English Language & Literature Teaching, Vol. 18, No.1 Spring 2012

Modality in Korean Learners' Spoken Interlanguage

Hyeson Park
(Keimyung University)

Park, Hyeson. (2012). Modality in Korean learners' spoken interlanguage. *English Language & Literature Teaching, 18*(1), 197-216.

This study examines spoken interlanguage of Korean learners of English, focusing on the distribution of modal verbs and devices of epistemic modality. (Semi-) spontaneous speech data were collected from four students participating in a self-organized study group for seven months, which produced a corpus of about 55,000 words. The data analysis reveals the following: 1) The frequency of the modal verbs produced by the learners was lower than that of native speakers; 1.99 vs. 2.32 tokens per 100 words. The range of the modal verbs used by the learners was also very limited, with over-reliance on *can* (43%). 2) The grammatical categories of the devices marking epistemic modality were in the order of adverbs, lexical verbs, and modal verbs, with a high frequency of a few items in each category. 3) Lexical items conveying certainty and modals of obligation were preferred over markers of weaker commitment, resulting in speech characterized by firmer assertions and a more authoritative tone, a potential cause for pragmatic failure. 4) A weak developmental change was observed in the frequency of modal verbs, but not in their functions over the seven month period of data collection. L1 influence, L2 proficiency, mode of communication, and instruction effects are discussed as possible variables involved in the distribution patterns observed.

[modality/epistemic/root/deontic/spoken interlanguage]

I. INTRODUCTION

The linguistic category of modality is related to the concept of possibility and necessity, which together with temporality plays an important role in describing situations or states beyond the actual here and now, a unique feature of human language (Von Fintel, 2006). The traditional division of modality generally accepted by logicians and linguists is

between epistemic and root[1] (Depraetere & Reed, 2006; Lyons, 1977; Nuyts, 2000, 2006; Palmer, 1986; Van der Auwera & Plungian, 1998). Epistemic modality, which conveys the degree of certainty a speaker has about a proposition she expresses, has a relatively well-defined boundary, while root modality, covering such diverse modal meanings as necessity, obligation, permission, prohibition, volition and ability, is difficult to delineate, though the overall meaning of root modality can be summed up as the necessity or possibility of the actualization of situations (Collins, 2009; Depraetere & Reed, 2006).

The modalized sentence consists of two parts: the unmodalized proposition and the semantic domain that expresses a speaker's attitude or opinion of the necessity or possibility of the stated proposition, which can be expressed via phonological, morphological, lexical, or syntactic structures (Bybee & Fleischman, 1995; Coates, 1983; Nuyts, 2000; Palmer, 1986). The selection and utilization of such modal devices by a language user is subjective and variable, influenced by such factors as text types, topics, socio-cultural values, and contexts etc., and it is the interaction of these multiple factors that renders control of modality one of the most difficult, if not overwhelming, tasks L2 learners have to struggle with.

The complexity involved in the learning and use of modality has motivated research on this topic in L2 acquisition, among which two general approaches may be discerned: longitudinal form-function mapping studies, which investigate gradual processes of the acquisition of modality, and corpus-based studies, which analyze characteristic features and patterns of modal expressions in cross-sectional interlanguage corpora, especially in written essays of L2 learners (Aijmer, 2002; Chen, 2010; Ramat, 1999; Hinkel, 2009; Hyland & Milton, 1997; Karkkainen, 1992; McEnery & Kifle, 2002; Oh, 2007; Stephany, 1995; Vold, 2006).

Building upon the findings of previous research, the current study attempts to integrate the cross-sectional and longitudinal approach; we examine distributional patterns of modal expressions in an L2 corpus of about 55,000 words collected longitudinally during a seven month period, which might exhibit some developmental changes in interlanguage grammar of modality. An analysis of spoken interlanguage is another feature that differentiates this study from previous ones, the majority of which have focused on modal expressions in written L2 corpora. Specific topics this study aims to explore are the following:

1) What are the distribution patterns of epistemic and root modal verbs in spoken interlanguage?

[1] Root modality is in turn divided into two subcategories, deontic and dynamic modality. The former includes such meanings as deontic possibility and necessity while the latter ability and volition (Depraetere & Reed, 2006).

2) How does the overall frequency of lexical devices used to express epistemic modal meanings in spoken interlanguage differ from those reported in published literature ?
3) Would there be developmental changes in the use of modal expressions?

This paper is structured as follows: Section 2 provides a review of relevant literature on modality in L2, followed in section 3 by an introduction of the corpus analyzed in this study. The results of data analysis with discussions of relevant issues are presented in section 4. A brief summary and suggestions for future research conclude the paper.

II. PREVIOUS RESEARCH ON MODALITY

The crucial role modality plays in communicative interaction has led scholars in language acquisition to examine the emergence of form-function mapping of modal expressions in developing grammars of L1 children (Choi, 1995, 2006; Papafragou, 1998; Stephany, 1995). In her review of acquisition of modality in Indo-European languages such as English, German, and Greek, Stephany (1995) observed that epistemic modality appeared later than deontic or dynamic modality in children's data, which she attributed to an interaction of several factors: children's cognitive development, the purpose of language use by children, and the nature of linguistic input children were exposed to. What Choi (1995) observed in L1 Korean, however, diverged from Stephany's finding in that in child Korean, epistemic modality began to appear no later than deontic modality, which according to Choi, might be due to the characteristics of the modal markers specific to the Korean language: the morphological saliency and discourse interactional role of the modal suffixes. These observations in L1 acquisition point to a possibility of cross-linguistic or cross-cultural differences in the acquisition and use of modality in L1 and L2.

