

# Epistemic Modality in the Argumentative Essays of Chinese EFL Learners

Chunyu Hu<sup>1</sup> & Xuyan Li<sup>1</sup>

<sup>1</sup>School of English for International Business, Guangdong University of Foreign Studies, China

Correspondence: Chunyu Hu, Associate Professor of School of English for International Business, Guangdong University of Foreign Studies, Guangzhou 510420, China. Tel: 86-20-3631-7183. E-mail: gwncy@gdufs.edu.cn

Received: March 9, 2015 Accepted: April 27, 2015 Online Published: May 28, 2015

doi:10.5539/elt.v8n6p20 URL: <http://dx.doi.org/10.5539/elt.v8n6p20>

## Abstract

Central to argumentative writing is the proper use of epistemic devices (EDs), which distinguish writers' opinions from facts and evaluate the degree of certainty expressed in their statements. Important as these devices are, they turn out to constitute a thorny area for non-native speakers (NNS). Previous research indicates that Chinese EFL learners differ significantly from the native speakers (NS) in marking epistemic modality. One problem of previous studies is that the essay topics are not well controlled, which makes it somewhat ambiguous as to whether the observed linguistic discrepancies are caused by the NNS/NS difference or by the topic differences. This paper sets out to explore much more comparable data from International Corpus Network of Asian Learners of English (ICNALE). The results show that while both NS group and NNS groups are heavily dependent on a restricted range of items, the manipulation of epistemic modality is particularly problematic for the L2 students who employ syntactically simpler constructions and rely on a more limited range of devices, as already discovered in the previous studies. Nevertheless, this study also shows that the most proficient L2 students modify their statements with less certainty markers and more tentative expressions than do their L1 counterparts, and that all learner groups, regardless of their overall language proficiency, use less boosters than L1 writers, which is in sharp contrast with previous studies. The ability to mark epistemic modality has much to do with L2 proficiency. While lower-band students exhibit a heavy reliance on a narrower range of items for strong assertions, higher-band students tend to be more tentative and demonstrate a more native-like use of some Eds. The observed patterns are explained in the light of the inherent properties of English EDs, the imperfect modal instruction and learner factors.

**Keywords:** argumentative writing, epistemic modality, hedges and boosters, learner corpus research, Chinese EFL learners, L2 proficiency

## 1. Introduction

Epistemic modality has been a fascinating area to philosophers, logicians as well as linguists, and has been approached from a large number of different perspectives. From a cognitive-pragmatic perspective, "epistemic modality concerns an estimation of the likelihood that (some aspect of) a certain state of affairs is/has been/will be true (or false) in the context of the possible world under consideration" (Nuyts, 2001, pp. 21-22; cf. also Coates, 1987).

Epistemic modality manifests itself via a great variety of means in English, including modals (*could, may, might, will, would*), adjectives (e.g., *possible, likely, certain, necessary*), adverbs (e.g., *maybe, perhaps, possibly, probably*), nouns (e.g., *possibility, probability, certainty*), mental verbs taking sentential complements (e.g. *think, believe, doubt*), and other forms. In this study, the linguistic expressions of epistemic modality will be labeled as epistemic devices (henceforth EDs). Hedges and booster are two types of EDs that are well recognized in the literature. According to Hyland (2005, p. 130), hedges include downtoners that reduce force of statements (*fairly, almost, partly*), frequency adverbs that make statements indefinite (*usually, sometimes*) and devices that decrease responsibility for truth (*probably, perhaps, may*); boosters are emphatics that reinforce truth value (*certainly, demonstrate, really*), and amplifying adverbs that strengthen verbs and adverbs (*totally, always*).

EDs are crucial to academic writing where authors have to distinguish opinions from facts and evaluate their statements in acceptable and persuasive ways. EDs not only convey the writer's confidence in the truth of proposition, but also help contribute to a relationship with readers. Therefore, the appropriate use of EDs has

been proved to be of great importance to L2 learners, because they “need to make claims and assertions which academic readers judge to be warranted and which reflect appropriate social interactions” (Hyland & Milton, 1997, p. 183). In addition, a wide range of knowledge of epistemic modality can help L2 writers “to have at their disposal a repertoire of devices that allow them to make claims with the exact degree of certainty or doubt that they intend” (McEnery & Kifle, 2002, p. 183).

Nevertheless, epistemic modality is generally acknowledged to be difficult for both first and second language learners to acquire (Holmes, 1988; M. Bloor & T. Bloor, 1991). There are a number of reasons to account for the well-observed phenomenon. To begin with, many EDs are polypragmatic, that is, they can simultaneously convey a range of different meanings. Take the modal *may* for example, it has at least three types of meaning: expressing *permission*, *root possibility* and *epistemic possibility*. Second, the meaning of a device often overlaps with many other modal expressions. Thus, in order to master the use of *might* one will have to not only distinguish it with *may*, but also differentiate it from *could* in context like *John might be in his office now* and *John could be in his office now*. The use of EDs is further complicated by the fact that they perform quite differently in different types of text or speech genres. *Maybe*, for example, occurs much less frequently in writing than in speaking. *Perhaps*, on the contrary, is rarely used in conversation but extremely common in academic prose (Biber, Johansson, Leech, Conrad, & Finegan, 1999).

