Academic Features Among the Non-academic: A Case of Chinese Argumentative Texts as Compared to BaREnLoB

Zhang Xin [1*]   Long Shaoyun [2]
1.Zhang Xin, Master, Foreign Language College, Jiangxi Normal University, Jiangxi, China
2.Long Shaoyun, Professor, Foreign Language College, Jiangxi Normal University, Jiangxi, China
*E-mail of the corresponding author: steven8810469@sina.com

Abstract
Compared with different degrees of academic corpora(RC): BaREnLoB, this study conducted a CIA analysis to find out whether the usage features owned by Chinese English majors in their argumentative texts in WECCL2.0 (the observation corpus,OC henceforth) can demonstrate different academic features except for non academic ones. Results show some academic features among the non-academic(i.e.almost the same strong narrative features, and insufficient epistemic comments, rare depiction and classification of particular things, i.e. similar to previous findings). It shows: (i) NN2 (mostly followed by should) and we (the usual form to refer to the author himself /herself or the research group themselves) are usually incorporated as subjects; (ii) The standard frequency (PMW) of the top 20 verbs(including be, have, find, know, become, see, etc.) and all the 55 pointed common stative ones(Zhang Z.B, 2003) between OC and RC differ much although the standard frequency(PMW) of stative verbs among the 20 top between differs slightly. (iii) the occurrences of " they/he/she + modal verbs" in OC are many times more than those in RC, even "we/you/I + modal verbs" between show the nonnatives employ much less than the natives. (iv) there are 4 out of 19 carefully-picked abstract nouns before modal verbs are in OC while none in RC. In short, this research finds that the English argumentative writings by Chinese English majors display some academic features as mirrored by the reference corpus except for some non-academic features as we found before. Finally, some implications for the teaching and research of modal verbs are discussed.
Keywords: deontic and epistemic modality, academic feature among non academic, Chinese argumentative texts

1.Introduction

Studies on English modal verbs has attracted much attention (e.g., Aijmer 2002; Biber et al.,1999; Bouhlal, F., Horst, M.,& Martini, J., 2018; Coates,1983; Cournane, A., 2014; deHaan,1997; Fan, X., 2016; Halliday, M.A.,1985; Hunston,S, 2004; Hsin-i Chen, 2010; Huddleston & Pullum, 2002; Hondo,J., 2012; Mindt, 1993; Kakzhanova, F. A., 2013; Khojasteh, L., & Mukundan, J., 2011; Kwary, D. A., Kirana, A., & Artha, A. F. 2017; Palmer, 2001; Quirk et al., 1985; Sweetser, 1990; Pastor, M. A. J., & Pérez-Guillot, C. ,2015), and they are inclusive. Involving research on keyness alone, there are some researches, for example, Bondi, M., & Scott, M., (2010), Scott, M., & Tribble, C., (2006), Gabrielatos, C.,(2018), and some other empirical studies. (e.g.Culpeper, J.,2009;Gabrielatos, C., & Marchi, A.,2012; Pojanapunya, p., &Todd, r.w.,2018;Pojanapunya, p.,&Todd, r. w., 2018;Rivers, D.,& Ross, A.,2018;Bowker, l.,(2018) et al.). Early studies on modality in China were mainly done from the perspective of system functional linguistics (e.g., Huang tao, 2006; Li Zhanzi, 2001, 2002, 2005; Zhu Yongsheng, 2006; Yuan xiaoning, 2000). There are also some studies done from the perspective of cognition linguistics (e.g. Liang Xiaobo, 2001), Also, there are studies on the modality of the Chinese language(e.g. Grammar Research and Exploration. xii from Chinese Language). However, studies on "modal sequences" in west countries are rarely involved, especially the literature with "modal sequence" as title or key word. In China, since 2008, there have been many such papers with "modal sequence" as topic or key word (e.g., Author, 2011,2012, 2013, 2014,2016; Chen Aibing, 2012; Feng Sha, 2017; Liang Maocheng, 2008; Liu Wenyan, 2009, 2013; Pang Jixian, & Chen Jun,2018; Qiao lingling, et.al,2014; Tang liling,2013;WangGuan 2013,2015; Zhao Lizhu,2018;Yan Pengfei,2017;Zhang Zhenghou,,2016; 2016;Zhang Xueyuan,2015; Zhang Hongqiong,2014; Zhang Jie, 2013). All the studies are presented in various ways because the corpora they used and the settings they established are quite different, so the conclusions obtained are not so comprehensive, consistent and reliable(e.g.no effect size is involved even in one single research). Also, whether there exists academic features among the non-academic in the observation corpora was seldom discussed before. Hence, there is a new way for this study to go, which, in turn, may confirm our intuition of the difference of the academic degrees of the reference corpora (including here quasi-academic and academic texts) and validate the case as Mike Scott(2016) checked to be. After all, whether the epistemic sequences hidden in observation corpus could betray some academic features against the reference corpora is a

very new research area that aims to help language researchers and instructors to learn something from the inter-language texts more confidently.

2.Modal(epistemic and deontic) Sequences and Modality: Definition and relationship among regarding Epistemic Sequence

Modals are general statements that represent the notion of the mind or events that may or may not take place in the future and reflect the speakers' attitudes about what he/she says (Palmer, 2001). Halliday et al. (2004) believed that modality is the estimation and uncertainty of language users' understanding of things, unlike the clear "affirmation" and "negation" of polarity. Because of the content or "proposal" or "proposition" exchanged by language users, there was a distinction between "modalization" and "modulation", and it was assumed that "modalization" could be described in terms of probability and frequency, while "intentionality" in terms of "duty" and "orientation".Halliday et al. (2014) also believed that "modalization" and "normality" are often embodied by modal verbs in traditional grammar, but in the view of functional grammar, they can also be expressed by adjectives, adverbs, nouns and their corresponding structures, and pointed out that when expressing semantics in negative sentences, attention should be paid to distinguish between topic negation and modal negation. In short, Halliday(1985,1994,2004,2014) made a new description of modality and other related topics under the framework of interpersonal meaning of systemic functional grammar, which broadened the concept of modality, thus providing a more complete analysis approach for the study of modality and its meaning from the perspective of discourse.

Modal verbs are the main carrier of modality, which roughly falls into two categories: deontic and epistemic (Biber et al., 1999; de Haan 1997; Larreya 2004; Coates, 1983; Palmer, 2001,etc). Deontic modality indicates the responsibility felt by the speaker or the acts that the speaker is allowed or required to perform, while epistemic modality communicates the speaker's degree of certainty, the truth or falsity of the issue under discussion, or the likelihood of the event involved (Biber et a1., 1999; Coates,1983).

Modal verbs have interpersonal meanings (Biber et al., 1999; Halliday, 1985), and since most of them are polysemy with their meanings overlapped (Aijmer, 2002; Mindt, 1993), it takes context to specify their modality in use (Hunston,2004; Mindt, 1993). Embodying attitudes, emotions ,and state of affairs, modal verbs can indicate likelihood, permission, obligation, necessity, prediction, will, intention, determination, and even more complicated and subtle emotions like pity, worry, happiness, euphemism, and dare, modal verbs are usually taken by second language learners as one of the most challenging grammatical points (Author , 2012, 2013).

The study of modality in the English language is regarded as the most persistent and fascinating area of philosophical and linguistic inquiry (Hoye, 1997). According to Quirk, et.al. (1985, 235),modality is "the manner in which the meaning of a clause is qualified so as to reflect the speaker's judgment of the likelihood of the proposition of the sentence being true".

Modality can express a wide range of semantic meanings, like obligation, necessity, permission, request and so on (Quirk et al., 1985). A variety of devices can be employed to convey modality. Lexical devices, for example, by using nouns (intention, determination, hope, presumption and expectation etc.), adjectives (certain, doubtful, likely, conceivable, possible, and sure etc.), adverbs (hardly, perhaps, possibly, probably, and evidently etc.) and verbs (doubt, believe, think, predict, and suggest etc.) to express modality (Hermerén, this study. Nine core modal verbs have been identified in past literatures as core modals, including can, could, may, might, shall, should, will, would and must (Quirk et al., 1985; Biber et al, 1999). According to Halliday (1976), they are distinguished from lexical verbs by grammatical properties including a lack of non-tensed forms, no person-number agreement, occurrence with a following verb in bare infinitival form, on-occurrence in imperative clauses and so on. Other modals have also been identified. Marginal modal verbs include need to, dare to, used to and ought to(Biber et al., 1999). Quasi- modals include had better, have to, have got to, be supposed to, be going to and can co-occur with modal verbs (Collins, 2009, p. 15). Expressions like call into question, chances are that, and it seems plausible that that define degree of certainty. According to the research by Gabrielatos & T. McEnery (2005), epistemic modality is mainly carried by the modal verbs listed above, which account for 83% of all modal sequences.

Modal verbs should not be taken as a pack of discrete terms with their meanings overlapped, but as components of unambiguous modal sequences, for they are usually combined with auxiliaries and notional verbs in the form of "subject + modal verbs + auxiliary /notional verbs." The expression constituted by a modal verb and the components before and/or after is considered as a whole and acknowledged as a modal sequence, no matter if it is a colligation or a collocation (Huston, 2001, 2004), and no matter that it is in the form of "subject + modal verbs" or "modal verbs + verbs + complements." It is Susan Hunston (2004) who first put forward the concept of modal sequence and she has conducted many research on it. For instance, she chose one meaning of modal verbs, summarized a variety of related sequences, and listed numerous examples. According to Susan Hunston(2004), a modal verb relies heavily on the collocation to tell the meaning of the sequence "modal verbs+ be+ adjective or verb form." For example, we can tell the sequences I must be going and I must be getting back indicate responsibility instead of prediction by examining them as collocations. I must, as seen by Susan Hunston, is a

different linguistic phenomenon from you must. She (2004) stated that polysemous modal verbs interacted with their contexts (verbs included) in a dependent, non- interferring way and a modal sequence, as a whole, made sense only when its constitutive modal auxiliary interacted and connected well with the immediate contexts (verbs included). Therefore, she is convinced that teaching students modal sequences is more helpful in their language acquisition than only telling them the meaning of modal verbs as auxiliaries (H.Susan, 2004).

The concept of modal sequence is also approved by Sinclair, who thinks that rather than individual lexical items, linguistic meaning stems from units of meaning, e.g., lexical sequences are constituted by lexical items (Sinclair, 1991, 2004). He states that the recognition of phraseology will give a full explanation to the use of lexis (Sinclair, 1995). His research finds that must can express a speaker's different intentions when used in different collocations and concludes that one should learn modal verbs by detecting the multiple meanings of the collocations they constitute, instead of treating them separately.

2.1 Relationship between Modal Sequences(Syntactic structures) and Modality

Colligations and collocations constituted by modal verbs and the components before or after modal verbs can be in the form of "animated/unanimated subject/there/other introducers ect.+ modal verbs", or "modal verbs + stative/dynamic verbs", or "modal verbs + perfective verbs/progressive verbs", or having an epistemic modal adverb between modal verbs and notional verbs.

Biber et al. (1999), Coates (1983), and Wärnsby (2003) think the modality of modal verbs can be predicted through the syntactic structure involving them, since there is a strong correspondence between the structure of modal sequences and their modality, which can be seen from table 1.

Table 1. Congruent Relationship between Syntactic Structures of Modal Verbs and Modality

Syntactic structures	Examples	Epistemic modality	Deontic modality
VM + epistemic adverbs	can probably do	Yes	No
Introducers + VM	it may be true that	Yes	No
VM + perfective verbs	must have done	Yes	No
VM + progressive verbs	must be working	Yes	No
VM + stative verbs	must be	Yes	No
VM + dynamic verbs	must work	No	Yes
animated subject + VM	you can	No	Yes
unanimated subject + VM	The weather may	Yes	No
there + VM + be	there must be	Yes	No

Note: VM=modal verb

2.2 Hedging and Modality

Hedges are included in a research paper (hereafter RP) or an high-stake time-limited argumentative essays to express meaning tentatively, leaving space open for readers' possible different views or interpretations. In this sense, they contribute to avoiding threatening peer readers or graders and to expressing modesty towards them (Myers 1989). In addition, it has been claimed that the use of hedges can respond to other pragmatic functions; their inclusion in academic or non-academic texts may respond to a need to convince readers/graders of their arguments so that they become accepted, taking into account the communities' norms and values (Hyland 1998).

Hedging has come to be seen as a key characteristic of academic discourse, due to the indeterminate nature of knowledge produced and distributed in the academia(Pilar Mur-Dueñas,2017).

Hedging is considered a problematic concept and there does not seem to be full agreement on what counts as a hedge. One of the problems that hinders homogeneity and consensus in what is to be considered a hedge is "the multiplicity of forms hedges may take" (Varttala 2001, 24). Hedges can be seen to constitute an open category, to which new realisations can be added. Another problem is that not all instances of a linguistic form can be considered hedges. They are context-dependent: "no linguistic items are inherently hedgy but can acquire this quality depending on the communicative context or the co-text" (Markkanen & Schröder 1997, 4).

However, hedging in this study refers to lexico-grammatical choices in the academic or argumentative text which help writers or test-takers withhold full commitment from a proposition, and which contribute to modulating the expression of (un) certainty, commitment and (im) precision in view of the readership the text is intended for. This can be the result of not having enough empirical evidence or assurance to claim that a proposition or an argument is absolutely true or certain, or the result of a desire not to express that certainty (Hyland 1998), even if the utterer(s)/writer(s) is(are) going to convey his/her/their degree of certainty or the more truth or falsity of some issues under discussion, or the likelihood of the event involved. Their use thus allows for alternative voices and viewpoints, opening up negotiation of meaning and space for possible opposing views, which may be expected in a given disciplinary community and a given genre.