The developmental gap between root and epistemic modality observed in some of the L1 acquisition studies has not been evidenced in L2 acquisition of modality, presumably due to the mature cognitive ability of adult learners and the relevance of both modalities in adult communication. L2 learners were found to first express both root and epistemic modality implicitly through discourse contexts, followed by a stage in which modality was conveyed by lexical verbs, adverbs, or modal verbs. Modal verbs were first employed to express root modality, while epistemic modality was being expressed by diverse means such as mental verbs and adverbs. Utilization of inflectional categories such as conditional or subjunctive to express modality lagged behind other devices (Dittmar & Ahrenholz, 1995; Ramat, 1999; Park, 2010; Stephany, 1995).

The majority of studies of modality in L2 have focused on distributional patterns of epistemic markers in academic writing, reflecting the significance of modal devices in

manipulating the degree and tone of commitment a writer has regarding the proposition she produces (Aijmer, 2002; Hinkel, 2009; Hyland & Milton,1997; Lee & Park, 2008; McEnery & Kifle, 2002; Min, 2010; Oh, 2007; Vold, 2006). Hyland and Milton (1997) is a representative study in this group, whose framework has often been adopted by studies of modality in L2. In their analysis of academic essays written by British English speakers and Cantonese learners of EFL with a similar age and educational background, Hyland and Milton examined the distribution and frequency of 75 epistemic devices which were grouped in terms of grammatical categories (i.e. modal verbs, lexical verbs, adverbs, nouns, and adjectives) and semantic categories (i.e. certainty, probability, possibility, usuality, approximation).

The general characteristics of the use of the modal devices by the NS and NNS groups were: 1) Though the overall frequency was similar between the two groups (1.83 vs. 1.82 tokens per 100 words for the NNS and the NS, respectively), the range of modal devices utilized by the NNS was more restricted than that of the NS; the five most frequent items accounted for 63% of the total devices in the NNS data, while 10 items were needed to reach 62% in the NS corpus. The learners employed 66 of the 75 epistemic devices while the NS used all of them. 2) The grammatical category of the modal device most frequently produced by the NNS was the modal verb, while both the modal verb and adverb were equally preferred by the NS. As the learners' L2 proficiency improved, the use of adverbials increased while the frequency of lexical verbs dropped. 3) When the modal items were analyzed in terms of the semantic categories, the markers of certainty were predominant in the NNS data, while certainty and probability markers were equally distributed in the NS data, supporting the common observation that NNS writings are characterized by firmer assertions and more authoritative tone, possibly due to their insufficient linguistic knowledge of the epistemic markers with various degrees of certainty and doubt. L2 proficiency was shown to be related to the learners' ability to control the degree of commitment, with the lower level learners producing more markers of certainty than the higher level learners.

Oh (2007), adopting Hyland and Milton's methodology, investigated epistemic modality in the writings of Korean learners of English in comparison with that of English native speakers. Two sets of corpus consisting of university students' academic essays - 100,000 words for NNS and 270,000 words for NS - were analyzed in search of similarities and differences between the two groups in the use of epistemic devices. The overall frequency was similar between the two groups, 2.81 (NNS) vs. 2. 83 (NS) per 100 words, while the range of devices was more limited in the NNS data (79 and 104 out of 110 items for the NNS and NS, respectively), a pattern similar to that found in Hyland and Milton (1997).[2]

[2] The epistemic devices included for analysis in Oh (2007) were 110 items, while they were 75 in

The epistemic marker most frequently used by the Korean students was lexical verbs (43%), followed by adverbs (28%) and modal verbs (16%), while in the NS data, it was modal verbs (30%) that were most frequent, with lexical verbs and adverbs equally split at 20% each. The distribution of grammatical categories observed in Oh is not consistent with that found in Hyland and Milton (1997); the devices preferred by the Korean students in Oh were lexical verbs, specifically *think*, *know*, and *feel*, while the Cantonese students in Hyland and Milton selected modal verbs over other devices. Modal verbs were also favored by the NS in both studies. If Hyland and Milton's observation that the use of lexical verbs and learners' language proficiency negatively correlate is correct, the Korean learners' lower language proficiency may be a factor that resulted in their over-reliance on a couple of lexical verbs.

The epistemic devices utilized by the learners in Oh (2007) were classified based on the semantic categories provided in Hyland and Milton (1997), which is shown in Table 1.

TABLE 1

Semantic Categories of Epistemic Markers

semantic category	NNS	NS
certainty	55%	46%
probability	15%	34%
possibility	11%	11%
usuality	15%	7%
approximation	4%	2%

(adapted from Oh, 2007, p. 160)

The high frequency of the devices conveying certainty indicates the difficulty learners face in manipulating degree of epistemic commitment and assertion, possibly resulting from insufficient language skills or transfer of an L1 rhetorical style, as was also noted by Hyland and Milton (1997).