It is thus of critical significance to examine how Chinese EFL learners express epistemic modality that either strengthen or weaken the writer’s commitment to the statement in argumentative essays, and how the use of EDs varies due to differing L2 proficiency.

## 2. Literature Review

The use of EDs by Chinese EFL learners is somewhat well-documented. Hu, D. Brown and L. Brown (1982) found Chinese L2 writers to be more direct and authoritative in tone and to make more use of strong modals than NSs. Having examined 27 English essays written by Hong Kong freshman students, Allison (1995) noticed that those writers tended to choose more boosting devices to modify their statements with a strong voice rather than an alternative voice.

The first systematic and influential investigation of EDs in L2 writers’ essays was conducted by Hyland and Milton. Having examined the expressions of doubt and certainty between Hong Kong and British high school leavers, Hyland and Milton (1997) found that L2 writers employed syntactically simpler constructions, relied on a more limited range of devices, offered stronger commitments to statements and exhibited greater problems in conveying a precise degree of certainty. L1 writers, however, employed more tentative devices, with about two thirds of EDs functioning as hedges (only a third in L2 writing). Their study also reveals that L2 students with higher English proficiency tended to make more tentative statements. Another learner corpus research also clearly demonstrates an underuse of hedging devices by Hong Kong EFL learners, leading to writing that is ‘too direct’ (Flowerdew, 2000).

Recent years have also witnessed a few studies of EDs employed by mandarin-speaking EFL learners in their argumentative essays. Chen (2012) reported that despite the significantly similar normalized frequencies, Chinese EFL learners employed a restricted range of EDs and made strong and unwarranted assertions compared to NS writers. This study also suggests L2 learners make progress in qualifying their statements with EDs as learners’ proficiency increases. Meng (2012) reports that Chinese students tend to depend far more heavily on several predominantly speech forms to express epistemic meanings, and that over half of the devices in the NNS data function as boosters, while most items in the NS sample are hedges, suggesting that the academic writing of Chinese students is indeed characterized by firmer assertions and stronger commitments than NS discourse.

The direct and unqualified writing is not only typical of Chinese EFL students, but seems to pertain to most L2 writers regardless of their L1 backgrounds. A recent study, for instance, shows that Korean EFL learners take a stronger stance in their statements and rely on a restricted range of EDs with simpler constructions compared to native speakers (cf. Kim & Suh, 2014). Using fewer hedges and more boosters is also found in the essays written by Arab students (Scarcella & Brunak, 1981), Bulgarian academics (Vassileva, 2001), Korean students (Oh, 2007; Kim & Suh, 2014). Interestingly, the problematic use of EDs is also evident in academic essays by L2 writers at postgraduate level (Back, 2011)

The seemingly universal pattern is challenged by McEnery and Kifle who have largely replicated Hyland and Milton’s study by comparing learner corpus compiled from 92 short argumentative compositions written by Eritrean second-year university students aged around twenty, and native corpus of argumentative essays by English native speakers around 16 years of age. Both corpora contain approximately 22,000 words. McEnery and Kifle (2002) found that in many respects there are striking differences in NS and NNS’s use of modal

expressions. For instance, MAY occurs twice frequently in NNS corpus than it does in NS corpus; WOULD, however, occurs almost 8 times frequently in NS corpus than it does in NNS corpus (McEnery & Kifle, 2002, p. 188). Nevertheless, this study reveals that non-native students use more tentative EDs, while the native English speakers employ stronger devices, conveying a higher degree of confidence. McEnery and Kifle (2002) attributed the more tentative use of EDs by L2 writers to modal instruction.

McEnery and Kifle's (2002) study suggests that topics might affect the way writers establish their epistemic stance. Considering that comparability is one of the most crucial requirements for learner corpus research (Hunston, 2002, p. 206) and that the essay topics in previous studies are not well controlled, this paper sets out to explore much more comparable data produced by L1 and L2 writers with the purposes of investigating how similar or different the two groups employ EDs and how L2 learners' use of EDs changes over their general English proficiency.

### 3. Method

#### 3.1 Research Questions

This study sets out to investigate how Chinese EFL learners use EDs to distinguish their opinions from facts and evaluate the certainty of their assertions when writing argumentative essays. More specifically, it aims to address the following three questions:

RQ1: In what aspects do the advanced Chinese EFL learners differ from the native speakers in their use of EDs in argumentative essays?

RQ2: To what extent does learner's mastery of EDs develop with their general English proficiency?

RQ3: What factors might be at work and how they interact to influence EFL learners' use of EDs?

#### 3.2 Data

The data comes from the written component of the International Corpus Network of Asian Learners of English (ICNALE), which consists of 1.3 million words of 5600 essays written by 2600 college students in 10 Asian countries and areas as well as 200 English Native Speakers (Ishikawa, 2013). The essays are of argumentative mode with two set topics:

Topic 1: *Is it important for college students to have a part time job?*

Topic 2: *Should smoking be completely banned at all the restaurants in the country?*

For the EFL participants, based on their scores in the standard L2 proficiency tests such as TOEIC or TOEFL or in the standard vocabulary size test (VST), their proficiencies were classified into four levels: A2 (Waystage), B1\_1 (Threshold: Lower), B1\_2 (Threshold: Upper), and B2+ (Vantage or higher). These are identical with the levels proposed in the CEFR (Common European Framework of Reference). Since this study focuses on Chinese EFL learners' use of EDs, only four sub-corpora are employed, as demonstrated in Table 1.