Modal verbs are considered the most prototypical realization of hedging(Pilar Mur-Dueñas,2017). Previous

analyses of hedging in written academic discourse include modal verbs as a first and main category (e.g. Salager-Meyer 1994; Crompton 1997; Hyland 1998, 2005; Varttala 2001). The study here presented, thus, focuses on the study of epistemic modality understood as "the speaker's assumptions or assessment of possibilities and, in most cases, it indicates the speaker's confidence (or lack of confidence) in the truth of the proposition expressed" (Coates 1983, 18). It draws on the expression of 'certainty' or 'possibility' meanings (Palmer, 1990) by means of modal verbs. The expression of possibility and of the degree of commitment on the part of the authors is especially relevant in academic writing/argumentative essays, as discussed above, in which disciplinary, language and culture-driven socio-pragmatic conventions need to befollowed.

Hedging, in general, and hedging modal verbs/sequences, in particular, have been shown by previous intercultural EAP/EGP/ESP researches to be used differently across different genres. It is, therefore, interesting to look into the use of this interpersonal feature made by writers in different academic /non-academic contexts and explore the extent to which their use of ELF resembles or differs from the use made of the same interpersonally-driven feature in ENL RPs. In particular, 9 core modals: may, might, can, could, will ,would, shall, should and must and the modal sequences regarding epistemic and deontic modality made from these modal verbs in sharply-different contexts will be analyzed, as they contribute to expressing the test– takers'/author's diversified degree of confidence in the propositions/claims they make.

2.3 Relationship between Hedging and Epistemic Sequences

Hedging has come to be seen as a key characteristic of academic discourse, and modal verbs are considered the most prototypical realization of hedging(Pilar Mur-Dueñas,2017). In this sense, we find some kind of relationship between modal verbs/sequences and academic and non-academic discourse. In order for proof-readers or test-graders to convey his/her/their degree of certainty or truth or falsity of some issues under discussion, or likelihood of the event involved, or mood that they are in to negotiate with to avoid to be attacked or criticized in their academic utterances, one of the most challenging jobs for them is to avoid threatening peer readers or graders but to expressing modesty towards them (Myers 1989) so as to get published or admitted. In a sense, the more hedging it is used, the more chance for papers to publish or essays to be highly scored. In another way, the more epistemic, the more academic, and the more chance. Thus, the more hedges, the more epistemic sequences in papers or argumentative texts.

2.4 Relationship between Epistemic Sequences and Academic feature

As can be seem from the epistemic modality in Table 1, it is easy to conclude that the more epistemic sequence structure a paragraph has, the more epistemic it is, the more epistemic it is, the more possibilities that the utterer(s)/writer(s) is/are going to convey his/her/their degree of certainty or the more truth or falsity of   some issues under discussion, or the likelihood of the event involved, or the more mood that they are in to negotiate with to avoid to be attacked or criticized in their utterances, which,we argue, is a kind of academic negotiation feature (short for academic feature here) related to modality(Author,2014), while such features embedded in deontic sequences in such sharp contrast to epistemic sequences as should, must ect. are overused and could and would are underused, so the overuse of verbs with unmarked aspect and voice after modal verbs, and the underuse of epistemic sequences such as can / could + VBN+(/RG.etc./)+VVN (e.g. can be greatly cancelled, could be fairly bought,etc.)should be called non-academic features here in the article.

Research(Z.H.Zhang,2016) reveals that in academic papers, regardless of the subject or themes, the more epistemic modality, the more mood it is in to negotiate with the readers or the grading teachers or blind reviewers, and the higher the scores, the more possibilities for the papers to be published. And the epistemic degree varies according to what kind of modal verbs, strong mood or weak mood of modal verbs. As Hyland (1998) argues, the purpose of academic discourse writers is to persuade readers to accept their views and then obtain permission to enter a certain discourse community. Li (2011)also believes that scholars tend to refuse to express absolute commitment in order to avoid unnecessary criticism or attack and increase the room for negotiation (quoted from Z.H.Zhang,2016). Academic discourse writers even ordinary essay writers intend to employ all above-mentioned linguistic strategies/means, i.e. the negotiating mood hidden in the epistemic modality sequences, precisely, the kind of structures demonstrated in epistemic sequences when they are writing/uttering something.

2.5 Hedging and other Related Features of Academic English and their Very Realization in Corpus Concordancing regarding Epistemic Sequence

Snow&Uccelli(2009) argued that academic English has the characteristic of integrating academic content into English, so its language usually contains highly information-intensive lexical forms and a large number ofcomplex sentence patterns. Academic English has many characteristics different from daily colloquial English in interpersonal standpoint, information capacity and information organization. Vocabulary selection alone is manifested in variable vocabulary, formal expression, accurate meaning, abstract concept, natural connection in structure, clear relationship and logical coherence.

Hyland(2017:26) summarized academic English into eight aspects:

- Be more explicit about its structure and purpose;
- Use more citations to support arguments;
- Focus on actions rather than actors;
- Use fewer rhetorical questions than students tend to use in school essays;
- Be generally intolerant of digressions;
- Be cautious in making claims;
- Package processes as things;
- Spell out steps in an argument and connections between sentences very clearly;

It can only be explained in some ways that using nouns and abstract nouns as subjects in particular, adverbs of degree as adverbial, logically clear subordinating conjunctions is characteristic of academic language concerning lexical and syntactic levels. Academic language is derived from colloquial language, compatible and apt to abstract language and figurative language but different from these language structures because it relates not only learning phases but also to subject matters (Fang, 2016:325). For example, natural science language contains mostly very professional technology information and scientific discovery knowledge, so its language tends to use phrases related to figurative language. However, human and social sciences texts tend to be abstract or colloquial language because they are more related to people's real life (Halliday&Martin,1993). Academic writing in English is linear, which means it has one central point or theme with every part contributing to the main line of argument, without digressions or repetitions. Its objective is to inform rather than entertain. As well as this, it is in the standard written form of the language. There are ten main features of academic writing that are often discussed. Academic writing is to some extent: complex, formal, objective, explicit, hedged, and responsible. It uses language precisely and accurately. It is also well organized and planned(Gillet,A.,2009).

There is no need for all these particular features to be discussed here, our focus is on the features related to modal sequences. By hedging, it means it is necessary for any academic writing authors to make decisions about their stances on a particular subject, or the strength of the claims they are making. Actually,different ways can be employed to do this,for example, we can use: (i). Certain modal verbs, e.g. will, must, would, may, might, could.(ii). Adverbs of frequency, e.g. often, sometimes, usually.(iii). Modal adverbs, e.g. certainly, definitely, clearly, probably, possibly, perhaps, conceivably. (iv).Modal adjectives, e.g. certain, definite, clear, probable, possible, and other structures like "That clauses",e.g. It could be the case that... . It might be suggested that There is every hope that and To-clause + adjective,e.g. It may be possible to obtain. It is important to develop.... It is useful to study etc. to express what you are to claim by hedges like the above- mentioned phrases and structures. Here, in relation to modal verbs, we would like to focus on the availability of the retrieval of some of the epistemic sequences for the purpose of the corpus concordancing. For example, such modal verbs as will, must, would, may, might, could ect. and the above-mentioned adverbs of frequency and modal adverbs,modal adjectives, and such modal nouns as assumption, possibility, probability, and some clauses/structures can be turned into the searching words of syntaxes or regular expressions for the realization of the researching of some of the academic features.Animated subject before and dynamic/ stative verbs and continuous and perfect tense patterns after modal verbs can also changed into searching words of syntax or regular expressions,which is in relation to the academic features, i.e. "Focus on actions rather than actors"and "Package processes as things"as is claimed by Hyland(2017). Here,we narrow down to only expressions related to epistemic and deontic modality structures as in Table1. Precisely, some modality structures above-mentioned except others for later exploration.

2.6.Analogue epistemic   sequences and their significance of academic tendency
24 epistemic & analogue epistemic sequences (12 epistemic sequences, 12 analogue epistemic sequences) includes:CAN RR VVI, CAN VBI VVN, COULD XX VVI, NN1 CAN VBI, NN1 CAN VBI, NN1 SHOULD VBI, NN1 SHOULD VBI, NN2 CAN VBI VVN, NN2 CAN VBI VVN, NN2 SHOULD VBI, NN2 SHOULD VBI, NN2 SHOULD VBI VVN, PPH1 CAN VVI, PPH1 CAN XX VVI,PPH1 COULD VVI, PPH1 COULD XX VVI, PPHS2 CAN VVI, PPHS2 CAN XX VVI, PPHS2 COULD VVI, PPHS2 COULD VVI, PPHS2 COULD VVI, PPIS2 CAN RR VVI, PPIS2 CAN RR VVI, PPIS2 CAN XX, PPIS2 CAN XX VVI, SHOULD VBI VVN was extracted from WECCL1.0&2.0 based on the CLAW7 coding technology based on the prediction of the relationship between syntactic structures and modality (see Table1)by Biber et al.(1999), Coates (1983), Warnsby (2003) et al. The epistemic sequences in learner corpus were not perfect and many of them could not be found.They are not so comprehensive but typical of students' argumentative articles since they are extracted after many procedures such as segmentation, coding and retrieval. Hence the term "analogue epistemic sequence".They are so-called because they , in my opinion, refer to a third modal sequence similar to epistemic sequence, which does not completely conform to the previous standard but resemble epistemic sequence in a way and must be a proper supplement to the previous research. For example: CAN RR VVI, COULD XX VVI should be epistemic sequence according to "VM+epistemic adverbs" structure although we cannot determine whether XX and RR,the

negative(negator:not&n't) and general adverbs, especially when they are behind some epistemic modal verbs, usually belong to epistemic adverbs(long ,2014)although CLAWS cannot distinguish between epistemic adverbs and general adverbs, so we consider it also belongs to epistemic sequence. However, because it is not complete and can't meet the aforementioned standard, it is called analogue epistemic sequence.In addition, "NN1/NN2 CAN VBI" and "NN1/NN2 SHOULD VBI" are only one VVN different from "NN1/NN2 CAN VBI VVN" and "NN1/NN2 SHOULD VBI VVN" (They are in the "animated + Subject +VM" and "VM+perfective verbs" structure). Likewise, they belong to analogue epistemic sequence because previous probes found that there was a high probability of nouns or adjectives(89%) after "NN1/NN2 CAN VBI" and "NN1/NN2 SHOULD VBI" (Long,2015), so they conform to people's cognition that the identification of the judgment and evaluation of the truth and falseness of proposition is the key consideration for epistemic sequence. In addition, "PPH1/PPHS2 +CAN /COULD+VVI" is only XX(negative adverb/negator:not&n't) different from PPH1/PPHS2 +CAN /COULD+XX+VVI, which belongs not to "animated + Subject +VM" but to "VM+epistemic adverbs" structures, and is similarly identified as analogue epistemic sequence .Similarly, "PPIS2+ CAN +RR/XX+ VVI" (belonging not to "animated + Subject +VM" but to "VM+epistemic +VM" structure), was also agreed on as analogue epistemic sequence because the subjects such as PPH1/PPHS2 and PPIS2 are commonly found in academic writings as they are in descriptive mood or usually referred to authors himself/herself/themselves.Such structures as "there+ VM+ be" , "guide words +VM" and "VM+progressive verbs" are not so common here in students' argumentation corpus, so they are here ignored. To this end, according to theoretical reasoning and practical proof of the relationship between epistemic sequence and academic tendency described (long, 2015,2020), it can be basically concluded that the more frequent the occurrence of epistemic sequence or analogue epistemic sequence, the stronger the logical of the text, or, the more academic it may be in the case of general argumentative essays. Of course, other registers(e.g.narrative and expository writing) need to be verified and discovered.

2.7 Previous Typical Researches on Modal Sequences in China and inspiration for further study
Although few overseas scholars did research key-wording as or titled as "modal sequence/s",quite a few researchers in China did (e.g. H.Liu, 2004; G. Ma et al,2007; M.C Liang,2008; J.C.Xie,2009; W.Y.Liu,2009; Author,2011,2012,2013,2014,2016; X.N.Chang ,2016; J.Chen，2017; A.B. Chen, 2012; S.Feng, 2017; Pang , et al.,2018; H.H. Wang ,2015; G.Wang ,2013; K.Wei,2009. J.H..Wu ,2016; P.F.Yan ,2017; J.Yu,2016; L.Z.Zhao,2018; H.Q. Zhang,2013; J.Zhang,2013; Z.H Zhang.,et al.,2016; X.Y. Zhang,2015; Yang,2018,etc.). And they obtained different findings from the different corpus they used. Such researches on modal sequence in China can be roughly divided into several parts to summarize, falling into two categories: non–academic and academic, from the perspective of RC the researches employ and similarities and difference they share among the same RC users although their OC are often in different form from one to another. Usually, the researchers have their own settings when having to use different software to help .i.e. different ways to set the settings.e.g. span of clusters or searching expressions.