Variable use of modals induced by L1 influence is shown in Aijmer (2002): modal verbs were generally overused by NNS compared to NS, with different modals preferred by learners from different L1 backgrounds- *can/could* by German, *may* by French, *might* by Swedish learners of L2 English. In addition to L1, essay topics, learners' socio-cultural background, and personal experiences were found to be relevant factors in explaining the distribution patterns of modals in learner essays (Hinkel, 2009).

Hyland and Milton (1997). This explains the higher number of tokens per 100 words in Oh (2.81-2.83) compared to that in Hyland and Milton (1.83-1.82).

Letica (2009) is one of the few studies which analyzed epistemic modality in spoken language. In order to understand the usage pattern of modality in L1 and L2, she compared the epistemic devices produced by the same speakers in their L1 Croatian and L2 English. The results demonstrated that epistemic modal devices were less frequently used in L2 (1.52/100 words) than in L1 (1.99/100 words), while the types of devices used in L1 and L2 were very similar; the range of devices was very limited, the most frequent ones being lexical verbs and adverbs. The learners' L2 proficiency was found to be weakly related to the number of types but not to frequency of modal devices produced, leading the author to note that other factors rather than L2 proficiency might be related to the use of modal markers. With regard to the semantic categories, medial probability was predominant (85%) over certainty (4%) or possibility (11%), a result diverging from Hyland & Milton's (1997) and Oh's (2007) findings. The author, seeing a similar distribution of semantic categories in the learners' L1 corpus, attributed this pattern to L1 influence.

The previous studies of modality involving participants of diverse L1 backgrounds and L2 proficiency and various types of texts covering various topics do not seem to provide a convergent picture, indicating the complex interaction of multiple factors in the use of modal devices. The current study examines modal expressions in speech data collected from Korean learners of English, hoping to shed some light on the usage pattern of modal expressions in spoken interlanguage, which has not been a focus of L2 study so far, and to observe any possible developmental changes in the use of modal markers for the duration of seven months.

III. METHOD

1. Participants and data collection

The data were collected from four students participating in a study group organized to practice oral English. Of the four, three were undergraduate students and one a graduate student, the organizer of the group. The students' English proficiency was judged to be at a mid- to high-intermediate level, with TOEIC scores ranging from low 800s to low 900s. The graduate student (a female) was an English Education major with a study abroad experience of eight months. Two of the undergraduate students (one male and one female) were English majors with no previous study abroad experience. The last one was a male student majoring in Tourism Management, who had spent one year in Japan.

The data collection started in June 2007, when the group had already been meeting regularly for three months, and continued till December 2007. The students met twice a week for about 50 minutes each, totaling 50 meetings over the seven month period. The members took turns in selecting discussion topics with accompanying reading materials. The topics covered were mostly related to, but not limited to, current issues selected from newspapers, which included such topics as the Miss Universe contest, capital punishment, child prostitution in Korea, single motherhood, gastric bypass surgery for severe obesity, on-line shopping, and English education in Korea etc. The students were expected to read the materials selected by the discussion leader prior to each meeting and bring questions or issues for discussion. Each meeting began with a brief introduction of the topic by the leader, who functioned as a moderator, eliciting opinions and reactions of the participants on a given topic and coordinating the discussion not to be dominated by a single member. Twelve of the 50 meetings were recorded with the permission of the participants using an MP3 player.

2. Data analysis

Three graduate assistants cooperated in transcribing the data, each in charge of four of the 12 sessions. The whole transcript was later checked and corrected by another graduate assistant to enhance the accuracy of the transcription. A corpus of 55,072 words was compiled out of about 600 minutes' of recording. A quantitative analysis was first conducted using the 'search' function of the Microsoft Word program, focusing on the distribution of root modal verbs as well as the 75 epistemic modal devices presented in Hyland and Milton (1997). Next, the meaning and function of each of the items were examined in light of the context they appeared in, discarding polysemous items which were not used as a modal marker in a given context. The remaining items were first classified according to their grammatical categories: modal verbs, lexical verbs, adverbials, adjectives, and nouns. The modal verbs were further classified based on Coates (1983), as shown in Appendix.

Of the whole modal devices found, epistemic modal devices were selected and classified according to the semantic categories proposed by Hyland and Milton (1997):

certainty: actually, certainly, definitely, indeed, in fact, know, think, will
probability: probably, quite, believe, seem, would
possibility: perhaps, possible(ly), maybe, possibility, may, might
usuality: always, often, usually
approximation: about, almost, nearly, kind of

IV. RESULTS AND DISCUSSION

1. Overall frequency of root and epistemic modal verbs

The types and frequencies of the modal verbs used to express root and epistemic meanings in the learner data are summarized in Table 2. A total of 1,095 modal verbs with 11 types were produced by the learners, which amounts to 1.99 tokens per 100 words. The three most frequent modals in the learner data were *can*, *will*, and *should*, which accounted for 68.1% of the total modals produced.

TABLE 2

Frequency of Modal Verbs in the Learner Data

	raw frequency	per 100 words	per total modals
can	471	.85	43.0
will	158	.29	14.4
should	117	.21	10.7
would	85	.15	7.8
could	76	.14	6.9
be going to	74	.13	6.8
have to	65	.11	5.9
need to	32	.058	2.9
must	11	.019	1.0
might	5	.009	0.46
may	1	.002	0.09
total	1095	1.99	

To compare with a native usage, the data provided in Biber, Johansson, Leech, Conrad, and Finegan (1999)[3] were referred to, focusing only on the statistics for the 11 modals in the conversation register. The frequency of each of the modal verbs was converted from per million words to per 100 words, as shown in Table 3.

