbs by Chinese Beginner Learners from a Conceptual Transfer Perspective. *International Journal of Corpus Linguistics, 19*(2), 252-279. http://dx.doi.org/10.1075/ijcl.19.2.05hui