    This literature review targets such research papers key-wording or titled as only "modal sequence" to summarize the findings in the way mentioned above.
(1)   With almost the same reference to LOCNESS,CLEC/WECCL(1.0&2.0) as OC, researchers almost achieved almost the same findings regarding non-academic features although they have their some unique findings due to their different OC or settings,etc. Findings(Liang, 2008; Liu ,H. ,2004, Ma ,G. et al,2007; Tang ,2013; Wang,2013; Author,2014,2016) reveal similarities like these:(i)overused personal pronouns before, and underuse verb forms with unmarked aspect and voice after modal verbs;(ii) can, will and should that teaching materials earlier involved are used too much, while epistemic modal verbs expressing politeness, gentleness, subtlety, such as might, could and would are too less used; (iii) learners overuse deontic modality and underuse epistemic modality;(iv) language proficiency affected the use of modal sequences.
(2)   With almost the same reference to BNC, CLEC/SECCL/TECCL as OC, researches had almost the same findings regarding non-academic features yet they have their unique findings due to their different OC or settings,etc. For example,with the same reference corpus BNC (oral corpus:X.Y.Zhang (2015),spoken monologue:A.B. Chen (2012)) ,with SWECCL2.0 as observation corpus (including oral tests from TEM4 and TEM8:Zhang)and SWECCL 2.0 [TASK3 (talking on a given topic) and TASK4(making the comment on a given topic): Chen], and the width of the cluster set as (2-8) but no mention of modal verbs :(X.Y.Zhang ,2015) concorded, researches had almost the same findings regarding non-academic features:(i)The occurrences of " animate subject + modal verbs" were excessive, (ii) verb forms with unmarked aspect and voice after modal verbs are often used, c).passive voice and perfect tense were less used after modal verbs; (iii) Structures formed by "modal verbs+negative abbreviations" and interrogative sentences formed by modal verbs + second-person pronouns are rarely used; (iv) modal sequences of modal verbs + stative verbs (Be, Have) are used less frequently and (v)modal sequences beginning with could and would are underused.
(3)   With the same reference corpus to LOCNESS and BNC, and with CLEC and SECCL1.0 (Zhang,2013)/CLEC

(st5 and st6): (Liu, 2009) / TECCL: (Feng, 2017) as the observation corpus, and the width of the word cluster set as (2-8) and 17 core modal verb variants (J.Zhang, 2013)/the width of the word cluster set as 2-3, and 13 core modal verbs (liu 2009)/the width of cluster set as (no mention),and 13 core modal verbs (Feng,2017)used, their findings share similarities like these :(i) overuse modal verbs can, should, will, ought to, need, must, shall and used to, while would, could and might are underused; (ii) use the first and second personal before and dynamic verbs after modal verbs with unmarked aspect and voice ; (iii) seldom use past tense and   the modality sequences guided by the past tense especially with the passive voice ; (iv) English majors overuse deontic modal sequence except for the overusing of modal sequence of "People + Modal Verb" and tending relatively to be close to the native of English speakers regarding the use of epistemic modality; (v) more likely to use such modal sequences as "modal verb + dynamic verb with unmarked tense or aspect" and "animate subject + modal verb";(vi) tend to use personal pronouns I, we, you, they, he/she before modals, avoiding those modal sequences such as "modal verb + verb with perfect aspect" or "the guide word of subject + modal verb";(vii)the frequency of English learners' use of the deontic modality in their written language is related to their language proficiency.

(4)   By referring to the students' presentations and colloquium in the oral presentations section of the MICASE (academic oral corpus), with the oral test section of TEM4 and TEM8 from SWECCL2.0 as observed and the width of the cluster set as (no mention) and no mention of core modal verbs used, Chang, X.N (2016)found that can, should, will, could, must and may are overused, whereas would and might are underused, as can be seen from every research, even in academic research articles, only some modals are missing and the order is somewhat different. Results show five obvious characteristics in Chinese English majors' oral English: (i) The first person plural pronoun and the third person plural pronoun are overused before modal verbs, and after unmarked aspect and voice are excessively used,as can be seen from Liu (2009),only the third personal pronoun was replaced by the send personal pronoun; (ii) When the modal verbs collocate with negative word(not),learners tend to use "modal verb + not",the same as in c)in Chen (2012) ; (iii) Modal sequences beginning with would and might are less used, and after which the verbs with unmarked aspect and voice are frequently used, as can be seen in Zhang(2013), Feng(2017), Author(2014,2016) and Tang (2013);(iv) Interrogative sentences made up of modal verbs and personal pronouns are obviously less used, especially "modal verb+ you"; The sequence of "people + modal verb" is also frequently used. Besides, after comparing with the two corpora of oral tests in TEM-4 and TEM-8, the conclusions are: i) As for modal verbs, low-proficiency learners use more could and would than high-proficiency learners; ii) In terms of modal sequences, high proficiency learners use more stative verbs after modal verbs, and not just perceive personal pronouns as subjects before modal verbs; iii) All learners employ more deontic modality, but the use of epistemic modality also assume a definite proportion; for learners of higher English levels, more epistemic modality has been used, which is what makes it special and characteristic of English majors as compared to MICASE .

With the same category of OC concerning Chinese majors, they seem to share similar feature. With reference to applied linguistics academic papers corpus (ALC) (2005-2009) from applied linguistics, the Chinese applied linguistics academic papers corpus(CALC) (2010-2011)are observed. Cluster width set as 2-8 set, 8 modal verb core (can, could, will, would, should, may, might , must) employed, J. Zhang (2013) found that: (i)should is overused while can is underused in the Conclusion section of CALC. may, could, might, would, will and must do not have significant difference (p > 0.05) in the two corpora.(ii)the proportion of low modality value in ALC is higher than that in CALC, while the proportion of high modality value is lower than that in CALC. (iii) Chinese scholars tend to use "VM + stative verbs" and "inanimate subject +VM" to convey epistemic modality. There are some characteristics of modal sequences used by Chinese scholars. (i)The majority of Chinese scholars avoid using some complicated patterns of modal verbs, such as"VM + deontic modality adverb", "VM +perfect aspect" and "It +VM". (ii)They usually use students, teachers, researchers and learners as the subject of modal verbs in the pattern of "animate subject + modal verb". (iii) Low value modality occupies the largest proportion in ALC.The paper authors used low and median value modal verbs which expressed euphemistic mood to convey their attitudes and ideas to readers. The proportion of high value modal verbs takes up 5.2‰ in CALC, higher than that in ALC. (iv)Chinese scholars tend to use the modal sequences of "VM + stative verbs" and "inanimate subject +VM" to convey epistemic modality in CALC,as   can be seen in Yu's finding(2016).

Even under academic category, using both the academic RC and OC, researchers can probably obtain the same findings, as is shown as follows:

(1)   By referring to BAWE, with the academic corpus of science and engineering learners as observed and with 10 core modal verb (can, could, will, shall, should, may, must, need)employed and the width of the cluster set as 2-3, Peng (2017) found that learners use more deontic modal sequences than epistemic modal sequences in science and engineering, as can be seen in Chang(2016)and Liu (2009) and Author (2014,2016), Chen (2012),and Yang (2018).More specifically, learners use more personal pronouns subjects before and dynamic verbs with unmarked aspect and tense and fewer epistemic modal adverbs modifying after modal verbs. In addition, learners seldom use attributive clauses containing modal verbs and THERE BE patterns to annotate

or evaluate the propositions mentioned above.

(2)   With JDEST as reference, the self -built academic corpus of Chinese soft and hard subjects as observation. And with the width of the word cluster set as 2-3 and no mention of core modal verbs employed, Yu(2016) finds that: Chinese scholars tend to overuse can, could, should and underuse would, may, must, should. However, there are some exceptions: As for modal sequence, they tend to overuse "modal verbs+stative verbs", and underuse "animate subject+modal verbs" to avoid modal sequence indicating deontic modality, which somewhat satisfied the academic writings' requirements of being objective and polite.

Luckily, we can see some others mentioning Chinese research abstracts.With reference to English research articles abstracts, Chinese research abstracts as observation, with no mention of the width of the word clusters set and core modal verbs used, Wei (2009) found: in both English and Chinese abstracts, simple themes dominate; ii) The usages of textual themes in English and Chinese abstracts are almost the same. Conjunctive adjuncts dominate in both English and Chinese abstracts; iii) Unmarked themes dominate in both English and Chinese abstracts, and both English and Chinese abstracts prefer object of research and epistemic nouns; iv) The marked themes of manner are used more frequently in English abstracts and the marked themes of angle are used more frequently in Chinese abstracts; v)as for TP patterns, Both English and Chinese abstracts use parallel progression pattern and continuous progression pattern more frequently. Inspired by previous studies, this research makes a comparison between the usage features of modal sequences constructed with modal verbs in the Chinese English major' argumentative essays collected in corpora WECCL2.0 (2003~2007) and the natives. Since the data covers writings produced during exams and compositions respectively with restriction of time or resources, this research, referring to the BaREnLoB(consisting of :1. BAWE, 2. RA, 3. ENS, 4. 5. LOCNESS, BNC essay respectively), would probably reveal further how modal sequences are used in Chinese English majors' argumentative essays, and thus confirm the usage features of modal sequences among Chinese English learners, with the hope of providing theoretical support for the teaching and learning of modal verbs and thus improving the acquisition of modal structures and modal meaning.

The authors here aim mainly to conduct an inquiry of the overall modal sequences used by Chinese English majors and the natives through CIA. What is really new about this study is that we employ a new reference corpus, which consists of five different corpus seemingly mainly of different layers of academic levels ,and we are going to find whether there exist difference between modal sequences produced by the reference corpora as compared with the previous findings,and above all,whether there are some academic features hidden in the overall features.

3.Research Design

Following previous research and their methodologies, this study is designed as follows:

3.1 Research Questions

Based on the analysis above, this study tries to address the following questions:

(1)   What are the academic features among the non-academic regarding modal verbs?

(2)   What are the academic features among the non-academic regarding modal sequence in terms of some epistemic modality structures. i.e"modal+stative verb","unanimated subject+modal"(e.g.plural noun/absract noun+modal)], and some deontic modality structures. i.e "animated subject+modal" (e.g."we/you/ I+modal"and "he/she/they+modal")?

3.2 Corpora Used and Their Pre-processing

This research employs the sub-corpora of Chinese English majors' argumentative essays: STU1 in WECCL2.0 (Wen Qiufang et al, 2004) as the observation corpora and BaREnLoB as the reference corpora.

(1)   Observation corpus(OC): the observation corpus used in this study is derived from SWECCL2.0 (Wen Qiufang et al., 2004). Spoken and Written English Corpus of Chinese Learners 2.0(SWECCL2.0),which was published in the year of 2008 and conducted by scholars like Professor Wen Qiufang, Liang Maocheng and Yan Xiaoqin. This paper will only employ its sub-corpus(WECCL) as our observation corpus. There are originally 4950 compositions of university students collected from freshmen to seniors. However, in this study, the token is recalculated in "tokens used for wordlist" in statistics embedded in the "Wordlist" in Wordsmith Tool and number of composition is also subject to the marked "sections".WECCL2.0 has 3,027 instead of the supposed 4,359 argumentative essays, and files like a2105nd.cls and a341xsd.cls, etc. are damaged when processed by Wordsmith, so only 3,027 argumentative essays are available to this study. Table 2 shows the specific information about Written English Corpus of Chinese Learners 2.0 (WECCL 2.0).The participants of the writings are from 34 universities in China. Those compositions were originally written on test papers, and then input into computers without alteration. It's up to the researchers to choose among Chinese non-English majors or Chinese English majors according to their research purposes. Also, they are free to choose the subjects of different grades. In this corpus, their writing levels are diversified in accordance with different grades of the students.

(2) Reference corpus(RC): In order to make a comparison between the relationship among different sequences and their different academic levels of semantic modal meanings between Chinese English majors and the natives, the study uses the different degrees of academic corpus as reference corpus, they are : 1. The BAWE, 2. RA, 3. ENS, 4. 5. LOCNESS, BNCessay respectively,so it is named BaREnLoB.

BAWE, created by Brookes, Oxford university, the University of Reading and the University of Warwick in 2004-2007, which includes 35 sub disciplines under four categories : humanities, life science and physical science and social science,written by undergraduate or graduate students, consisting of 2761 academic texts, a total of 6, 506, 995 tokens. In this study,2,062 texts of English academic corpus (BAWE) were randomly selected, totaling 5,251, 926 tokens.Obviously ,they are intuitively academic.

RA (abbreviated for Research Articles):It is our collective self-compilation. Ranging from 2002 to 2014,

39 articles were extracted from three international journals:Applied Linguistics(AP), Studies on Second Language Acquisition(SL) and System(SYS). The selected articles were named respectively from ap01-ap13, sl01-sl13, and sys01-sys13. The authors are mostly from English-speaking countries. The subjects are all about applied linguistics, belonging to Arts and Humanity.Undoubtedly ,they are intuitively academic.

ENS belongs to ICNALE (International Corpus Network of Asian Learners of English),which is an international corpus among Asian learners of English. ICNALE has all of the text based on the composition of two given topic: "It is important for college students to attend a part-time job" and "Smoking should becompletely banned at all the restaurants in the country".The length of the composition is 200 ~ 400 words, and the time is 20 ~ 40 minutes. In order to facilitate the comparative analysis, the corpus also collects the writing works of native English speakers from the United States, Britain, Australia, New Zealand and Canada. This study combines the corpus of English native speakers (ENS1 and ENS2), called ENS in this paper,strictly speaking they are between, so named here quasi-academic corpus, and the feature they display is called quasi- academic feature intuitively.

LOCNESS:(abbreviated for Louvain Corpus of Native English Essays)..It includes brsur1,brsur2, brsur3, USARG.,usmixed5 large files (297 small files), 149,574 words of argumentative essays written by American university students 18,826 words of literary-mixed essays written by American university students,and 95,695 words of argumentative and literary essays written by British university students ,and 60,209 words of British A-level argumentative essays(here we have no A-level texts). Nearly all are argumentative essays on different topics related to culture tradition and some literary commentary. Besides some historical expository essay,topics are everything related to their culture life (e.g. euthanasia ,controversy in the classroom ,capital punishment, values and consequences of school interaction, Adolescent suicide, gender roles in our society, Premarital sex, Talk shows, Great inventions and discoveries of 20th century and their impact on people's lives, An Aspect of Studying Ethnic American Literature, and Aspects of Social Psychology, etc.). Timed or untimed, and no reference tools is used, as a consequence, strictly speaking ,they are intuitivelyquasi-academic.