Table 1. Corpus information

	CHN_A2_0	CHN_B1_1	CHN_B1_2	ENS1
Number of participants	50	100	100	100
Number of essays	100	200	200	200
Tokens in corpus	22315	46390	49010	44694
Types in corpus	1798	2861	3366	3393
Average sentence length	15.26	15.75	16.19	24.42

The ICNALE-Written is a new learner corpus designed for a reliable contrastive interlanguage (IL) analysis of varied English learners in Asia. The conditions for writing (topics, time, length, etc.) are strictly controlled, which guarantee a high comparability of data, allowing the observation of a more reliable comparison of the range and frequency of EDs among groups under investigation.

#### 3.3 Data Analysis Procedure

The overall frequency and density levels were assessed using WordSmith Tools 5.0. The searching words were based on a list of 163 items from Kim and Suh (2014), which, in turn, was based on Hyland's (2005)

classification of hedges and boosters. The concordance output was then subjected to a qualitative analysis in order to isolate true EDs from other entries such as *will* as in 1) which is used as a noun instead of a modal, *would* in 2) which is actually the wrong use of *had*, or *could* in 3) which expresses ability instead of epistemic meaning:

1) If you are busy with study, you don't need to take it. But if you have the time and urge and the **will** to challenge yourselves, you'd better have a try. (CHN\_A2\_0) N: non-modal

2) Today it is more and more popular to take a part-time job in the university, while many people think students **would** better not to do it. (CHN\_A2\_0)

3) In addition, he thought if he **could** smoke a good cigarette, it seemed to be that he was a rich man as well. (CHN\_A2\_0) (Ability)

The resulting data were presented with comparison in terms of distribution of semantic category, grammatical category, and individual devices of epistemic modality.

#### 4. Results

##### 4.1 Semantic Distribution

The total number of lexical devices used to express epistemic meanings in the corpora is shown in Table 2. Clearly, all EFL learner groups used fewer hedges and boosters, compared with their native counterparts. The analysis also reveals remarkable similarities in the overall frequencies among the three learner groups. In addition, all learner groups, like the native group, used notably more hedges than boosters in their essays, which is in sharp contrast with many previous studies reporting that L2 writers tended to take a stronger stance because they used more boosters than hedges except those in the advanced proficiency band (cf. Hyland & Milton, 1997; Chen, 2012; Meng, 2012; Kim & Suh, 2014).

Table 2. Frequencies of hedges and boosters in each group (per 100,000 words)

	CHN_A2_0	CHN_B1_1	CHN_B1_2	ENS1
Hedges	1698**	1664**	1696**	1938
Boosters	1138**	1183**	1102**	1363
Total	2837**	2848**	2797**	3300

*Note:* 1) this study adopts the log-likelihood ratio test to check whether the difference in L1 and L2 students' frequency of a certain category/device is statistically significant, which is more precise than chi-square test; 2) \* signifies the existence of striking difference ( $p < 0.05$ , log-likelihood ratio is greater or equal to 3.84), \*\* signifies the existence of highly significant difference ( $p < 0.01$ , log-likelihood ratio is no less than 6.63).

Perhaps it is more illustrating to examine the semantic categories on a more specific scale. As evident in Table 3, higher ability students modify their statements with less certainty markers and more tentative expressions than do their L1 counterparts. This is clearly different from previous research in that "the native speakers employ a higher proportion of tentativeness than any NNS group" (Hyland & Milton, 1997, p. 196). Another striking difference is that the L1 essays contain 84% more items expressing probability than L2 essays written by the most advanced group.

Table 3. Semantic distribution of epistemic devices (per 100,000 words)

	CHN_A2_0	CHN_B1_1	CHN_B1_2	ENS1
Certainty	1788	1837	1616 *	1747
Probability	426**	351**	449**	832
Possibility	350	403 *	471 **	345
Usuality	220	192**	196**	255
Approximation	54**	65**	65**	121

When comparison is made between CHN\_A2\_0 and CHN\_B1\_2, it is evident that higher ability students employ a higher proportion of tentativeness than lower ability ones, suggesting that the more proficient students tend to be less assertive and more tentative in making statements, which confirms findings in many previous studies.

#### 4.2 Grammatical Distribution

It is also useful to categorize the items into grammatical classes for comparison. Table 4 shows broad agreement between the most advanced L2 group and the L1 group in the use of adverbials and nouns to express degrees of certainty, but marked differences in the use of modals, lexical verbs and adjectives.

Table 4. Grammatical distribution of epistemic devices (per 100,000 words)

	CHN_A2_0	CHN_B1_1	CHN_B1_2	ENS1
Modals	1170*	1196	1132*	1269
Lexical verbs	784**	783**	761**	1049
Adverbials	753	694**	788	812
Adjectives	126*	173	116**	168
Nouns	5	2	0	2

A comparison between the lower proficiency band (A2\_0) and the higher proficiency band (B1\_2) reveals substantial similarities, indicating that the increase of proficiency may not lead to marked changes in their preferences for grammatical realization of epistemic meanings.