The total frequency of the modals in the native data is higher than that of the learners: 2.32 tokens per 100 words. The modals *will*, *can*, and *would* were most frequent in the native usage. Though the total percentage of the three most common modals does not differ greatly between the NNS and NS groups (68.1 vs. 61.1%), the frequencies of the three common modals are equally split in the NS data while the NNS usage is heavily skewed toward *can* (43%), which will be discussed below.

[3] Biber et al. (1999) is a comprehensive reference based on a corpus of over 40 million words of English text from conversation, fiction, newspapers, and academic prose.

TABLE 3

Frequency of Modal Verbs in Native Data

	per million words	per 100 words	per total modals
will	5600	.56	24.1
can	4400	.44	18.9
would	4200	.42	18.1
be going to	2200	.22	9.5
could	2000	.20	8.6
have to	1800	.18	7.8
should	1000	.10	4.3
must	800	.08	3.4
might	800	.08	3.4
may	200	.02	0.86
need to	200	.02	0.86
total	23200	2.32	

(adapted from Biber et al., 1999)

Root modality such as possibility, ability, intention, necessity, volition and obligation was mainly conveyed by the learners with such modal verbs as *can*, *will*, *should*, and *have to*, the most frequent one being *can*. The tendency to overuse *can* was also observed in Japanese students' writings analyzed by Hinkel (2009), who attributed this tendency to the Japanese speakers' transfer of L1 pragmatic strategy; that is, "qualifying the weakness of personal judgment or delimitedness of one's view of an event's realization" (p. 677). However, according to Hinkel (2009), the Chinese and Korean students in her study did not produce as many *cans* as the Japanese learners. Rather, these two groups' use of *can* was comparable to that of native English speakers. The difference between Hinkel and the current study with regard to the use of *can* might be related to such factors as the types of the data (written vs. spoken), topics covered in the essays and the discussions, and the learners' language proficiency[4].

Let us examine the usage pattern of *can* in more detail since it accounts for half of the modals produced by the learners. The majority of *can* found in the learner data signaled root possibility (1a-b) (52.9%), followed by ability (1c) (12.3%).

(1) a. They really want to show their body and advertise them and then they *can* earn a lot of money. (file 1)

[4] Oh (2007), which focused on epistemic modality, did not include *can*, a deontic modal marker, in her analysis. An investigation of the use of *can* with all the relevant variables controlled is needed to determine the cause of the overuse of *can* by the Korean learners in our study.

b. People *cannot* live maybe, three days without water and seven days without food. (file 6)

c. She maybe she *cannot* take care of their child very well. (file 10)

One noticeable finding is that *can* was also used as an epistemic marker in assertive contexts (7.9%) as shown in the examples in (2), which is not usually sanctioned in native usage, though a recent development of epistemic *can* has been reported in native usage even in assertive contexts (Coates, 1995; Collins, 2009).

(2) a. Even though she did on purpose, I think a lot of volunteer work and then, she can change, a lot of judges *can* just put a value her experience. (file 1)

b. Even though someone doesn't have a good diploma she or he *can* have good ability but there are no many chances for her or him. (file 3)

c. If they, they don't want they cannot ... what they want, they *can* kill the rest of hostages maybe, (file 5)

The modal verb better suited to express the intended meaning of the speakers in (2) appears to be *may*, delivering the speakers' tentative judgment or uncertainty regarding the proposition their utterances express. As probable causes for the overuse of *can* in epistemic contexts, L1 transfer and influence of formal instruction may be considered.

The Korean expression corresponding to *can* is the peripheral modal construction *-swu issta*[5]. Differently from English, in which each of the modal auxiliary verbs has dual functions as root and epistemic markers, Korean periphrastic modal constructions are not polysemous. One exception, however, is the modal construction *-swu issta*, which is four-way ambiguous: root possibility, epistemic possibility, permission, and ability (Ammann & Van der Auwera, 2002; Kim, 1998; Lim, 2003). The Korean learners in our study appear to have transferred the multiple functions of *-swu issta* including the epistemic possibility when they used *can*.

Analyses of modal verbs in secondary school English textbooks published in Korea have shown that *can* is the most common modal verb. For example, Gwon's (2009) analysis of eight middle school English textbooks revealed that *can* accounted for 31.3% of 4863 tokens of modal verbs, with *will* (29.8%) and *would* (14.1%) next on the list. A similar result was found in Choi's (2011) analysis of six high school textbooks: *can* (33.2%), *will* (21.6%) and *would* (16.5%). Considering the role of textbooks as the main source of

[5] Modal auxiliaries or periphrastic modal constructions are a very productive way of conveying modal meaning in Korean. The canonical structure of the periphrastic modal construction is <main verb + connective + auxiliary verb + tense/aspect + sentence ending suffix >.

language input for Korean learners of English, it is not unexpected to see the high frequency of *can* in the learner data.

Instruction effect may also be related to the use of *can* in epistemic contexts. The functions of *can* introduced in well-known English grammar books include 'possibility' together with 'ability' and 'permission' while the main function of *may* is also described as 'possibility'[6]. By using one and the same terminology in describing the two different types of possibility, i.e. root vs. epistemic, grammar books and teachers might have led the learners to erroneously regard *can* as synonymous to *may*.