BNC-essay (the university and school essay section based on BNC Indexer,hereafter BNC-essay) is topic discussion papers by the British and American university students. It here contains 10 texts: H8W, HD5, HDA, HPG, HPS, HSC, HUB, KA1, KA3, KAY. They are almost essays (including schoolboys/girls' essays, course essays,unpublished essays, misc unpublished university essays,academic essay etc.),and some English literature papers. As for school essays, they mixture of expository, argumentative, narrative and imaginative essays, but mostly English literature essays (expository & argumentative).But for university essays,they are all on history, politics, computer science; organizational behaviour; psychology lab reports; lecture notes. Consequently,they are intuitively quasi-academic.

As can be seen from above, the reference corpora in this research reflects the following characteristics: 1. total amount is large, almost 6 times larger than that of observation corpus(WECCL2.0), which is in a good accord with international norms and standards (Berber – Sardinhat,2004), 2. Academic.BAWE, BNC essay and self-built RA are typical academic corpus, and the others are almost quasi -academic.3.unbalance and homogeneous whether in the size or the genre. However,where and what is the indicator to define the exact difference between these corpora regarding academic and non academic norms?And whether there exist some academic features among the overall features is the focus of the study.

According to Mike Scott's research (2016), it is not related to the size of reference corpus, as well as the strangeness and normality regarding keyword generation. There is no such thing as a bad reference corpus as for the precision value1 of keywords. Reference corpus can be a mixture of different genres, and the bigger ,the better, of course within the standard: "a moderate sized RC may suffice" (). At the same time, his research also concluded that, depending on the corpus text, the keywords will be different, but the situation is complex and involves more than one thing. Regarding the genre of reference corpus of this study, they are only about the academic and general arguments, a mixed bag of only two kinds of grains .And they are basically in recent 20 years in time, basically about issues of daily classroom or academic topics concerning university graduates and undergraduates, so they will produce homogeneity of keywords,with its thematic data basically credible and effective.

To observe the use features of modal sequences, the research has built its own corpus of Chinese English majors' argumentative essays by extracting all tagged argumentative essays with the title of STU1, extracted from

WECCL2.0 by a subcorpus generator. The sub-corpus of the natives' academic or argumentative essays of BaREnLoB is employed as the reference corpora as can be seen from below . More detailed information on the corpora is shown in Table2.

Table 2 .Basic Information of the Corpor

Corpora	No.of tokens	Percentage	No.of texts	Percentage
OC: STU1 in WECCL2.0	1,005,607	14.14%	3,027	51.94%
RC: BaREnLoB	6,107,989	85.86%	2,801	48.06%
TOTAL	7,113,596	100.00%	5,828	100.00%
RC Including 1):BAWE	5,251,926	85.98%	2062	73.62%
2)ENS	88,458	1.45%	393	14.03%
3)BNC(essay)	202,720	3.32%	10	0.36%
4)LOCNESS	263,963	4.32%	297	10.60%
5)RA	300,922	4.93%	39	1.39%
TOTAL	6,107,989	100.00%	2801	100.00%

To better observe the syntactic features of the modal sequences in the above-mentioned corpora,we,consulting the processing methods from Arts & Granger (1998) and To-no (1999), dealt with the tagged texts with Power Grep. The results are exemplified below in Table 3.

Table 3. Exemplification of Text Preprocessing

	Active	Passive
Original text	We should set the standard carefully.	The standard should be set carefully.
Tagged text	We－PPSI2 should－VM set－VVO the－AT standard－NN1 carefully－ RR.	The AT standard－NN1 should－ VM be－VBO set－VVN carefully－RR.
Processed text	PPSI2 should VVO AT NN1 RR.	AT NN1 should VBO VVN

Following this procedure, this research then processed the data with WordSmith Tools like Concord, Wordlist, and Keyword so that all modal sequences in the processed texts can be easily extracted and observed. To be effective in measuring the effect, we employed ELL(effect size of log-likelihood) as effect size index,ELL varies between 0 and 1 (inclusive). As Johnston et al. (2016) said, "(its) interpretation is straightforward as the proportion of the maximum departure between the observed and expected proportions". Being a function of the data χ , the likelihood ratio is therefore a statistic. The likelihood ratio test rejects the null hypothesis if the value of this statistic is too small. How small is too small depends on the significance level of the test, i.e., on what probability of Type I error is considered tolerable ("Type I" errors consist of the rejection of a null hypothesis that is true). The numerator corresponds to the likelihood of an observed outcome under the null hypothesis. The denominator corresponds to the maximum likelihood of an observed outcome varying parameters over the whole parameter space. The numerator of this ratio is less than the denominator. The likelihood ratio hence is between 0 and 1. Low values of the likelihood ratio mean that the observed result was less likely to occur under the null hypothesis as compared to the alternative. High values of the statistic mean that the observed outcome was nearly as likely to occur under the null hypothesis as the alternative, and the null hypothesis cannot be rejected.

4.Procedures of the Analysis

4.1Analysis of Modal Verbs and Modal Sequences

In order to conduct a keyword analysis, this research, first, by using such words (+tagsets) e.g."can, ca, may, must, will,*'ll, wo, shall, sha, could, might, would*'d, should, need, use *K, oughtK, dare" to search in STU1 in wecc12.0, we created wordlists from the pre-processed corpora, then, by using the Wordlist in WordSmith Tool, we calculated the use frequency of all modal verbs here mentioned (occurrences per 100 words, as defaulted in WordSmith), and then examined the results to see which modal verbs in the observation corpus are overused or underused.

Then, in the same way ,we built a list of clusters of 3~8 width of span for each of the two corpora(as defaulted in this study), then compared these two lists by means of the KeyWords Tool in WordSmith Tool by the index, and uncovered all clusters that are overused or underused in the observation corpus, i.e., clusters with plus or minus keyness. Since the corpora have been pre-processed, the keywords clusters being analyzed bear the form of tag sequences. For example, the tag sequence "should VBO VVN" refers to the modal sequence constituted by the modal verb should plus the passive form of a notional verb (be + past participle), like should be canceled. Therefore, the keywords analysis actually reveals how different syntactic structures in the corpora are used.

To calculate all modal sequences in the observation and reference corpora, the research first found all keyword clusters that start or end with modal verbs, i.e. "tag 1+ tag 2 …+modal verb" ,and "modal verb + tag 1+ tag 2 + …",and then classified them by the varied forms of the modal verbs, and in this way we revealed all the overused and underused modal sequences and found the difference between the observation and reference corpora regarding

academic levels.

4.2 Identification of the Modality of Modal Sequences

In line with the syntactic structures of deontic and epistemic modal sequences shown in Table 1, this research grouped the overused and underused modal sequences into two types. For example, the sequence PPSI2 should (e.g.,we should) conforms with the syntactic structure of "animated subject + modal verbs" and therefore is classified into the category of deontic modal sequence, while should VBP (e.g.,should be) conforms with "modal verbs + stative verbs" and thus is classified as an epistemic modal sequence.

4.3 Procedures concerning Stative/verbs Identifying

Such procedures are precisely presented below the related titles in the study.Here is the generalization. With the typical modal verbs(with its tagged sets) as search terms (span set as certain length), after searching in the tagged WECCL2.0 , we saved the result as EXCEL spreadsheets, deleted unrelated verbs(leaving some numbers of columns behind), then conduct a counting based on the verb frequency behind the modal verbs.After frequency-data processing via Pivot table, we set R1 as the center and conduct order- processing in a descending way first, then R3 and R5 are added, and the frequency of the verbs after the modal verbs are counted after the order-descending processing. Almost in the same way we treated the corpus for the Animate Subjects research but the settings are in the left side of the center.

4.4 Procedures concerning the modal verbs after the abstract nouns

Abstract nouns from the top 100 plural nouns cab be picked up after we put NN2 in Concord in Wordsmith Tool, and then we put the abstract nouns along with the 13 modal forms of in the Concord again to find the modal verbs after the abstract nouns in both the observation corpus (henceforth Oc)and the reference corpus(henceforth Rc).

5.Research Results

5.1.Some non-academic and academic features of Oc against Rc.

5.1.1 Use Features of Modal Verbs

This research calculated the relative frequency (occurrences per 100 words) of all the modal verbs in these corpora. Results show that Chinese English majors show a strong feature to overuse modal verbs in their English writing. Though the use frequency of shall, need, ought to, dare, might and used to only show unremarkable difference (partly because their use frequency is relatively low), the use frequency of should, can,ca(deleted form of "can't") in the observation corpus is significantly higher than that in the reference corpus, and the learners' use frequency of could and would is much lower than the natives'. The specific results are shown below in Table 4.

Table 4. Overused and Underused Modal Verbs

Rank	Keyword	NNS		NS		Keyness	Sig	ELL
		Occur	Freq (%)	Occur	Freq (%)			
4	SHOULD	7069	0.64	7585	0.13	8536.88	0.00	0.00
5	CAN	11930	1.08	21279	0.36	8168.82	0.00	0.00
11	CA	1498	0.14	227	/	4298.17	0.00	0.00
34	COULD	1043	0.09	8136	0.14	-136.48	0.00	0.00
45	WOULD	1252	0.11	13445	0.22	-652.90	0.00	0.00

As Table 4 shows, should, can, and ca(deleted form of "can't")(all classified hear as Type D Modal Verbs: short form of 'Deontic Modals' henceforth,) are highly overused, among which should stands out with its extraordinary use frequency between the two corpora (being 0.64 and 0.13 respectively). ELL here means for the study that the effect here is so low,so unstable because low values of the likelihood ratio mean that the observed result was less likely to occur under the null hypothesis as compared to the alternative. It is somewhat consistent with the findings of Liu, H. (2004), Ma, G. et al.,(2007) ; Liang, M.C., (2008); Xie, J.C., (2009); Author(2014, 2016); Liu, C.Y., (2013); Liu, W.Y. (2009); Feng, S.(2017), Chang, X.N. (2016); Zhuang, X.Y., (2015); Wang, Y. (2018); Tang, L.L. (2013); Wang, G. (2013); Chen, A.B., (2012); Yan, P.F.,(2017); Yu, J., etc, which suggest that English learners in China face similar problems with their Swedish counterparts as reported by Aijimer (2002). And the result is consistent with Liang 's(2008) study on Chinese non-English majors, except that some modals will should (Liu , 2009) may, must or will may or will, may must need, ought to (Author's2016)or will, ought to, need, must,shall,used to (Feng,2017),or will, could ,must, may(Chang,2016)or will (Yang, 2018)or will ,must(Qiao et al,2014)and the rank of these keywords are somewhat different,which, in a way, means that learners with the same mother tongue face basically the same problems in their English acquisition. Could and would (classified here as Type E Modal Verbs: short form of 'Epistemic Modals',henceforth) are clearly underused in the observation corpus, and the use frequency of would is only 1/2 that of the reference corpus (Liang, 2008),and somewhat consistent with Author (2014),Tang (2013), Zhang, J. (2013) , Zhang, H.Q. (2015),Yu (2016), Feng (2017),Chang (2016)and Yan (2018).However,ca(deleted form of "can't")are here highly overused, which

can also account for some academic trends in learners'wrtitings, as Zhang, Z.H. , (2016) said, in the "we + weakened modal verb(Zhang, Z.H.,2016) + not/n't "pattern, the co-occurrence of we and can is the most, followed by the co-occurrence of could on the whole. This is what this finding reveal, as is evidence in Table 9 and Table5 and Table 12.

5.1.2 Features of Modal Sequences

Via a Keywords Analysis on the list of clusters drawn from the two corpora, this research has produced several keyword clusters, from which all modal sequences are picked out. Table 5 shows the top 20 modal sequences with the highest plus keyness. which are the most overused modal sequences in the observation corpus.

Table 5 Most Overused Modal Sequences in the Learner Corpus (top 20 keywords with the highest plus keyness)

Rank	Modal sequences	NNS		NS		Keyness	Sig	ELL
		Occur	Freq (%)	Occur	Freq (%)			
1	CAN VVI APPGE	656	0.059331991259			1435.09204	10.000000.	00004
2	SHOULD VVI TO	458	0.041423861109			1183.72876	0.000000.	00004
3	SHOULD VVI TO VVI	436	0.03943406897			1147.24182	10.000000.	00004
4	CAN VVI AT	800	0.072356082846		0.01414641697	9.09594730.000000.		00003
5	SHOULD VVI AT	482	0.043594543237			959.80181880.000000.		00003
6	NN2 SHOULD VVI TO	297	0.02686219725			936.22332760.000000.		00004
7	CAN VVI AT1	603	0.054538399468			931.70080570.000000.		00003
8	CAN XX VVI	879	0.0795012491094		0.01829335792	5.65783690.000000.		00003
9	NN2 SHOULD VVI TO VVI	291	0.02631952623			924.52862550.000000.		00003
10	SHOULD VVI DAR	263	0.02378706318			849.55340580.000000.		00003
11	CAN VVI PPIO2	270	0.02442017934			801.72766110.000000.		00003
12	PPIS2 SHOULD VVI AT	233	0.02107371 15			757.57696530.000000.		00002
13	CAN VVI APPGE NN1	340	0.030751336128			757.56884770.000000.		00003
14	SHOULD VVI DAR NN1	222	0.0200788149			751.84576420.000000.		00003
15	CAN RR VVI	779	0.0704567361088		0.01819302772	6.83911130.000000.		00002
16	CAN PPIS2 VVI	255	0.02306350346			705.67321780.000000.		00003
17	NN2 SHOULD VBI VVN	490	0.044318102438			685.70214840.000000.		00002
18	NN2 SHOULD VBI VVN TO	247	0.02233994246			678.63989260.000000.		00003
19	SHOULD VVI APPGE NN2 SHOULD VBI VVN TO	280	0.02532463 91			657.86950680.000000.		00003
20	VVI	237	0.02143549 44			651.62634280.000000.		00003

The top 20 most overused modal sequences can be divided into two types according to the syntactic constituents before or after the modal verbs. The first type is in the form of "subject + modal verbs." Specifically, they are: 1) sequences constituted by the plural first person pronoun (we) + modal verbs should; 2) sequences by plural noun phase (NN2) + the modal verb should.In short, the most overused modal sequences in the form of "subject + modal verbs" are "NN2 +should,"or "we + should". The second type is in the form of "modal verbs + verbs." Under this category, they are: 1) sequences constituted by modal verbs can, should + verbs with unmarked tense and aspect; 2) sequences by can /should+ verbs with unmarked tense and aspect + qualifiers (e.g.,the ,a or your)/(e.g.,more ,few, less)/(us), or by can/should + verbs with unmarked tense and aspect + complex objects, or by can + general adverbs/negative adverbs +(/verbs with unmarked tense and aspect), should + verbs with unmarked tense and aspect + to (+ verbs with unmarked tense and aspect). In summary, the most overused modal sequences in the form of "modal verbs + verbs" are "can (should/) (+ general adverbs) + verbs with unmarked tense and aspect (+ qualifiers :e.g.,the /a / your) /(us + verbs with unmarked tense and aspect)" and "should + verbs with unmarked tense and aspect + to (+ verbs with unmarked tense and aspect)."