#### 4.3 The Use of Individual EDs

The broad comparisons of the grammatical and semantic distribution of EDs provide an overall understanding of the differences and similarities among these groups. However, such an overall insights is far from sufficient, as different preferences for certain devices may be compromised in the overall frequency comparison. Therefore, it is necessary to track down the differences in the use of individual devices.

Table 5 lists the most frequent ten EDs used by different groups. At first glance, there are considerable similarities of usage, with *think*, *will*, *would*, *really*, *may*, *always*, and *could* occurring among the top 10 most frequently used devices of both L1 and L2 writers, although with strikingly different frequencies. Further investigation shows that the top 10 EDs accounts for 35% of total EDs in ENS1, 61% in CHN\_B1\_2, 75% in CHN\_B1\_1, and 73% in CHN\_A2\_0, strongly suggesting that learners in this study have a relatively limited range of EDs at their disposal.

Table 5. The top ten EDs in four groups

Rank	CHN_A2_0		CHN_B1_1		CHN_B1_2		ENS1	
	EDs	Freq.	EDs	Freq.	EDs	Freq.	EDs	Freq.
1	will	834	will	804	will	661	think	604
2	think	385	think	373	think	324	will	557
3	may	233	may	254	may	320	would	445
4	in my opinion	143	know	164	know	153	really	181
5	really	121	would	91	really	104	may	141
6	know	103	in my opinion	88	would	86	feel	101
7	maybe	85	always	80	in my opinion	80	always	76
8	always	67	really	80	maybe	71	could	72
9	believe	67	maybe	58	never	63	might	67
10	would	58	could	54	could	61	probably	65

Table 6 and 7 present the EDs that are overused and underused by EFL learners compared to the native group. Clearly, learners show a general inclination to overuse modals *will* and *may*, lexical verbs *know*, and some adverbs such as *maybe*, *generally*, *of course* and *in my opinion*. On the other hand, learners tend to underuse modals *would* and *might*, lexical verbs *think* and *feel*, and modal adverbials such as *probably*, *in general*, and *really*.

Table 6. The overused EDs by Chinese EFL learners (per 100,000 words)

	CHN_A2_0	CHN_B1_1	CHN_B1_2	ENS1
will	834**	804**	661 **	557
may	233**	254**	320**	141
must	40 **	9	12*	4
know	103**	164**	153 **	54
as we (all) know	5 **	24 **	39 **	0
known	54 **	39	49 **	25
(as) (it) is known to	31 **	24 **	43 **	0
obviously	0 **	24 **	22 **	7
of course	45 **	54 **	37**	13
in fact	22 **	30 **	24 **	4
surely	4	9*	31 **	2
generally	22 **	15 **	12 **	0
maybe	85 **	58 **	71 **	18
from my perspective	13 **	9 **	10 **	0
in my opinion	143 **	88 **	80 **	2

Table 7. The underused EDs by Chinese EFL learners (per 100,000 words)

	CHN_A2_0	CHN_B1_1	CHN_B1_2	ENS1
would	45**	91**	86 **	445
might	0**	17**	22 **	67
think	390*	373**	324 **	606
I think	273**	241**	198 **	425
feel	13 **	4 **	8 **	101
I feel	0 **	2 **	0 **	72
appear	0 **	0 **	0 **	13
really	121 **	80 **	104 **	181
definitely	4 **	9 **	14 **	29
clearly	13	13	2 **	18
almost	18 *	24	14 **	36
sure	13 *	11 **	2 **	29
always	67	80	41 **	76
in general	0 **	6 *	0 **	16
probably	0 **	15 **	18 **	65
perhaps	0 **	22	4 **	18
somewhat	0 **	2 **	4 **	16

It is by no means easy, based on the above three tables, to draw any solid conclusion regarding the use of EDs among three learner groups. Generally speaking, the two higher bands are closer to the native norm in terms of the frequency of individual EDs. Some more complex modals (*would* and *might*), for example, are clearly more frequent in CHN\_B1\_1 and CHN\_B1\_2 than in CHN\_A2\_0. In general, it seems that most EDs overused by lower proficiency bands belong to the expression of certainty, which is in sharp contrast with the most advanced group in this study (cf. also Table 3).

If we look at the context of EDs used by learners, we will probably find something more interesting. For instance, students in the lowest proficiency band frequently used the modal *must*. A closer investigation indicates that some occurrences of *must* can be replaced by other EDs. It might be more appropriated to use *would* in example 4) and 5).

4) So if the fire causes in a very short time, there are a number people **must** be died. (CHN\_A2\_0)

5) So why not try to find a part time work? And then our college life **must** be more colorful. (CHN\_A2\_0)

As presented in Table 6, learners relied heavily on *know*, and it was often used in the phrase *as we (all) know, as is known to all* and *it is known to all*:

6) **As we all know**, the harm of the second hand cigarettes to people is greater than to smokers. (CHN\_SMK\_112\_B1\_1)

7) **As it is known to us**, college tuition is not so low that every student can afford it. (CHN\_PTJ\_019\_B1\_1)

Native students, however, often used *know* in the context like 8) and 9):

8) It is commonly known that cigarettes cause cancer and **it is well known that** the damage to the lungs of secondhand smokers is often severe, if not fatal. (ENS\_SMK\_080\_XX.0)

9) **I know that** smoking is bad for my health, and I do plan to quit one day. (ENS\_SMK\_042\_XX.0)

The overuse or underuse of some EDs needs to be analyzed with caution. In fact, the underuse of *think* or *I think* by Chinese EFL learners as attested in this study is in sharp contrast with L2 writers' overuse of *think* as reported in the previous research (Hyland & Milton, 1997; Ringborn, 1998; Kim & Suh, 2014). A possible explanation is that L1 writers in this study were required to write an essay around 300 words within a short period of time. It is likely that they transferred patterns of use from spoken English into the written medium, and notably a high degree of topic sensitivity in the use of particular modals. In addition, it has to be noted that not all native writers are equally competent in the manipulation of rhetorical aspects of argumentation. In fact, L1 writers in this study are also heavily dependent on a narrow range of items, as demonstrated in Table 4.