The modals *will*, *would*, *could*, and *be going to* were used as root and epistemic modal markers in accordance with the native usage (Coats, 1983). *Will*, for example, was produced 158 times, of which 69 were used to express prediction or predictability, that is, epistemic meanings, while 48 were employed to mark root modality such as willingness or intention. Forty one tokens were difficult to determine their meaning. In the learner data, however, *should* (117) and *must* (11) were used only in root contexts as a marker of obligation or necessity, while *may* and *might* were used only for epistemic markers. The low frequency of *must* conveying the strongest obligation in comparison with the weaker obligation modal *should* and the phrasal modal *have to* in the learner data may reflect the change in the native usage, in which a gradual increase of the latter two modals has been reported (Leech, 2003). The distribution of modal verbs used as epistemic markers will be discussed in the next section together with other lexical devices for epistemic modality.

2. Epistemic modal devices

A total of 1,027 tokens of lexical devices conveying epistemic modal meanings were found in the data set, of which more than half (52.6%) were adverbials with 540 tokens and 14 types. Lexical verbs were produced 284 times (27.7%) with five types, followed by epistemic modal verbs with 199 tokens (19.4%) and seven types. *Possible* (1 token) and *possibility* (3 tokens) were the only adjective and noun used as epistemic markers.

The number of types of epistemic devices employed by the learners was 28 out of the 75 target items, much fewer than 66 produced by the learners in Hyland and Milton (1997) or 49 in Oh (2007)[7], but similar to that in Litica (2009), in which the Croatian learners of English produced 31 types of epistemic devices in picture description tasks. The fact that Hyland and Milton and Oh analyzed written essays while Letica and the current study

[6] For example, Greenbaum & Quirk (1990) use the expression 'possibility' in describing the meanings of *can* and *may*. Sengmwun Comprehensive English, a popular grammar book in Korea, presents the meaning of *may* as 'possibility', adding that *can* is equivalent to *may* in this meaning (Song, 2005, p. 131).

[7] The number of target items in Oh is 110, of which 75 were common with the items analyzed in Hyland and Milton (1997). Of these 75, the learners used 49.

examined spoken data may explain the observed discrepancy in the range of epistemic devices. The number of tokens per 100 words was 1.86, a frequency similar to what was observed in Hyland and Milton's (1997) written interlanguage. However, this similarity of frequency between the spoken and written interlanguage diverges from what Holmes (1988, cited in Hyland and Milton, 1997) noted; that is, the epistemic devices were twice as frequent in conversation as in written corpora. Differently from the NS, the NNS with a limited L2 proficiency may have been hindered from using modal devices by the processing pressure of the spontaneous speech.

There was no single predominant modal verb conveying epistemic modality in the learner data; rather, such diverse modals as *will* (69), *be going to* (49), *can* (38), *would* (28), *could* (9), *might* (5), and *may* (1) were used to express epistemic modal meanings. The (mis)use of *can* as an epistemic marker was relatively frequent as was discussed in the previous section.

The grammatical categories of the epistemic devices preferred by the L2 learners in previous studies vary: modal verbs in Hyland and Milton (1997), lexical verbs in Oh (2007), lexical verbs and adverbials in Letica (2009), and adverbials in the current study. However, what is common among the various studies is a heavy dependence on a few items to express epistemic modality: for example, *think* (237), *maybe* (233), and *kind of* (108) in our study.

Table 4 presents the epistemic markers classified according to the grammatical and semantic categories presented in Hyland and Milton (1997).

TABLE 4

Grammatical and Semantic Categories of the Epistemic Markers

	Certainty	Probability	Possibility	Usuality	Approximation	Total
Adverbs	actually 73 definitely 10	probably 9 quite 8	maybe 233 possibly 4 likely 4	usually 36 always 31 often 6	kind of 108 almost 11 about 6 nearly 1	540 (52.6%)
Lexical verbs	think 237 know 26	believe 11 seem 1	guess 9			284 (27.7%)
Modal verbs	will 69 be going to 49	would 28	can 38 could 9 might 5 may 1			199 (19.4%)
Adjectives			possible 1			1
Nouns			possibility 3			3
Total	464(45.3%)	57(5.6%)	307(29.9%)	73(7.1%)	126(12.3%)	1027

A few things are worth noting in the above table. First, the learners' overreliance on epistemic devices conveying certainty (45.3%) is consistent with what was observed in previous research, supporting the view that L2 learners are not as skillful as native speakers in qualifying their opinions or attitudes, leading to assertive or even aggressive tone in delivering their opinions, presumably due to their limited L2 proficiency or transfer of L1 speech style (Hyland & Milton, 1997; Oh, 2007). Second, the lexical items employed by the learners for each of the five semantic categories were restricted to one or two, producing rather monotonous tone with no subtleties of meaning that might have been conveyed with diverse expressions. In the certainty category, for example, the lexical verb *think* was the most common, followed by the adverb *actually* and the modal *will*. *Maybe* (233 tokens) surpassed all the other items in the possibility category, a few examples of which are shown in (3). In contrast, there was only one instance of *may* with the meaning of epistemic possibility.

(3) a. *Maybe* they have some standard that we don`t know. (file 1)
b. As time goes by, *maybe* people notice because on the internet they could see whenever they want they could see the picture with their name. (file 4)
c. *Maybe* the guy has a problem; they need to try fix them. (file 7)

Another expression the learners favored to qualify their opinion or judgment was *kind of* in the approximation category.