Based on the results above, we see that Chinese English majors employ modal verbs in a different way from the argumentative essays found in the reference corpora. The Chinese students usually incorporate agents (particularly NN2 or we) as sentence's subject, and then discuss the responsibility and obligation of their respective student community. As expected, it is appropriate that they put NN2 before modal verbs, like should, which increases author's objectivity and thus maintains the arguing effects,which is different from the findings from Author's finding(2014), and somewhat consistent with Chang's(2016) and Feng's, Zhang's(2013), Yu's(2016)(2017)studies .And the effect size(ELL) here means for the study that CAN VVI APPGE,SHOULD VVI TO, SHOULD VVI TO VVI, NN2 SHOULD VVI TO are relatively more likely to occur under the null hypothesis as the alternative than PPIS2 SHOULD VVI AT,CAN RR VVI,NN2 SHOULD VBI VVN,and the null hypothesis cannot be rejected.

Different from Chinese non-English majors(Liang,2008), however, Chinese English majors are capable of adopting plural nouns before modal verbs as the sentence's subject in order to discuss the responsibility and obligation of groups they are not part of. This process is similar to that of the natives. At the same time, they rarely put I before the aforementioned modal verbs, a sign that they do have a better acquisition of the epistemic modality

than do Chinese non-English majors, another sign that they are getting close to a trend of being academic, saying things in a negotiatory way. While some modal sequences are overused, others are underused among these majors. Through Keywords Analysis, this research reveals the top 20 modal sequences with highest minus keyness (see Table 6).

Table 6 Most Underused Modal Sequences in the Learner Corpus (top 20 keywords with highest minus keyness)

Rank	Modal sequences	NNS		NS		Keyness	Sig	ELL
		Occur	Freq (%)	Occur	Freq (%)			
1	CAN VBI VVN	487	0.044046767	6747	0.112820178	-530.3961792	0.00000	0.00001
2	COULD VBI VVN	48		1904	0.031837799	-373.8757935	0.00000	0.00001
3	CAN VBI VVN II	167	0.015104333	2954	0.049395408	-319.9978638	0.00000	0.00001
4	NN1 CAN VBI VVN	163	0.014742552	2670	0.044646494	-264.2997131	0.00000	0.00000
5	NN1 COULD VBI VVN	15	/	696	0.011638187	-146.4420471	0.00000	0.00000
6	NN1 CAN VBI VVN II	69	/	1242	0.020768143	-137.135498	0.00000	0.00000
7	WOULD VBI JJ	49	/	931	0.015567747	-108.8482895	0.00000	0.00001
8	COULD VBI VVN II	25	/	679	0.011353921	-107.2691422	0.00000	0.00001
9	CAN VBI VVN II AT	38	/	796	0.01331034	-102.2843933	0.00000	0.00000
10	NN1 NN1 CAN VBI VVN	15	/	490	/	-86.93881989	0.00000	0.00000
11	CAN VBI VVN CST	11	/	428	/	-83.25879669	0.00000	0.00000
12	PPH1 CAN VBI VVN CST	10	/	399	/	-78.55500031	0.00000	0.00000
13	PPH1 CAN VBI VVN	44	/	744	0.01244082	-76.50708771	0.00000	0.00000
14	MAY VHI VVN	11	/	393	/	-73.23538971	0.00000	0.00000
15	CAN RR VBI VVN	49	/	765	0.012791973	-71.2289505	0.00000	0.00000
16	NN2 CAN VBI VVN	114	0.010310742	1294	0.021637663	-70.9580307	0.00000	0.00000
17	CAN VBI VVN TO	34	/	625	0.010450958	-70.58866119	0.00000	0.00000
18	WOULD VVI II	35	/	626	0.010467679	-68.63321686	0.00000	0.00000
19	CAN VBI VVN II AT NN1	22	/	494	/	-67.47689056	0.00000	0.00000
20	NN2 COULD VBI VVN	10	/	349	/	-64.23249054	0.00000	0.00000

Table 6 lists the top 20 modal sequences with the highest minus keyness, which means that the modal sequences are commonly used among the natives but less frequently used in the observation corpus. The table suggests while the natives often adopt nouns(singular or plural), and the demonstrative pronoun it(PPH1)before modal verbs can/may/could as sentences' subject, and followed by (+ /RR) +VBI +/(/VVN) +/[(/II) +/(/NN1)]+/[(/JJ)] +/[(/CST)] +/[(/TO)], while Chinese English majors rarely used such modal sequences as constituted by can and could in this study. It is not consistent with the fact that verbs with marked aspect and voice (e.g., passive voice and perfective aspect) in the reference corpus are used much more often than verbs with unmarked aspect and voice after modal verbs, which is consistent with Liang (2008),Liu (2009), Chen (2012), J.Zhang (2013),Author (2014),X.Y. Zhang (2015), Feng (2017), Yan (2017)and Chang (2016). And the effect here means for the study that CAN VBI VVN, COULD VBI VVN, CAN VBI VVN II,WOULD VBI JJ, COULD VBI VVN II are relatively speaking more likely to occur under the null hypothesis as the alternative, and the null hypothesis cannot be rejected than the others here in table 6.

5.2 Some Particular Features of Modal Sequences in Relation to Modality Type
By matching the overused modal sequences in Table 5 and the underused modal sequences in Table 6 with the syntactic structures of deontic and epistemic modality, it becomes evident that the overused modal sequences in the observation corpus are generally those with deontic modality, but what about the epistemic sequences?
5.2.1Stative verbs features after modal verbs
With the typical 11 modal verbs(with its tagged sets) as search terms (span set as R0 - R6), after searching in the tagged WECCL2.0, we saved the result as EXCEL spreadsheets, deleted unrelated verbs(leaving 57 columns behind), then conducted a counting based on the verb frequency behind the modal verbs. After frequency-data processing via Pivot table, we set R1 as the center and conduct order- processing in a descending way first, then R1 and R2 are added, and the frequency of the verbs after the modal verbs are counted after the order-descending processing.

The structure of "modal verbs + dynamic verbs" is a typical deontic modal sequence. To see whether VVI in the overused structures in Table 5 like CAN VVI APPGE,SHOULD VVI TO,SHOULD VVI TO VVI,CAN VVI AT,SHOULD VVI AT,NN2 SHOULD VVI TO,CAN VVI AT1,CAN XX VVI,NN2 SHOULD VVI TO

VVI and SHOULD VVI DAR are dynamic, this research has identified the modality of all verbs following modal verbs in the observation corpus, and the results show that while non-native learners adopt dynamic verbs, native speakers tend to use stative verbs in most cases. Table 7 lists the top 20 verbs that are most heavily used among the Chinese English majors in descending order of frequency.

Table 7 .Top 20 Verbs Used Most Frequently after Modal Verbs among RC (N)and OC(NN)

	OC		RC			OC		RC	
Rank.	Freq	verbs	Freq	verbs	Rank.	Freq	Verbs	Freq	verbs
1	6250	be	157931	be	11	508	give	453	use
2	1808	learn	34987	have	12	470	know	489	give
3	1548	make	5980	lead	13	450	become	454	increase
4	1475	have	733	take	14	438	use	457	become
5	1327	do	764	help	15	428	live	451	do
6	922	get	676	see	16	407	say	444	result
7	755	help	680	make	17	398	see	385	affect
8	699	take	697	provide	18	341	improve	427	occur
9	573	find	542	cause	19	331	go	371	argue
10	533	pay	511	need	20	300	bring	346	find
In total	15890		203501			407		4277	

Likewise,with the typical 11 modal verbs(with its tagged sets) as search terms (span set as R0 - R6), after searching in the tagged WECCL2.0, we saved the result as EXCEL spreadsheets, deleted unrelated stative verbs(leaving 50 columns behind), then conducted counting based on the frequency of the stative verb (connected to the VM and including tagged sets and the original word)) behind the modal verbs. After frequency-data processing via Pivot table, we set R1 as the center and conduct order- processing in a descending way first, then R3 and R5 are added, and the frequency of the STATIVE verbs after the modal verbs are counted after the order-descending processing.

As can be seen from Table 7,in OC and RC ,among these 20 frequently-used verbs, only be, have, find, know, become, see can be used as stative verbs(so the others are theoretically used somewhat as typical dynamic verbs), with the exception that the order is a little different. In addition, the standard frequency (per 1000) of all the stative verbs listed here in OC and RC is 19.84 %and 34.01% respectively3.And the standard frequency(per 1000) of all the common stative verbs listed here like be, have, find, know, become, see in OC and RC is 9.56% and 31.83% respectively, even the standard frequency(per 1000) of be is 6.21 and 25.86 respectively. The normalized ratio(‰) of the stative verbs be and have is 7.7 and 31.6 respectively. It is evident that the natives employed 3 times more stative verbs after most frequently when compared with Chinese English majors, who, however,used only 5 times more frequently than did non-majors(Liang, 2008),which is similar to Author's(2014) and Feng's (2017) studies.

The research has tagged the 55 typical stative verbs4 (Zhang, Z.B. (2003))of varied forms in the first category into be_VB0, BE_VB0, were_VBDR, was_VBDZ, be_VBI, are_VBR, is_VBZ,have_VH0, HAVE_VH0, had_VHD,HAD_VHD, am_VBM,have_VHI,HAVE_VHI ,then they are searched in the two independent, pre-processed corpora respectively, using the stative verbs spanning four categories. Results (as seen in Table 8)show that the sequence of "VM+stative verbs" appeared 9,897 times in the observation corpus, accounting for 9.84 ‰ of all tokens, while the same sequence appeared 40,156 times in the reference corpus, accounting for 0.657 ‰ of the total, much lower than those in the observation corpus. Meanwhile, the occurrence of stative verbs of the four categories is (38/55=0.69) in the observation corpora and a slightly lower (42/55=0.76)than those in the reference corpora. Furthermore, the relative ratio is still much lower:$38 \div 1,005,607 = 3.77 \times 10^{-5}$ and $42 \div 6106,163 = 6.796 \times 10^{-6}$. According to these, we can see Chinese English majors use more stative verbs than natives and the relative ratio is still much higher among the typical stative verbs in the four categories here mentioned (as seen in table 8). If the verbs in the sequence of "modal verbs + verbs" are stative, the sequence expresses epistemic instead of deontic modality (see Table 1).By combining Table 6 with Table 1, we can also see that many modal sequences underused in the observation corpus, such sequences as WOULD CAN VBI VVN,COULD VBI VVN,CAN VBI VVN II,NN1 CAN VBI VVN,NN1 COULD VBI VVN,NN1 CAN VBI VVN II,WOULD VBI JJ,COULD VBI VVN IICAN VBI VVN II AT,VVI II ,and COULD VBI VVN are all rather typical epistemic sequences theoretically, although not so much used in Oc,the feature to use stative verbs after modals is still very strong. And the standard frequency of stative verbs(at least the above 55) shares a contrastive comparison ,another evidence to say Chinese English learners have an academic trend when talking about something they concern. And ELL here means for the study that the effect here is so low and so unstable because low values of the likelihood ratio mean that the observed result was less likely to occur under the null hypothesis as compared to the alternative. This provides further evidence that the modality of the modal sequences used by Chinese English majors is related to their linguistic proficiency, a finding consistent with all the conclusions of Liang's (2008) , Liu's (2009) , Tang's(2013), Wang's(2013), H.Q.Zhang's(2013), Author's (2014), Chang's (2016), Feng's(2017)findings. It is also clear that Chinese English majors show a greater variety when using the sequence "modal verbs + stative verbs" than do Chinese non-English majors, which can be seen from Liang(2008).

Table8.Comparison of stative verbs after VM between Oc and Rc

	Observation	corpus	Reference	corpus	ELL
	With	occurrences	With	occurrences	
SUM(VM)	VM	9896	VM	40154	0.00002
SUM(VMk)	VMk	1	VMk	2	0.00000
Freq(Sum)	VM+VMk	9897	VM+VMk	40156	0.00002
Standard Freq	VM+VMk	0.09841	VM+VMk	0.00658	0.00003

5.2.2 Animated/unanimated subjects before modal verbs

Subject in the structure of "animated/unanimated subject + modal verbs" is another key to telling whether the sequence is deontic or epistemic. The sequence is generally deontic if the subject before the modal verbs is animated (Coates 1983; Wärnsby 2003). By referring to Table 9, we can see that "personal pronouns (such as I, we ,and you) + modal verbs" is a typical deontic sequence employed most frequently in the observation corpus. Thus,the heavy usage of this structure is another prominent feature among the Chinese English majors.