## 5. Discussion

The aim of this section is to discuss the empirical findings described in Section 4 and in the related literature with a view to seeking possible answers to RQ3. Following previous research (VanPatten, 2004; Hu, 2010), we divide factors that influence L2 writers' use of EDs into three dimensions: the inherent properties of English EDs, L2 modal instruction and learner factors.

### 5.1 The Inherent Properties of Epistemic Devices in English

There is no doubt that the nature of the target language itself is an important factor affecting L2 acquisition. The inherent properties of English EDs are well discussed by Hyland and Milton (1997, p. 185) in terms of semantic and pragmatic complexity: 1) many EDs can simultaneously convey a range of different meanings; 2) EDs not only convey the writer's confidence in the truth of referential information, but also help contribute to a relationship with the reader; and 3) epistemic meanings can be signaled in many different ways.

Needless to say, semantic complexity plays an important role in the acquisition and use of EDs. *Would*, for example, is very complex in that it has quite a number of meanings (Coats, 1983; Perkins, 1983; Palmer, 1990), which normally baffles EFL learners regardless of their L1 background (Hinkel, 2002). This accounts for the relative infrequencies of this modal in learner groups of this study.

In addition to semantic and pragmatic complexity mentioned by Hyland and Milton, there are other inherent properties that cannot be ignored. Form complexity, for instance, is an important factor to determine which item learners prefer, as is in the case of a competing adverbs marking epistemic possibility. In view of ease of articulation, *probably* is formally more complex than *maybe*. The complex forms require more effort to use, and may have the potential to hinder the use of them. Thus when two forms compete for the same or similar function, the shorter and easier one is constantly gaining the upper hand. For instance, the easier pronunciation of *maybe* makes it much preferred in spoken discourse, and once it has become entrenched, learners are likely to transfer it

habitually in writing, if they are unaware that *maybe* is rarely used in academic writing.

Frequency and saliency are also important properties that affect L2 acquisition and use of EDs. It is well observed that some EDs are more frequent and salient than others, and these items are normally noticeable to learners. Saliency is related to frequency in many intriguing ways. The frequent form is normally the one easily to get noticed, thus salient. However, frequency cannot guarantee saliency. For instance, *might* in the normal speech is possibly of low saliency in the language stream, since it is usually pronounced very fast by the native speaker. To the contrary, it is much easier for learners to notice *maybe*, because it is normally used in the initial position. Many modal adverbials such as *in my opinion* and *from my perspective* often appear in the initial position of the sentence. Infrequent as they are, they are quite salient for L2 learners, and are most likely to be frequently used by learners, as attested in this study.

### 5.2 Modal Instruction

It is well recognized that instruction plays an important role in SLA, as it provides structured, differentiated input that assists noticing by focusing attention on and enhancing awareness of language features (Skehan, 1998). Also, according to Schmidt (1990), instruction may prime learners to notice features by establishing expectations about language. In terms of epistemic expressions, previous studies suggest that the significance of these devices is largely ignored or misrepresented in writers' handbooks, style guides and ESL textbooks (Holmes, 1988; Hyland & Milton, 1997).

In this study, the role of modal instruction is decomposed into three aspects: providing instructional input that may or may not reflect the inherent properties of the English EDs, providing the explicit modal knowledge that may help learners to understand the subtlety of the target system or may confuse learners instead, awareness-raising activities that help learners notice the gap between their output and the target norm, and provide the chance for practicing.

In order to examine how instructional input affects output production, Table 8 lists frequencies of five EDs marking epistemic possibility appeared in the six-year middle school textbooks and in the three learner groups as well as the native groups for the sake of reference.

Table 8. Comparison of instructional input and learner output (per 100,000 words)

	Textbooks	CHN_A2_0	CHN_B1_1	CHN_B1_2	ENS1
may	61	238	254	320	141
might	33	0	17	22	69
maybe	19	85	58	71	18
perhaps	20	0	22	4	18
probably	4	0	15	18	65

Note: Frequency data of the five EDs in the textbooks are reprinted from Hu (2010, p. 384).

As demonstrated in Table 8, the most frequent item *may* in the textbooks is also the most frequently used one by learners, and the infrequent form *probably* in the textbooks is also infrequent in learners' output, suggesting that there indeed exists frequency effect as proposed by Ellis (2002). It is also evident that not all learners' modal behavior is traceable to input frequency. For instance, although *might* and *perhaps* are slightly more frequent than *maybe* in the textbooks, it is the latter that was the second most frequently used device to mark epistemic possibility.