(4) a. But as I mentioned that it's *kind of* maybe a way to so prevent this kind of corruption. (file 2)
b. But before leaving Korea, they, they got *kind of* warnings from the government and they left for Iraq after writing will. (file 5)
c. I think it's *kind of* natural happen natural happening like cause they are they are doing something wrong. (file 10)

Overall, the use of the epistemic markers appears to have been restricted by the learners' L2 proficiency and spoken mode of communication in that the inventory of lexical items exploited to convey the various degrees of epistemic commitment was not diverse.

3. Developmental change in the use of modal expressions

According to Coates' (1983) classification, the functions of the 11 modal verbs analyzed in this paper amount to 31 (see Appendix). A possible developmental change in the use of the modal verbs by the learners was examined by counting the functions and tokens of

modal verbs in each file (Table 5). A positive correlation was observed between the file number[8] and the number of tokens per 100 words (r=.712, p=.001), but not between the file number and the number of functions (r=.115, p=.721), or between the number of functions and tokens of the modal verbs per 100 words (r=.177, p=.583).

TABLE 5

Functions and Tokens of Modal Verbs

File	Functions	Tokens/100 words
1	13	1.53
2	14	1.69
3	14	1.41
4	16	2.02
5	12	1.93
6	20	1.58
7	15	2.37
8	14	2.16
9	13	1.63
10	17	2.78
11	13	2.04
12	15	2.87
average	14.5	1.99

This result suggests that the frequency of the modal verbs increased in proportion to the students' English proficiency, but not the functions of the modal verbs, meaning that the learners tended to use the modals with the same function repeatedly, adding few new functions to their inventory of modals.

The frequency and the types of the epistemic devices per file were also analyzed according to the grammatical categories, which is presented in Table 6. There was no correlation observed between the file number and the number of adverbials, lexical verbs, or modal verbs, as shown by the Pearson's r for each category in Table 6. This indicates that the use of these epistemic devices did not increase gradually during the seven month period. The total number of the whole epistemic devices also did not exhibit a gradual improvement (r=.180, p=.575). The types of the epistemic devices also did not increase; rather they decreased in the latter part of the file, as evidenced by the negative Pearson's r in the right most column of Table 6.

The number of epistemic devices per 100 words for each semantic category was calculated for each file. No correlation was observed between the file number and the number of devices in each of the categories (Table 7). The categories of probability and approximation showed negative Pearson's r though statistically not significant, implying

[8] Based on the assumption that the learners' English proficiency will improve in proportion to the time they invest in practicing English, we hypothesized that the file number correlated with the learners' L2 proficiency: the higher the file number is, the higher their L2 proficiency is.

that the devices in these two categories tended to decrease in number over the duration of seven months.

TABLE 6

Epistemic Modal Markers per File

File	Adverbs	Lexical verbs	Modal verbs	Total	Type
1	43(1.12)	19(0.49)	15(0.39)	78(2.02)	18
2	37(0.95)	21(0.54)	20(0.51)	79(2.02)	18
3	34(0.74)	33(0.71)	12(0.26)	79(1.71)	14
4	39(0.95)	21(0.51)	16(0.39)	76(1.85)	17
5	46(0.78)	22(0.38)	45(0.77)	113(1.93)	19
6	39(0.76)	28(0.55)	10(0.20)	77(1.50)	19
7	43(0.80)	8(0.15)	18(0.34)	69(1.29)	15
8	81(1.57)	28(0.54)	13(0.25)	124(2.41)	20
9	46(0.86)	26(0.49)	10(0.19)	82(1.54)	14
10	35(0.82)	34(0.79)	20(0.47)	89(2.08)	15
11	48(1.29)	29(0.78)	8(0.22)	85(2.29)	17
12	49(1.30)	15(0.40)	12(0.32)	76(2.02)	14
total	540(0.98)	284(0.52)	199(0.36)	1027(1.86)	
Pearson's r	.357	.121	-.329	.180	-.357
p	.255	.708	.296	.575	.254

TABLE 7

Semantic Categories per File

file	possibility	certainty	probability	usuality	approximation
1	0.70	0.70	0.13	0.13	0.36
2	0.61	1.00	0.13	0.05	0.23
3	0.41	0.84	0.04	0.11	0.30
4	0.51	0.71	0.29	0.12	0.22
5	0.31	1.11	0.15	0.05	0.31
6	0.47	0.76	0.04	0.10	0.14
7	0.54	0.43	0.15	0.00	0.17
8	0.99	0.78	0.04	0.33	0.27
9	0.24	0.68	0.02	0.47	0.13
10	0.56	1.29	0.05	0.05	0.14
11	0.54	1.24	0.19	0.03	0.30
12	0.98	0.69	0.05	0.08	0.21
Pearson's r	.234	.184	-.285	.124	-.440
p	.465	.568	.369	.702	.153

In sum, there was no developmental progress observed except for the frequency of modal verbs, leading to the tentative conclusion that seven months may be too short a time to make any visible change in the use of modal devices in spontaneous speech in L2.

V. SUMMARY AND CONCLUSION

This paper set out to analyze a spoken corpus of L2 learners with three focuses: 1) to examine the overall distribution pattern of English modal verbs, both root and epistemic, in spoken interlanguage of Korean learners, 2) to analyze devices used to express epistemic modal meanings, and 3) to explore possible developmental changes in the use of modal devices during the seven months of data collection period.