Table 9 Use Frequency of "Personal Pronouns + Modal Verbs" between RC and OC

Sequences	RC		OC		ELL
	Freq.	ST. Freq. (per10,000)	Freq.	ST. Freq. (per10,000)	
We + VM	3594	5.88	8795	87.46	0.00003
You + VM	593	0.97	2855	25.75	0.00000
I + VM	2083	3.41	668	6.64	0.00013
TOTAL1	**6270**	**10.26**	**12318**	**119.85**	/
She + VM	409	0.69	204	2.03	0.00003
He + VM	787	1.28	627	6.24	0.00004
They + VM	2242	3.67	3494	34.74	0.00005
TOTAL2	**3438**	**5.64**	**4325**	**43.01**	/
In total	**15969**	**26.14**	**18495**	**18.39**	**0.00034**

Note: VM stands for modal verbs.

Almost in the same way, with the typical 11 modal verbs(with its tagged sets) as search terms (span set as L1 - L3), after searching in the tagged WECCL2.0, we saved the result as EXCEL spreadsheets, deleted unrelated personal pronouns(leaving 57 columns behind), then conducted counting based on the frequency of the personal pronouns (connected to the VM and including tagged sets and the original word) before the modal verbs. After frequency-data processing via Pivot table, we set L1 as the center and conduct order- processing in a descending way first, then L1 and L3 are added, and the frequency of the personal pronouns before the modal verbs are counted after the order- descending processing.

The data in Table 9 show that the use frequency of "personal pronouns + modal verbs" in the observation corpus is 0.70 times as high as that in the reference corpus (18.39:26.15).

However,the occurrences of "we/you/I + modal verbs" in the observation corpus are many times (5.89, 0.97 and 3.41 respectively)fewer than those in the reference corpus. These results are inconsistent with the conclusion of Liu (2009) that non-native speakers tend to overuse the first and second person pronouns before modal verbs as subject. And the different use frequency of "they/he/she+ modal verbs" between the learners   and the natives show that natives employ epistemic modality much less often(:42.74:5.63) than the nonnatives. Compared with the results by Liang (2008), Chinese English majors' usage of "personal pronouns + modal verbs" in this research is remarkably fewer, even the structure "I + modal verbs" occurs fewer times than those   in the reference corpus ,another sign to say Chinese majors may be more objective in talking something they concern. However, the effect size (ELL) here means for the study that the natives are relatively using They + VM,He + VM and SHE+VM more as shown in table 9 and the null hypothesis cannot be rejected. This result confirms that Chinese English majors have a better grasp of some of the features of epistemic modality and is consistent with Liu's(2009) that Chinese English majors' use of modal sequences is closer to the native speakers than to the Chinese non-English majors but inconsistent with Tang (2013), Long (2014) , Feng (2017) Yan (2017)and Chang (2016).

5.3Other academic features in OC.

Table 10.The first 20 plural nouns in STU1

RANK	FREQ	FREQ(L)	FREQ(R)	STAT	COLLOCATE
1	5630	5630	0	3.27581	students
2	4193	4193	0	3.27117	children
3	2767	2767	0	3.27632	others
4	2095	2095	0	3.27494	things
5	1855	1855	0	3.27477	games
6	1773	1773	0	3.27632	animals
7	1716	1716	0	3.27632	cards
8	1238	1238	0	3.27283	parents
9	1117	1117	0	3.27503	friends
10	1032	1032	0	3.27353	problems
11	987	987	0	3.27048	words
12	848	848	0	3.27632	dictionaries
13	685	685	0	3.27632	skills
14	667	667	0	3.26769	women
15	666	666	0	3.10688	lives
16	643	643	0	3.27632	kinds
17	637	637	0	3.25162	functions
18	631	631	0	3.27632	countries
19	619	619	0	3.27632	advantages
20	616	616	0	3.27632	books

5.3.1 Some particular academic feature shown in majors' writings

Chinese English majors usually incorporate NN2 or we as a sentence's subject, and their use frequency of "We + modal verbs" is many times more than that in the reference corpora, which are some academic features shown in non-academic writing as mirrored by the (quasi-)academic reference corpora. "NN2+VM" pattern: As is shown in table 10, the top 9 plural nouns students, children, others, things, games, animals, cards, parents, friends here work as subjects, indicating that the choice of words in learners' written expressions mainly depends on the topic and theme. They are theme-related because words such as students, children, games, animals, cards, parents, friends are in writing test section in TEM-4 or TEM- 8 as test items. In order to find out some modal verbs after them, we picked up 19 abstract nouns (problems, skills, lives, kinds, functions, advantages, reasons, abilities, disadvantages, aspects, opinions, effects, measures, rights, efforts, difficulties, chances, benefits, opportunities) from the top 100 plural nouns. It turns out that only 4 abstract nouns before modal verbs in observation corpus (as seen in table 11)while there is none in reference corpus, which, to some degree, indicates Chinese English learners are somewhat subject to adopting negotiating strategies when talking about something they believe though they are all strong mood of modal verbs(Z.H.Zhang,2016). The emergence of the sequence of "abstract subject + modal verb" indicates that learners pay some attention to the weak position of one's own mood in their academic discourse (Z.H.Zhang, 2016).

Table 11.Abstract nouns before modal verbs in OC

Word	With	Texts	Total Right
SHOULD	measures	78	60
WILL	problems	30	17
WILL	lives	11	11
MUST	measures	11	5

As shown in table 12, the VM (after NN2) standard frequency difference between both observation and reference corpus appears almost the same pattern: Type D modal verbs occurs more than Type C modals. However, it seems that ca(delete form of can't) here in the observation corpus carries more "possibilities" than those in the reference corpus. Actually, we can see a little more academic features advances at least in such patterns. And "may" here embodies more "possibilities" than "can"in such a pattern though it is not so polite and indirect in mood(Z.H.Zhang,2016), another confirmation with the finding here above in table 5. "We + can/could + n ' t/not" sequence is often used to express the limitation of individual ability or the limitation of individual ability by objective conditions (Z.H.Zhang,2016).And it is appropriate that they put NN2 before modal verbs, like should, which increases author's objectivity and thus maintains the arguing effects,which is different from the findings from Author's finding(2014), and somewhat consistent with Chang's(2016) and Feng's, J.Zhang's(2013) , Yu's(2016) studies.

Table 12.Standard frequency difference of VM after NN2 between OC and RC

	Group	N	Mean	Std. Deviation	Std. Error Mean
FREQ	1	14	.00509130	.007745238	.002070002
	2	14	.00169979	.001995296	.000533265

6. Discussion

The research has revealed the shared features of modal verbs /sequences used in Chinese English majors' argumentative essays. Firstly, modal verbs are used with uneven frequency: Type D Modal Verbs like should,can,and ca(deleted form of "can't")are overused ,while Type E Modal Verbs like could and would are underused. Secondly, compared with Chinese non-English majors, Chinese English majors usually incorporate agents (particularly NN2 and we) as the sentence's subject though they overuse verbs with unmarked aspect and voice after modal verbs, and their use frequency of stative verbs has surpassed the natives, their relative use frequency ratio is still higher. Thirdly, Chinese English learners rarely used modal sequences constituted by can and could in the way that natives did. For example ,the natives often adopt nouns(singular or plural), the demonstrative pronoun it(PPH1)before modal verbs can /may/could as the sentence's subject, and followed by (+RR) +VBI+(/VVN) +[(/II)+(/NN1)]+/[(/JJ)]+/[(/CST)]+/[(/TO)],while Chinese English majors use more stative verbs than the natives and in a much higher relative ratio, And the occurrences of " they/he/she + modal verbs" in the observation corpus are many times more than those in the reference corpus, even the occurrences of "we/you/I + modal verbs" between show that the nonnatives employ much more less than the natives. In contrast to the observation corpora, the reference corpora can betray some hidden features when compared with the observation corpus. Anyway, in short, Chinese English majors' writings here show comparatively strong narrative features and are somewhat lacking in epistemic comments on the issue under discussion due to their feature to overuse deontic modality and underuse epistemic modality. However, some academic features can be found to be hidden in the students' writings:

   Ca (deleted form of "can't") are highly overused, which can account for some academic trends in learners' writings, and as is shown in the "we + weakened modal verb (Zhang,Z.H.,2016) + not/n't" pattern, the co-occurrence of we and can is the most. Chinese English learners put NN2 before modal verbs, like should, which increases author's objectivity and thus maintains the arguing effects .The standard frequency of stative verbs of Chinese English learners shares a contrastive comparison in   that they use absolutely more stative verbs and in a much higher standard ratio as a whole. The occurrences of "we/you/I + modal verbs" in the observation corpus are many times fewer than those in the reference corpus. And the use frequency of "they/he/she+ modal verbs" between the learners and the natives show that native speakers employ epistemic modality much less often than the nonnatives. Only 4 out of 19 carefully-picked abstract nouns before modal verbs in observation corpus while there is none in reference corpus. And different keywords can be generated rather different reference corpus in the meantime ,and a rank order regarding academic trend related to RC can be seen.

   Many reasons contribute to the features above, this research will try to explain them from a perspective of language ontology or social culture or pragmatics or language use.

1) Negative transfer of mother tongue. Learning the passive voice of English, the Chinese English learners rarely can avoid the influence of the theme-theme patterns of the Chinese language meaning that Chinese English learners take the nomial phrase at the beginning of the sentence as the theme instead of the subject (Tang, 2006). Meanwhile, since they still have no grasp of the grammatical rules of passive voice ,especially of the complex usage of English verbs (Zhuang, 2005), they rarely choose to adopt passive voice.

2) Restrained deontic modality. Shu (2010) has grouped the modality of English modal verbs into two categories, deontic/dynamic (both being event modality) and epistemic (a kind of propositional modality), stating that the epistemic modality is a result of the de-categorization of deontic/dynamic modality. Such de-categorization shows different features. First, during the development from deontic to epistemic modality, the typical features of deontic modality gradually give way to epistemic features, but dynamic modality rarely shows up as an intermediary. Second, during the transition from dynamic to epistemic modality, the typical features of dynamic modality gradually disappears while epistemic features gain prominence, and in the meantime, deontic modality usually play the role of intermediary. Both deontic and dynamic modality can show up as prototypes and serve as the others intermediary.

3) It is reasonable to say that deontic and dynamic modality are closely related. As mentioned above, during the transition from dynamic to epistemic modality, deontic modality usually shows up as an intermediary. This means the illocutionary acts of first-person subjects are delivered through the intermediary of obligation and responsibility, which confirms that deontic sequences expressing responsibility and obligation generally have subjects of first person pronouns ("modal verbs + dynamic verbs" and "animated subjects + modal verbs" can only express deontic modality) (Liang, 2008). Therefore, "animated subjects + modal verbs +dynamic verbs" is a deontic sequence in all cases, and the conveyance of deontic modality is based on animated subjects or dynamic verbs. As long as the subject before a modal verb is animated, the modal sequence

expresses deontic meaning. The representative structure "personal pronouns (we/you) + modal verbs" is one of the modal sequences most frequently employed among the learners, and to overuse it is a remarkable linguistic feature of them (Liang, 2008).The modality of responsibility and obligation is often expressed by dynamic verbs describing actions or events instead of stative verbs describing psychological, sensory, and emotional activities, so the learners often have the dynamic verbs instead of the stative ones after modal verbs. The reason why they tend not to give a sufficient objective description on events is that deontic and dynamic modality are both event modality instead of propositional modality.

4) Chinese English majors usually used the agents (in particular: we, you,I) as subjects of the sentences on most occasions, which is clearly related to the fact that English learners have the confidence in their intentions, but less in others'. And they feel that their duties are very important when talking about problems and, then they stand as masters and show great responsibility to the social issues. Due to the overuse of this kind of sentence pattern, the sentence structure is pretty monotonic, lacking changes when it is necessary (W.Y.Liu,2009) And you/we +modal verbs makes learners more subjective, which lowers or ruins the arguing effect of the texts (Y.W. Liu,2009) .

5) However ,here in the paper, the occurrences of "we/you/I + modal verbs" in the observation corpus are many times fewer than those in the reference corpus. And the use frequency of "they/he/she+ modal verbs" between the learners and the natives show that native speakers employ epistemic modality much less often than the nonnatives, which reflects the improvements in social pragmatic transfer. Learners can be more objective and rational even showing somewhat academic trends when they talk about their own responsibilities and obligations, or those of groups including themselves as unexpected by researchers here.

6) Different awareness of responsibility from non-native cultures. Palmer (1986) thinks modal verbs like must and should bear obligatory modality, which implies the existence of external requirement. When it comes to showing the responsibility of being filial and respectful to one's parents, people in different countries are faced with varied levels of pressure. For example, the parental authority in China as well as in Japan is stronger than that in the US. The social culture of the three countries is very different in their criteria on parental authority, disciplining children, as well as on the obligation of sons and daughters. When talking about traditional topics ,like family, and friendship, overseas students from Asian countries like China, Singapore, North Korea, Japan, Indonesia, and Vietnam often focus on the aspects of harmony, familial, or communal responsibility and tend to draw from the values of Confucianism, Taoism, and Buddhism, showing the existence of obligations imposed on them from the outside. Therefore, it is safe to say that the learners' choice and employment of modal verbs are related to their cultural convention and contexts. And their extremely strong awareness of responsibility originates from a thinking pattern rooted in Confucianism, Taoism, and Buddhism, and such thinking pattern has produced the obligations and requirements imposed from the outside and contributed to their posture as mastery (Hinkle, 1995). Modal verbs with obligatory modality are very dependent on genuine facts (Coates, 1983)— respects of certain authority and moral principles (Lyons, 1977) and inseparable from a force imposed on man from outside (Sweester, 1990). Meanwhile, obligations and requirements tinted with subjectivity often convey responsibility in a manner approved by a certain culture (Collins, 1991).So when the learners use modal verbs, they usually bring into their English writings the social norms and values of their own culture (Basham & Kwachka, 1989). The reason why non-native learners show a feature to overuse deontic modality(such as "should") is that they basically have gained a wrong understanding of the British and American culture and of the pragmatic implications of English modal verbs, a phenomenon of asymmetric cultural information (Atman, 1990).