Another important role of instruction is to provide learners with instructional support to ease the burden of learning. Instructional support can take on many different forms. Does the instructor draw learners' attention to a modal form? Is the instructor able to help learners to distinguish the subtle differences between semantically close modal forms? Does the instructor make learners aware of the gap between their modal production and the native norms? Does the instructor provide the opportunity for learners to use the modal expressions that learners have learned? In short, if the learner is given strong scaffolding and support, acquisition will be facilitated (for the role of social support based on Vygotsky's (1962, 1978) sociocultural theory, see Lantolf (2000), Lantolf and Thorne (2006), Johnson (2004). Otherwise, if there is no appropriate scaffolding and support, learning may be delayed or even distorted. Previous research indicates that at least devices of epistemic possibility are not well



taught in China (Hu, 2010).

### 5.3 Learner Factors

There are also a number of learner factors that are likely to affect L2 acquisition and use of EDs. In SLA research, factors such as L1 influence, the One-to-One Principle and L2 proficiency are well-recognized (cf. VanPatten, 2004).

#### 5.3.1 L1 Influence

It is well-known that adult L2 learners, especially at the beginning stages of acquisition, have important cognitive limitations related to attention and memory (VanPatten, 1996; Skehan, 1998; Robinson, 2003; among others), and one way to overcome the processing limitations is to resort to processes from their L1 knowledge. Since there already exists a powerful L1 modal system, learners will probably transfer everything they can in order to establish form-meaning relationship. For example, *zhong suo zhou zhi* (*as we all know*) appear frequently in Chinese writing, so Chinese EFL learners tend to transfer them into their L2 writing, leading to the frequent occurrences of *as we all know* and *as is known to all* in their writing.

#### 5.3.2 The One-to-One Principle

The One-to-One Principle states that one form is mapped onto a single meaning. Andersen (1984, p. 79) believes that this principle is a constraint guiding “the construction of a minimal but functional IL system” in the initial stage. There is ample evidence for the work of this principle. When L2 writers opt for *generally*, *in general* seems unnecessary. Similarly, when *maybe* becomes the dominant form to mark epistemic possibility, other devices such as *perhaps*, *possibly* and *probably* becomes less significant. It is clear that adherence to the One-to-One principle results in initially non-targetlike form-meaning association, since an initial connection between a given form and an additional meaning were delayed or suppressed altogether. This situation has been remedied when learners’ overall proficiency improves over time.

#### 5.3.3 L2 Proficiency

The acquisition of EDs is not only anchored in learners’ L1 modal system, but also in the prior L2 knowledge. Perhaps the most prominent role of L2 proficiency lies in the fact that many other factors change with proficiency level (cf. Pienemann, 1989). First, learners with different levels of proficiency have different processing capacity (VanPatten, 1996) and different attentional focus (Gass, 2004). As a result, it is unlikely for the less advanced learners to notice the grammaticalized forms (modals in our case). Advanced learners may also become more sensitive to the stylistic aspect of modal devices: *must*, *maybe* are used much less frequently by advanced learners, indicating that these learners are more sensitive to the register knowledge. Secondly, the role of input factor (frequency, saliency and complexity) may also change when the general proficiency level improves. The infrequent forms usually go unnoticed for beginning learners, yet something which is unusual because of its infrequency may as well stand out for a learner at more advanced stage of learning (cf. Gass & Silinker, 2008). Thirdly, the One-to-One principle may stop to work for L2 learners with advanced proficiency. Finally, advanced learners may rely less on their L1 modal system. Instead, they may resort to the context in which a modal construction appears, and they may abstract some implicit knowledge or patterns regarding a modal form-meaning relationship from richer experience of input, as L1 children do.

### 5.4 The Dynamic Integration of Contributing Factors

Although the identified factors can be investigated in isolation and their significance can be determined, in reality they interact continuously in intricate ways, and they should also be investigated as interacting and converging factors so that we are able to truly see how they interact to produce the observed pattern.

Once again, take devices of epistemic possibility for example. We will first address the question as to what factors make *maybe* the dominant adverb marking epistemic possibility. Of many possibilities, *form saliency* and *ease of articulation* may be two important factors that make *maybe* win out. Compared with *possibly* and *probably*, the two forms that can also be used independently in an utterance, *maybe*, however, is the form that is easy to pronounce. Previous studies in L2 modal acquisition show that *maybe* is the first device to mark epistemic possibility and most frequently used by learners of different L1 backgrounds (Salsbury, 2000). Once *maybe* becomes deep-rooted in the learners’ grammar, it may block other modal adverbs. As for the striking rarity of *might*, there may be a number of possible explanations. Compared with *may*, *might* is the so-called secondary modal, or “modal of modal” (Perkins, 1983), which is more semantically complex and thus is supposed to be acquired later.

## 6. Conclusion

Argumentation is the art and science of civil debate, dialogue and persuasion (Glenn, Miller, Webb, Gary, & Hodge, 2004), which involves presenting the writer's statement, discussing each side of the issue, and justifying the writer's conviction with a focus on the reader. Crucial to successful argumentative writing is the ability to qualify one's statements and express an appropriate degree of certainty and doubt.