(Semi-)spontaneous discussions on various topics were recorded from four students participating in a self-organized English study group, producing a corpus of 55,072 words. The data analysis revealed that the frequency of modal verbs employed by the learners was a little lower than that found in the NS conversation data presented in Biber et al. (1999): 1.99 vs. 2.32 tokens per 100 words for the NNS and NS, respectively. The learners' use of modal verbs was skewed toward one modal *can*, which accounted for 43 % of the 1095 modals. L1 transfer and instruction effect were discussed as two main factors that might have resulted in the excessive use of *can*.

In addition to being used to mark root possibility and ability meanings, *can* was used in epistemic contexts, which is not fully accepted in native usage. Such an extension of *can* in epistemic contexts is considered a byproduct of formal instruction which does not pay careful attention to the necessary distinction between two types of possibility: root and epistemic.

Of the 1,027 epistemic modal devices utilized by the learners, adverbs were most frequent (540 tokens), followed by lexical verbs (284) and modal verbs (199). Previous research does not provide a unifying picture regarding the grammatical categories preferred by L2 learners, possibly due to diverse data collection methods, classification schemes, and language proficiency of L2 learners.

The learners' utterances were marked by epistemic devices denoting strong commitments to the truth of the propositions they produce (45.3% of 1,027 markers): modals *will* and *be going to*, and a lexical verb *think*. The relative high frequency of *should* and *have to*, modals of deontic obligation, also adds to the assertive and authoritative tone of L2 speech, confirming the view that L2 learners are not effective in qualifying their opinions or ideas, being restricted by their L2 proficiency. The learners' speech was also characterized as monotonous and repetitive due to a heavy reliance on a few items in each of the five semantic categories: *think* (237), *maybe* (233), and *kind of* (108).

No clear developmental progress was detected over the seven month period except for the gradual increase of the tokens of modal verbs. The lack of any visible development in the range of modal devices even after seven months' participation in a self-organized study group underlines the difficulty L2 learners experience in controlling the forms and meanings of various modal markers. As was noted by Hyland and Milton (1997), learners may need the assistance of explicit instruction to acquire the multiple functions and meanings of modal verbs and the full range of epistemic category.

The use of modal expressions is likely to be subjective and context sensitive, influenced by such factors as mode of communication, topics, speakers' L1 backgrounds and L2 proficiency, and socio-cultural backgrounds. Due to the multiple variables that need to be controlled, the results obtained in previous research as well as the current study were shown to diverge from each other in some respects, failing to give a complete picture on how learners acquire and use modal devices. In order to better understand how learners gain control over the complex system of modality, more longitudinal research is needed examining both written and spoken data and with a longer data collection period. Further, studies examining the effect of explicit instruction may help teachers in selecting an effective approach to modality in L2 classrooms.

REFERENCES

Aijmer, K. (2002). Modality in advanced Swedish learners' written interlanguage. In S. Granger, J. Hung & S. Petch-tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 55-76). Amsterdam: John Benjamins.

Ammann, A., & Van der Auwera, J. (2002). Korean modality: Asymmetries between possibility and necessity. In H-D. Ahn & N. Kim. (Eds), *Selected Papers from the Twelfth International Conference on Korean Linguistics* (pp. 101-115). Seoul: Kyengjin Publishing.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English.* London: Longman.

Bybee, J., & Fleischman, S. (1995). Modality in grammar and discourse: An introductory essay. In J. Bybee & S. Fleischman (Eds.), *Modality in grammar and discourse* (pp. 1-14). Amsterdam: John Benjamins.

Chen, H-I. (2010). Contrastive learner corpus analysis of epistemic modality and interlanguage pragmatic competence in L2 writing. *Arizona Working Papers in SLA & Teaching, 17*, 27-51.

Choi, J-E. (2011). *A corpus based analysis of modal verbs between high-school textbooks and English exam of college scholastic ability test (CAST).* Unpublished master's thesis. Korea University of Foreign Studies, Seoul.

Choi, S-J. (1995). The development of epistemic sentence-ending: Modal forms and functions in Koran children. In J. Bybee & S. Fleischman (Eds.), *Modality in grammar and discourse* (pp. 165-204). Amsterdam: John Benjamins.

Choi, S-J. (2006). Acquisition of modality. In W. Frawley, E. Eschenroeder, S. Mills & T. Nguyen (Eds.), *The expression of modality* (pp. 141-171). Berlin: De Gruyter Inc.

Coates, J. (1983). *The semantics of modal auxiliaries*. London: Croom Helm.

Coates, J. (1995). The expression of root and epistemic possibility in English. In J. Bybee & S. Fleischman (Eds.), *Modality in grammar and discourse* (pp. 55-66). Amsterdam: John Benjamins.

Collins, P. (2009). *Modals and quasi-modals in English*. Amsterdam: Rodopi.

Depraetere, I., & Reed, S. (2006). Mood and modality in English. In B. Aarts & A. McMahon (Eds.), *The handbook of English linguistics* (pp. 269-290). Oxford: Blackwell.