7) The usage features of modal verbs/sequences can be better understood from the perspective of practical linguistic application. A more likely reason why learners tend to overuse Type D Modal Verbs is that these verbs are introduced earlier in their textbooks and thus learners feel more confident about using them. The monotonous, repeated overuse of can and can't is probably due to the fact that learners are still far from the comfort of using other modal carriers (such as possibly, it is (un)likely, etc.) with ease and the fact that they are afraid of making grammatical mistakes. As for ca, it is due to the fact that Wordsmith tools tend to separate the abbreviation of negative form of words into two separate parts(e.g.ca +n ' t), here mostly meaning "impossibly being able to ". Chinese English majors use more abbreviations such as " ca + n' t " in their written English, which show the negative forms of can are prone to be can't, in the way that it can extend students' thinking time. This is consistent with Wen (2006) that the features of spoken and written English are still obscure to Chinese students, and written production is a bit like oral English and vice versa(Chang,2016).

8) The underuse of Type E Modal Verbs is probably because could and would are using in relatively complex, unrealistic conditions and are introduced relatively later in their textbooks. Thus, the learners are not sure how to use them. Their feature to use the most simple, colloquial, and grammatically safe sequences of "personal pronouns/singular or plural nouns + modal verbs" and "modal verbs + verbs with unmarked tense" is probably a result of having grasped Type D Modal Verbs so earlier and so better that they have grown

reliant on them and feel it is difficult to use other more complicated expressions.

9)   Learners' pragmatic inappropriateness due to their rare use of passive voice in modal sequences. Chinese English majors are unaware that modal verbs, like should, are often used to indicate collective obligations and that incorporating passive voice among the modal verbs can save the trouble of indicating who is obliged. In learners' written English, modal sequences containing could and would are inappropriately used, probably due to their ignorance of these sequences being able to convey interpersonal meta-functions (Halliday, 1985) and more polite tones (Biber et al., 1999; Quirk et al., 1985). The underuse of could and would, to some extent, adds to the abruptness and bluntness of the text and damages its interpersonal functions, all giving the argumentative essays an air of dogmatism and thus reducing their reasoning effects. In the meantime, their writings show comparatively strong narrative features yet are lacking in epistemic comments on the issues under discussion.

10)   Chinese English majors share more similarity with the natives than Chinese non-English majors do in their employment of modal sequences. One more possible reason is that they have spent more time on the target language and thus, consciously or unconsciously, they get to understand how modal sequences are learned and used among the natives. Also, the learners' early grasp of deontic modality and then epistemic modality of modal sequences shows their cognitive development, which is in line with humanity's development of cognition and their progress in acquiring a new interlanguage. As they spend increasingly the longer time learning a foreign language, their linguistic performance moves further from the features of their mother tongue and get to be closer to those of the target language(Skehan,1998). For this reason, they often employ plural personal pronouns and singular/plural nouns instead of I before modal verbs as the subject, and also the subject shows a greater variety than Chinese non-English majors. However, they also overuse stative verbs with unmarked tenses or voices behind modal verbs, which is indeed a sign of interlanguage(Liang, 2008),and because of this typical sequences that are earlier acquired due to earlier involved in textbooks. Besides, Chinese English majors use more active sentence patterns in expressing dynamic meaning ,only to result in misunderstanding of readers, which is attributed by the differences of thinking patterns between English learners and native speakers(Liu, 2009).

To persuade to accept and then obtain permission and to to avoid unnecessary criticism or attack or being lowly graded Chinese English learners consciously and unconsciously show some academic features in their time-limited prompt writings about some concerns in their times, the only purpose is to increase their room for negotiation for further discussion in order to gain a better arguing effect to be better graded, only to get a better pragmatic effect as if they are doing a very academic paper trying to persuade their readers, so they are consciously and unconsciously using some kind of academic writing strategies to persuade to accept and then obtain permission and to avoid unnecessary criticism or attack in order to increase the room for negotiation. As Hyland (1998) argues, the purpose of academic discourse writers is to persuade readers to accept their views and then obtain permission to enter a certain discourse community. Li (2011)also believes that scholars tend to refuse to express absolute commitment in order to avoid unnecessary criticism or attack and increase the room for negotiation (Z.H.Zhang,2016).Chinese English learners are practicing to be like a negotiator with their grading teachers like an academic researcher is with his blind reviewers .In short, learners' most use of the sentence patterns: 1) animate subjects (e.g. we) +can + not+/epistemic modal adverb /+ dynamic verbs (active voice, unmarked aspect), more used in general English writing, expressed the obligation of typical modal semantics, which fully shows Chinese students learned early in the process of learning English vocabulary, grammar and sentence patterns. And 2) inanimate subjects (proper noun) + can +/non–epistemic modal adverb/ + dynamic verbs (active voice, unmarked aspect), a certain understanding of modal semantics, is used for academic discourse to explain some and applications, etc., but non-epistemic , unmarked dynamic modal verbs like can, are mostly used in revealing, to some extent, learners' relatively monotonous and shallow in modal semantic connotation construction type. These features that the subject of the sentence is mostly they/he/she less you/we/I , followed mainly by modal verb can, and then by dynamic verbs modified by non-epistemic modal adverbs, active voice or unmarked aspect are often used, and existential sentences or (non-) restrictive attributive clauses are seldom used in the sentence pattern indicate, to a large extent, that learners are aware of notes on academic propositions, and full of necessary markers of author's position, and of euphemism, and of construction of evaluation meaning in academic writing, while all these features can be hidden in general essay writing like TEM-4 or TEM-8.

7.Conclusion

Drawing up the methodology from Aarts & Granger(1998) and Tono(1999), the research processed the tagged texts with PowerGrep, and analyzed the employment of modal verbs and modal sequences used in Chinese English majors' argumentative essays. With WordSmith Tools like Concord, Wordlist, and Cluster Analysis, we have found that the usage features of modal verbs in these argumentative writings. Results show different compound features(including non-academic and academic features). Regarding non-academic features: Firstly, should,can, are overused while could and would are underused, but ca(deleted form of "can't") is used here to express mostly

the limitation of individual or individual ability by objective conditions. Secondly, Chinese English majors usually incorporate NN2 or we as sentence's subjects, and overuse verbs with unmarked aspect and voice after modal verbs. They underuse modal sequences constituted by can and could.(e.g. can be cancelled, could be bought etc.). However, NN2s are followed mostly by should, which increases author's objectivity and thus maintains the arguing effects, and we as sentence's subjects, the usual form to refer to the author himself /herself or the research group themselves indicating their action or responsibilities he/she /they are doing or taking. Chinese English majors employ absolutely more stative verbs, although the standard frequency(per 1000) of all the stative verbs between non-native and native corpora is 23.14 and 34.03 respectively. And the standard frequency(per 1000) of all the common stative verbs in non- native and native corpora is 9.56 and 31.83 respectively, even the standard frequency(per 1000) of be is 6.21 and 25.86 respectively among these 20 frequently-used verbs can be used as stative verbs, showing an academic trend when talking about something they concern, And the standard frequency of stative verbs(at least the above 55) shares not a very contrastive comparison ,another evidence to say Chinese English learners have an academic trend when talking about something they concern. And the occurrences of "we/you/I + modal verbs" in the observation corpus are many times fewer than those in the reference corpus, the use frequency of "they/he/she+ modal verbs" between the learners and the natives show that native speakers employ epistemic modality much less often than the nonnatives, indicating less subjectivity relatively. And it turns out that only 4 out of 19 picked abstract nouns before modal verbs in observation corpus while there is none in reference corpus, which, to some degree, indicates Chinese English learners are somewhat subject to adopting negotiating strategies when talking about something they believe.

Moreover, it also approached this issue from the perspective of practical language use aside from other perspectives. The conclusion will surely provide some implications for China's English teaching, especially in the teaching   of modal verbs in terms   of modal teaching, academic   writing   and/or oral training.   etc. (Long ,2011,2012,2013,2014, 2016). The transformation from deontic modality to epistemic modality is conceived as a manifestation of the cognitive development law, which implies that this process seems to be automatic and learning in advance is not a prerequisite; however, conscious teaching and explicit learning would undoubtedly help save learners from detours. Accordingly, this paper proposes that while the learner's cognitive law should be followed in textbook compiling, classroom teaching and even after-class summarizing, the role that syntax sequence of modal verbs play should be consciously highlighted. To this end,specified suggestions are given as follows:

（1）   We should bear in mind when compiling textbooks, making sure epistemic modal verbs with euphemism modality should be more exposed to students who are badly needed in early junior/senior high school experiences. Detailed usage of modal sequences with euphemism modality such as "could" and "would" should be summarized and be made part of the revision sections. Spiral methods with repeated examples can be applied to the texts followed to help students consolidate and digest these knowledge.

（2）   Native speakers' feature to use modal verbs should be explicitly clarified in class, e.g., native speakers tend to use more personal pronouns (in plural form) and inanimate subjects before the modal verbs while verbs with markers of tense and aspect after the modal verbs; they prefer to use stative verbs to convey speakers' assessment on the proposition. Meanwhile, a variety of sentence patterns of modal sequences should be consciously practiced both in and after class with reference to how they are used by native speakers.

（3）   To develop students' sensitivity to modal verbs with high automaticity, examples of epistemic modal verbs that native speakers tend to use should be consciously presented both in and after class, particularly in students' language output practices like argumentation writing.

（4）   The proper pragmatic meaning of modal verbs, the basic value view and social philosophy of Anglo-American Culture involved, as well as the differences in cultural tradition and value between East and West should be underlined in English modal verbs teaching. Moreover, students should be given more chances of self-discussing and modal verbs' context analyzing in relation to modal verbs' social and cultural implications.

（5）   The analysis of modal sequences is relatively new and complicated in corpus-based study. This analysis can reveal the morphological and syntactic sequences that are beyond the reach of Concordance and Wordlist   tools ,and the new approach to use the typical modal sequence in the observation corpora for the ranking of academic feature here provided is a new angle for researchers. With WECCL 2.0, this research has proved effective in finding the usage features of modal verbs/sequences of Chinese English majors. To provide pedagogical implications for modal sequences, more research can be conducted that studies WEECL1.0 and 2.0 respectively or compares the two groups' academic features of modal sequences in the course of 11 years, or more research that compares the academic features of modal sequences between students in different grades, or more research that compares the academic features of modal sequences in students' oral English based on WECCL1.0 and 2.0 or any given nonnative corpus ,such as TECCL,ICLANE etc, or more research that compares the academic features of modal

sequences in their written and spoken English or any given corpus regarding different registers of English or other language in EFL or ESL contexts, such as fiction, newspaper or academic writings, etc.

Declaration of Conflicting Interests
The Author(s) declared no potential conflicts of interest with respect to the research, authorship,and/or publication of this article.

Notes
Precision :Oakes Mike defines that it is the "the proportion of retrieved items that are in fact relevant ,the number of relevant items obtained divided by the total number of retrieved items" (Oakes,M1998:176).
This size(token) is subject to the original raw corpora practically published, but it is recalculated in "tokens used for wordlist" in statistics embedded in the "Wordlist" in Wordsmith Tool and number of composition is   also subject to the marked "sections".WECCL2.0 has 3,027 instead of the supposed 4,359 argumentative essays, and files like a2105nd.cls and a341xsd.cls, etc. are damaged when processed by Wordsmith, so only 3,027 argumentative essays are available to this study.
The frequency here is calculated on the basis of tagged texts; the reference corpus has 6,107,989 tokens and the observation corpus has 1,005,607 tokens.
Zhang, Z.B. (2003, 163) has classified stative verbs into four types:
1)   be or have used as main verbs (in the sense of "to exist")；
2)   verbs with similar meanings as be and have, such as apply to, belong to, differ from, cost, weight, measure, fit, hold, lack and resemble;
3)   verbs of perception, such as feel, hear, see, smell and taste;
4)   verbs concerning mental or psychological status, such as assume, believe, consider, detest, fear, hate, hope, imagine, know, like, love, mean, mind, notice, prefer, regret, remember, suppose, think, understand, want and wish.

References
Altman, R. (1990), Developing Communicative Competence in a second language. New York: Newsbury House.
Aijmer, K. & Altenberg, B. (1991). English Corpus Linguistics. London: Longman.
Aarts, J., & Granger, S. (1998). Tag sequences in learner corpora: A key to interlanguage grammar and discourse. Learner English on computer, 132, 141.
Aijmer, K. (2002). Modality in advanced Swedish learners written interlanguage. Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching. 55-76. https://doi.org/10.1075/lllt.6.07aij
Basham, C. & Kwachka, P. (1989). Variation in modal use by Alaskan Eskimo student writers. Variation in second language acquisition, 1, 129-143.
Berber-Sardinha, T. (2000). Comparing corpora with WordSmith Tools: How large must the reference corpus be?. In Proceedings of the workshop on Comparing corpora (pp. 7-13). Association for Computational Linguistics. http://dx.doi.org/10.3115/1604683.1604687
Biber, D. S. et al. (1999). longman Grammar of Spoken and Written English. Harlow: Longman.
Bitchener, J., Storch, N., & Wette, R. (Eds.). (2017). Teaching Writing for Academic Purposes to Multilingual Students: Instructional Approaches (Vol. 65). Taylor & Francis.
Bouhlal, F., Horst, M., & Martini, J. (2018). Modality in ESL Textbooks: Insights from a Contrastive Corpus-Based Analysis. Canadian Modern Language Review, (aop), 1-26. https://doi.org/10.3138/cmlr.3075.
Brown, H. D. (2000). Principles of language learning and teaching.
Chang, X.N. (2016). Corpus-based study of modal sequences in Chinese English majors' spoken language. Zhongbei university.
Chen, J. (2017). An investigation of promotion discourse in international journal research articles. Zhejiang University. Chen,A B. University of Nationalities, 18(01), 19-20.
Chen, H. I. (2010). Contrastive learner corpus analysis of epistemic modality and interlanguage pragmatic competence in L2 writing. The Arizona Working Papers in Second Language Acquisition and Teaching, 17, 27-51.
Coates, J. (1983). The Semantics of Modal Auxiliaries. London: Croom Helm.
Collins, P. (2009). Modals and quasi-modals in English. Brill. https://doi.org/10.1163/9789042029095. Cournane, A. (2014). In search of L1 evidence for diachronic reanalysis: mapping modal verbs. Language
Acquisition, 21(1), 103-117. https://doi.org/10.1080/10489223.2013.855218.
Corder, S. P. (1973). Introducing Applied Linguistic. Harmondsworth, Middlesex: Penguin Books.
Crompton, P. (1997). Hedging in academic writing: Some theoretical problems. English for specific purposes, 16(4), 271-287.
De Haan, F. (1997). The Interaction of Modality and Negation: A Typological Study. New York: Garland.