By examining more comparable data from ICNALE, this study reveals that while both NS group and NNS groups are heavily dependent on a narrow range of items, the manipulation of epistemic modality in academic writing is particularly problematic for the L2 students. The Chinese EFL learners employ syntactically simpler constructions and relied on a more limited range of devices, as already discovered in the previous studies. Nevertheless, this study also shows that the most proficient L2 students modify their statements with less certainty markers and more tentative expressions than do their L1 counterparts, and that all learner groups, regardless of their overall language proficiency, used less boosters than L1 writers, which is in sharp contrast with previous studies. The ability to mark epistemic modality has much to do with L2 proficiency. While lower bands students exhibit a heavy reliance on more limited items for strong assertion, higher bands students tend to be more tentative and demonstrate a more native-like use of some Eds.

The findings reported in this study have some pedagogical implications for the instruction of epistemic modality. First, teachers can provide more input-processing tasks and focus-on-form activities. By default, L2 learners drive their attention to the content words that are information-loaded rather than the grammatical words. Input-processing tasks, however, can alter the way in which learners perceive and process the input exposed to them (VanPatten, 1996). By directing learners' attention away from their normal ways of processing input, learners will notice plenty of modal input such as *could*, *might* and *would* in native speakers' everyday conversation or writing. Closely related to input-processing instruction is the notion focus-on-form instruction. Focus-on-form instruction, which "overtly draws students' attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication" (Long, 1991, pp. 45-46), can facilitate language acquisition (Long & Robinson, 1998). In our case, teachers need to maximize the chances for a particular modal form to be used in a meaningful situation, so that learners' attention will thereby be drawn to its form.

Second, provide more effective explicit instruction. In L1 modal acquisition, the implicit exposure to the input is usually sufficient for the child to acquire the target modal system. In L2 modal acquisition, although the subtle sense of a modal device could only be fully appreciated through the exposure to sufficient context, formal accuracy may still be an unlikely result without explicit instruction, especially for adult learners who rely much more on the existing L1 modal system in the initial stage to construct the IL modal system. For instance, it is insufficient to teach students that *would* is the past tense form of *will*, that it can be used in the hypothetical context such as *If I were you, I would ...*, and that it can be used to express tentative or polite request such as *Would you like ...*. Instead, it would be more instructive to inform students how the semantic chains "Would: Past form of *will* → Explicit hypothetical → Implicit hypothetical → Tentative/Polite" have evolved by explaining why the expression *I was wandering/hoping/thinking ...* is more polite.

Third, teach a certain epistemic device when learner are developmentally ready. Teachers need to be aware that teaching does not necessary lead to learning. Learners are not passive receivers; rather, they construct their L2 modal system based on previous knowledge such as the Chinese modal system as well as the modal input. Thus, a greater understanding of how these factors impact learners' processes will be an important step in advancing the research agenda in the instructed modal learning.

Last but not least, provide more awareness-raising activities. Learner must also consciously notice the gap between their typical modal output and forms used by native speakers. In doing so, learners may be able to reflect on what is noticed, endeavor to understand its significance, and then reanalyze and restructure the current IL modal system. Given that many learners are unable to recognize the mismatch between what is present in the input and the existing L2 knowledge, they have little motivation for changes in their knowledge base. It is therefore up to teachers to be facilitators to such strategy rather than to leave the learners for their own fumbling.

## Acknowledgements

This study was supported by the 2012 Project of Postgraduate Teaching Reform (12GWYJSJG-04) sponsored by Guangdong University of Foreign Studies and the 2012 Project of the Twelfth Five-Year Plan of Guangdong Social Sciences (GD12XWW04) sponsored by Guangdong Social Science Foundation.