Dunning, T.(1993). Accurate Methods for the Statistics of Surprise and Coincidence. Computational Linguistics19 ,1:61-74.

Ehrman, M. (1966). The meaning of the modals in present-day American English. TheHague: Mouton. https://doi.org/10.1515/ling.1966.4.28.46.

Fan, X. (2016). Reconstruction of the Semantic-Development Path of the Modal Verb Hui (*) with Crosslinguisitic and Dialectal Evidence. Language and Linguistics, 17(2), 195.

Feng, S. (2017). A corpus-based study on English modal sequences in Chinese college students' written language. Xihua University.

Granger, S. and P. Rayson (1998). Automatic Lexical Profiling of Learner Texts. In Granger (ed.), 119-131.

Granger, S. (2002). A bird's-eye view of learner corpus research. Computer learner corpora, second language acquisition and foreign language teaching, 6, 3-33. https://doi.org/10.1075/lllt.6.04gra

Gabrielatos, C., & McEnery, T. (2005). Epistemic modality in MA dissertations.

Guo, X. (2005, July). Modal auxiliaries in phraseology: A contrastive study of learner English and NS English. In Corpus Linguistics.

Halliday, M. A. (1976). System and function in language: Selected papers.

Halliday, M. A. K., & Martin, J. R. (1993). Writing science: Literacy and discursive power. Pittsburgh. Hermerén, L. (1978). On modality in English: A study of the semantics of the modals (Vol. 53).

LiberLäromedel/Gleerups.

Hondo, J. (2012). Unveiling learners' attention during language processing: the case of epistemic meanings of a modal verb. Language Awareness, 21(3), 249-265. https://doi.org/10.1080/09658416.2011.598529.

Hinkel, E. (1995). The use of modal verbs as a reflection of cultural values. TESOL Quarterly, 29(2), 325- 343. https://doi.org/10.2307/3587627.

Hoye, L. (2014). Adverbs and modality in English. Routledge.

Huddleston, R. & G. Pullum. (2002). The Cambridge Grammar of the English Language. Cambridge. https://doi.org/10.1017/9781316423530.

Hunston, S. (2001), Colligation, lexis, pattern, and text. Pattern of Text: In Honour of Michael Hoey. 13- 33. https://doi.org/10.1075/z.107.03hun.

Hunston, S. (2004). "We can broke the forest": Approaches to modal auxiliaries in learner corpora. TaLC6, Granada.

Hyland,K.(1998).Hedging in scientific research articles (Vol.54). John Benjamins Publishing. http://dx.doi.org/10.1075/pbns.54.

Hyland, K. (2005). Metadiscourse: Exploring writing in interaction. London: Continuum, 13-15.

Hyland, K. (2017). Learning to write for academic purposes: Specificity and second language writing. In Teaching Writing for Academic Purposes to Multilingual Students (pp. 38-55). Routledge.

Kakzhanova, F. A. (2013). What modals are: Modal verbs, modal words, and auxiliary modals. European researcher. Series A, (10-2), 2530-2535. https://doi.org/10.13187/er.2013.61.2530.

Khojasteh, L. & Mukundan, J. (2011). A Pedagogic Corpus Analysis: Modal Auxiliary Verbs in Malaysian English Textbooks. Journal on English Language Teaching, 1(2), 45-55. https://doi.org/10.26634/jelt.1.2.1460.

Kwary, D. A., Kirana, A., & Artha, A. F. (2017). The distribution of verb tenses and modals in journal articles' abstracts. Southern African Linguistics and Applied Language Studies, 35(3), 229-244. https://doi.org/10.2989/16073614.2017.1373366.

Lado, R. (1957).Linguistics across cultures. Applied linguistics for language teachers.University of Michigan Press. https://doi.org/group/276/article/605161.

Larreya, P.(2004).Types of modality and types of modalisation. [OL]. http://www.univ-pau.fr/ANGLAIS/modal-ity/posters/larreya.pdf (accessed 09/08/2008).

Leech, G. (1998). Corpus Linguistics. London: Longman.

Li, J.J.et al.(2013) Discourse behavior of phrase sequences in academic texts. Foreign language teaching and research,45(2),200-213+318-319.

Li ,L.H. (2011).The use of modal verbs will and may in English academic paper writing and news style -- a cross-stylistic study based on corpus. Foreign language teaching,32(06),38-43.

Liu, W.Y. (2009). A Corpus-Based Study on the Written Form of Chinese Chinese English majors' Use of Modal Sequences, Changsha, HuNan University.

Long Shaoyun. (2011). A study of modal sequences in English major's argumentations. China English education, 2, 1- 13. https://doi.org/10.19716/j.1672-4720.2014.02.003.

Long Shaoyun. (2012). The relationship between English major's tendency to use modal sequence and their study time at school,Foreign language education, 4, 1-12.

Long Shaoyun. (2013). A study of modal sequences' grade features in English major's argumentations, China foreign language education, 1, 3-14.

Long Shaoyun. (2014). A Study on modal sequence usage characteristics in argumentative essays of Chinese

English majors,Foreign language and literature, 31, 90-102. https://doi.org/10.16263/j.cnki.23-1071/h.2016.01.025.

Long Shaoyun. , et al. (2016). Usage characteristics of modal verbs in argumentations of major students. Foreign language journal,1,124-131.

Liu, H. (2006). An Analysis of the use of SHOULD and MUST by Senior English majors in China. Journal of NingboUniversity(EducationalScienceEdition),28(1),86-89.https://doi.org/10.3969/j.issn.1008-0627.2006.01.025

Ma , G. &.Lv, X. J (2007). A Study of the Modal Verbs Based on Chinese Learners' English Corpus. Foreign language Audio-Visual Teaching, 3, 17–21. https://doi.org/10.3969/j.issn.1001-5795.2007.03.004

Mur-Dueñas, P. (2016). Modal hedging verbs in English as a Lingua Franca (ELF) Business Management research articles. Kalbotyra, (69), 153-178.

Myers, Greg. (1989). The pragmatics of politeness in scientific articles. Applied Linguistics10 (1), 1–35.

Johnston, J.E., Berry, K.J. and Mielke, P.W. (2006). Measures of effect size for chi-squared and likelihood- ratio goodness-of-fit tests. Perceptual and Motor Skills, 103, 412-414. https://doi.org/10.2466/ pms.103.2.412-414. https://doi.org/10.2466/pms.103.2.412-414.

Kratzer, A., Eikmeyer, H. J., & Rieser, H. (2002). The notional category of modality. Formal semantics: The essential readings, 289-323.

Markkanen, Raija & Helmut Schröder.(1997). Hedging: A challenge for pragmatics and discourse analysis. Hedging and discourse: Approaches to the analysis of a pragmatic phenomenon in academic texts ,3–18.

Mauranen, A. (1997). Hedging in language revisers' hands. Hedging and discourse: Approaches to the analysis of a pragmatic phenomenon in academic texts, 115-133.

M.C.Liang.(2008). A corpus-based study of modal sequences in Chinese tertiary EFL learners' written production, Foreign Language Teaching and Research (bimonthly)(1), 51-58.

Mindt, D. (1993). An Empirical Grammar of the English Verb: Modal Verbs, Berlin: Cornelson.

Neff, J., Dafouz, E., Herrera, H., Martinez, F & Rica, J.P. (2003). Contrasting learner corpora: the use of modal and reporting verbs in the expression of writer stance. Language and Computes, 48: 211-230.https://doi.org/10.1163/9789042029248_017.

Oakes, M., (1998),Statistics for Corpus Linguistics. Edinburgh: Edinburgh University Press. Palmer, F.R.(2001). Mood and Modality. Cambridge: CUP. https://doi.org/10.1017/CBO9781139167178.

Palmer, F. R. (1990). Modality and the English modals. London & New York: Longman.

Pang, J.X.& Chen J. (2018). A Comparative analysis of cognitive modality sequence of discourse in international journals.Journal of Fujian Normal University (philosophical social science edition), 2018(01), 97-109.

Jaime Pastor, M., & Pérez-Guillot, C. (2015). A comparison analysis of modal auxiliary verbs in Technical and General English. Procedia Social and Behavioral Sciences, 212, 292-297.

Quirk, R., S. Greenbaum, G. Leech & J. Svartvik. (1985). A Comprehensive Grammar of the English Language[M]. London: longman. https://doi.org/10.1016/j.sbspro.2015.11.375.

Qiao, L.L.et al. (2014),Thinking difference attribution of modal verb and modal sequence misuse [J]. Journal of Changchun university of science and technology (social science edition) 27(05), 93

Römer, U. (2004). A corpus-driven approach to modal auxiliaries and their didactics. How to Use Corpora in Language Teaching. 185-199. https://doi.org/10.1075/scl.12.14rom.

Salager-Meyer, Françoise. (1994). Hedges and textual communicative function in medical English written discourse. English for Specific Purposes 13 (2), 149–170.

Scott, M. (2016). In search of a bad reference corpus. In What's in a Word-list?. Routledge. 99-112.

Skehan, P. (1998). A cognitive approach to language learning. Oxford University Press.

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

Sinclair, J．et aL, (1995). Collins COBUILD English Dictionary. London: Harper Collins．

Sinclair, J. (1998). The lexical item. AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE SERIES 4, 1-24.

Sinclair, J. M. (Ed.). (2004). How to use corpora in language teaching (Vol. 12). John Benjamins Publishing.

Shu, M.Q. (2009). Decategoriarization approaches to English modality:its evolution from prototype. Yunnan Normal University：Yunnan Normal University.

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. The Cambridge handbook of literacy, 112-133.

Sweetser, E. (1990). From Etymology to Pragmatics.Cambridge: CUP. https://doi.org/10.1017/CBO9780511620904.

Tang, L. L. (2013),A survey of the corpus of semantic development of modal verbs for Chinese English majors. Contemporary foreign language research,6,32-36+77.

Tang, J.C. (2006). Syntactic transfer. Guangxi Normal University.

Thompson, P. (2002). Modal verbs in academic writing. In Teaching and learning by doing corpus analysis (pp.

305-325). Brill Rodopi.https://doi.org/10.1163/9789004334236_023.

Tono, Y. (2000). A corpus-based analysis of interlanguage development: analysing POS tag sequences of EFL learner corpora. Practical Applications in Language Corpora, 123-132.

Varttala, T. (2001). Hedging in scientifically oriented discourse. Exploring variation according to discipline and intended audience. Tampere University Press.

Wang H.H. (2015). A corpus-based on modality of modal verbs in English majors' compositions. Northeast University.

Wang, G. (2013). Analysis on relevance between modality and Chinese EFL learners' English proficiency based on corpus. Journal of Huaihai institute of technology (humanities and social sciences edition), 11(07):52-55.

Wei, K. (2009). A contrastive study on theme and thematic progression of in English and Chinese abstracts of research articles. Northeast normal university.

Wen, Q.F. et.al, (2008). Spoken and Written English Corpus of Chinese Learners (version 2.0).Beijing: Foreign Language Teaching and Research Press.

Wu J.H. (2016). Promotion function of modal verbs in the introduction sections of international journal articles, a corpus-based analysis. Zhejiang University.

Wärnsby, A. (2003). (De)coding epistemic modality in English and Swedish. [On-line].Available: http://www.englund.lu.se (2017,June 18).

Xie, J.C. (2009), the study of modal verbs in the English-Chinese parallel corpus --A example of the modal verb "must".Journal of the Yangtze university journal (social science edition) 6 , 581-61.

Yan, P.F. (2017). Research on Modal Sequences in Chinese Science and Engineering Learners'L2 Academic Writing. Shandong foreign language teaching, 38(03), 40-51. https://doi.org/10.16482/j.sdwy37-1026.2017-03-005.

Yang, X. (2018). A Corpus-Based Study of Modal Verbs in Chinese Learners' Academic Writing. English Language Teaching, 11(2), 122-130.

Yu ,J. (2016). A corpus-based comparative study of modal sequences in English academic papers by Chinese and native scholars. Overseas English,01, 200-202+218.

Zhao, L.Z. (2018). A contrastive study of evaluation chunks in Chinese and native' s academic texts. Research on Contemporary Foreign Languages, 01, 31-35+108. Author, (2011). A study of modal sequences in English major's argumentations. China English education, 2, 1-13. https://doi.org/10.19716/j.1672-4720.2014.02.003.

Zhang Z.B,(2003), "A new English grammar course (fourth edition)", Shanghai: Shanghai foreign language education press.