## References

- Allison, D. (1995). Assertions and alternatives: Helping ESL undergraduates extend their choices in academic writing. *Journal of Second Language Writing, 4*, 1-15. [http://dx.doi.org/10.1016/1060-3743\(95\)90020-9](http://dx.doi.org/10.1016/1060-3743(95)90020-9)
- Andersen, R. W. (1984). The one-to-one principle of interlanguage construction. *Language Learning, 34*, 77-95.
- Back, J. (2011). A corpus-based study of hedging in Korean EFL students' academic essays. *Studies in Modern Grammar, 65*, 279-307.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. London, LD: Longman.
- Bloor, M., & Bloor, T. (1991). Cultural expectations and socio-pragmatic failure in academic writing. In P. Adams, B. Heaton, & P. Howarth (Eds.), *Socio-cultural issues in English for academic purposes* (pp. 1-12). Basingstoke: Modern English Publications/British Council.
- Chen, Z. (2012). Expression of epistemic stance in EFL Chinese university students' writing. *English Language Teaching, 5*(10), 173-179. <http://dx.doi.org/10.5539/elt.v5n10p173>
- Coates, J. (1983). *The Semantics of the modal auxiliaries*. London and Canberra: Croom Helm.
- Coates, J. (1987). Epistemic modality and spoken discourse. *Transactions of the Philological Society, 35*, 101-131. <http://dx.doi.org/10.1111/j.1467-968X.1987.tb00714.x>
- Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*, 143-188.
- Flowerdew, J. (2000). Investigating referential and pragmatic errors in a learner corpus. In L. Burnard, & T. McEnery (Eds.), *Rethinking language pedagogy from a corpus perspective* (pp. 145-154). Frankfurt: Peter Lang.
- Gass, S. M. (2004). Context and second language acquisition. In B. Vanpatten, J. Williams, S. Rott, & M. Overstreet (Eds.), *Form-meaning connections in second language acquisition* (pp. 77-90). Mahwah, NJ: Lawrence Erlbaum Associates.
- Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Glenn, C., Miller, R. K., Webb, S. S., Gary, L., & Hodge, J. C. (2004). *Hodges' harbrace handbook* (15th ed.). Boston: Thompson Heinle.
- Hinkel, E. (2002). *Second language writers' text*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Holmes, J. (1988). Doubt and certainty in ESL textbooks. *Applied Linguistics, 9*(1), 21-44.
- Hu, C. Y. (2010). Modal expressions of epistemic possibility in Chinese EFL learners' spoken and written discourses. *Modern Foreign Languages, 33*(3), 379-386.
- Hu, Z., Brown, D., & Brown, L. (1982). Some linguistic differences in the written English of Chinese and Australian students. *Language Learning and Communication, 1*, 39-49.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Hyland, K. (1996). Writing without conviction? Hedging in science research articles. *Applied Linguistics, 17*(4), 433-454. <http://dx.doi.org/10.1093/applin/17.4.433>
- Hyland, K. (1998). Boosting, hedging and the negotiation of academic knowledge. *Text, 18*(3), 349-382.
- Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students writing. *Journal of Second Language Writing, 6*(2), 183-205. [http://dx.doi.org/10.1016/S1060-3743\(97\)90033-3](http://dx.doi.org/10.1016/S1060-3743(97)90033-3)
- Ishikawa, S. (2013). The ICNALE and sophisticated contrastive interlanguage analysis of Asian Learners of English. In S. Ishikawa (Ed.), *Learner corpus studies in Asia and the world, 1* (pp. 91-118). Kobe, Japan: Kobe University. Retrieved from <http://language.sakura.ne.jp/icnale/>
- Johnson, M. (2004). *A philosophy of second language acquisition*. New Haven and London: Yale University Press.
- Kim, C., & Suh, H. (2014). Epistemic rhetorical stance: Hedges and boosters in L1 and L2 students' English writing. *The Linguistic Association of Korea Journal, 22*(2), 61-93.
- Lantolf, J. P. (2000). *Sociocultural theory and second language learning*. Oxford: Oxford University Press.

- Lantolf, J., & Thorne, S. (2006). *Sociocultural theory and the genesis of second language development*. Oxford: Oxford University Press.
- Long, M. H. (1991). The design and psycholinguistic motivation of research on foreign language learning. In B. F. Freed (Ed.), *Foreign language acquisition research and the classroom* (pp. 309-320). Lexington, MA: Heath.
- Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. J. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15-41). Cambridge: Cambridge University Press.
- McEney, T., & Kifle, A. (2002). Epistemic modality in argumentative essays of second-language writers. In J. Flowerdew (Ed.), *Academic discourse* (pp. 182-195). London: Pearson Education Limited.
- Meng, Y. (2012). *Epistemic modality in written English of Chinese English majors: An interlanguage pragmatic study* (Unpublished doctoral dissertation). Northeast Normal University, Jilin, China.
- Nuyts, J. (2001). *Epistemic modality, language, and conceptualization a cognitive-pragmatic perspective*. Amsterdam: John Benjamins Publishing Co.
- Ou, S. (2007). A corpus-based approach to epistemic modality in Korean college students' writings in English. *English Teaching*, 62(2), 1-29.
- Palmer, F. R. (1990). *Modality and the English modals* (2nd ed.). London: Longman Ltd.
- Perkins, M. (1983). *Modal expressions in English*. London: Frances Pinter.
- Pienemann, M. (1989). Is language teachable? Psycholinguistic experiments and hypotheses. *Applied Linguistics*, 10, 52-79. <http://dx.doi.org/10.1093/applin/10.1.52>
- Robinson, P. (2003). Attention and memory during SLA. In C. J. Doughty, & M. Long (Eds.), *The handbook of second language acquisition* (pp. 631-678). New York: Blackwell.
- Salsbury, T. (2000). *The Grammaticalization of unreal conditionals: A longitudinal study of L2 English* (Unpublished doctoral dissertation). Indiana University, Bloomington.
- Scarcella, R., & Brunak, R. (1981). On speaking politely in a second language. *International Journal of the Sociology of Language*, 27, 59-75. <http://dx.doi.org/10.1515/ijsl.1981.27.59>
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129-158. <http://dx.doi.org/10.1093/applin/11.2.129>
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Norwood, NJ: Ablex.
- VanPatten, B. (2004). Input and output in establishing form-meaning connections. In B. VanPatten, J. Williams, S. Rott, & M. Overstreet (Eds.), *Form-Meaning connections in second language acquisition* (pp. 29-47). Mahwah, NJ: Lawrence Erlbaum Associates.
- Vassileva, I. (2001). Commitment and detachment in English and Bulgarian academic writing. *English for Specific Purposes*, 20(1), 83-102. [http://dx.doi.org/10.1016/S0889-4906\(99\)00029-0](http://dx.doi.org/10.1016/S0889-4906(99)00029-0)
- Vygotsky, L. S. (1962). *Thought and language*. Cambridge, Mass.: MIT Press.
- Vygotsky, L. S. (1978). *Mind in society*. Cambridge, Mass.: MIT Press.

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).