Dittmar, N., & Ahrenholz, B. (1995). The acquisition of modal expressions and related grammatical means by an Italian learner of German in the course of 3 years of longitudinal observation. In A. G. Ramat & G. C. Galèas (Eds.), *From pragmatics to syntax: Modality in second language acquisition* (pp. 197-232). Tübingen: Tübingen: Gunter Narr Verlag.

Greenbaum, S., & Quirk, R. (1990). *A student's grammar of the English language*. London: Longman.

Gwon, E-H. (2009). *Yeonge yangsang cotongsa can-kwa will-uy uymicek ssuim pwunsek mit ihay yeonkwu* (An analysis study of semantic uses and learners' comprehension on the modal auxiliaries 'can' and 'will'). Unpublished master's thesis. Korea National University of Education, Cheongju.

Hinkel, E. (2009). The effects of essay topics on modal verb uses in L1 and L2 academic writing. *Journal of Pragmatics, 41*, 667-683.

Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing, 6*(2), 183-205.

Karkkainen, E. (1992). Modality as a strategy in interaction: Epistemic modality in the language of native and non-native speakers of English. *Pragmatics and Language Learning, 3*, 197- 216.

Kim, J-E. (1998). *Wurimal yangtae yongen kwumwun yeonku* (Modal constructions in Korean). Seoul: Hankwuk Mwunhwasa.

Lee, E-P., & Park, S-W. (2008). A study of the epistemic modality in college writing. *English Language and Literature Teaching, 14*(1), 45-58.

Leech, G. (2003). Modality on the move: the English modal auxiliaries 1961-1992. In R. Facchinetti, M. Krug, F. Palmer (Eds.), *Modality in contemporary English* (pp. 223-240). Berlin: Mouton de Gruyter.

Letica, S. (2009). Use of epistemic modality by non-native speakers of English. In R. Lugossy, J. Horvath & M. Nikolov (Eds.), *UPRT 2008: Empirical Studies in English Applied Linguistics* (pp. 119-134). PecsL Lingua Franca Csoport.

Lim, D-H. (2003). On the modal system in modern Korean. *Korean Semantics*, *12*, 127-153.

Lyons, J. (1977). *Semantics 2*. Cambridge: Cambridge University Press.

McEnery, T., & Kifle, N. (2002). Epistemic modality in argumentative essays of second language writers. In T. Flowerdue (Ed.), *Academic discourse* (pp. 182-195). London: Longman.

Min, S-J. (2010). A corpus-based analysis of EFL learners' use of hedges in cross-cultural communication. *English Language and Literature Teaching*, *16*(4), 91-106.

Nuyts, J. (2000). *Epistemic modality, language, and conceptualization: A cognitive-pragmatic perspective*. Amsterdam: John Benjamins.

Nuyts, J. (2006). Modality: Overview and linguistic issues. In W. Frawly, E. Eschenroeder, S. Mills & T. Nguyen (Eds.), *The expression of modality* (pp. 1-26). Berlin: Mouton de Gruyter.

Oh, S-Y. (2007). A corpus-based study of epistemic modality in Korean college students' writings in English. *English Teaching*, *62*(2), 147-175.

Palmer, F. (1986). *Mood and modality*. Cambridge: Cambridge University Press.

Papafragou, A. (1998). The acquisition of modality: Implications for theories of semantic representation. *Mind and Language*, *13* (3), 370-399.

Park, H-S. (2010). Acquisition of modality by L2 learners of Korean. *Studies in Modern Grammar*, *60*, 233-261.

Ramat, A. G. (1999). Grammaticalization of modality in language acquisition. *Studies in Language*, *23*(2), 377-407.

Song, S-M. (2005). *Sengmwun Conghap Yeonge* (Sengmwun Comprehensive English). Seoul: Sengmwun Publishing.

Stephany, U. (1995). Function and form of modality in first and second language acquisition. In A. G. Ramat & G. C. Galèas (Eds.), *From pragmatics to syntax: Modality in second language acquisition* (pp. 105-120). Tübingen: Gunter Narr Verlag.

Van der Auwera, J., & Plungian, V. (1998). Modality's semantic map. *Linguistic Typology*, *2*, 79-124.

Vold, E. V. (2006). Epistemic modality markers in research articles: A cross-linguistic and cross- disciplinary study. *International Journal of Applied Linguistics*, *16*(1), 61-87.

Von Fintel, K. (2006). Modality and language. In D. Borchert (Ed.), *Encyclopedia of philosophy*. Detroit: MacMillan Reference USA. Retrieved from http://mit.edu/ fintel/ www/ modality.pdf.

APPENDIX

Classification of the Modal Verbs

	root	epistemic
can	possibility, ability, permission	
could	past of *can* hypothetical possibility, ability, permission	possibility
will	intention	prediction
would	past of *will* hypothetical intention	past of *will* hypothetical prediction
may	permission	possibility
might	past of *may* hypothetical permission	possibility past of *may* hypothetical possibility
must	obligation/necessity	logical necessity, confident inference
should	obligation(weak)	inference
have to	necessity	
need to	necessity	
be going to	intention	prediction

(adapted from Coates, 1983)

Examples in: English

Applicable Languages: English

Applicable Levels: Tertiary

Hyeson Park
Dept. of English Language & Literature
Keimyung University
1095 Dalgubeol Daero, Dalseo-gu
Daegu 704-701, Korea
Tel: 053-580-5091
Email: parkhy@kmu.ac.kr

Received in January, 2012
Reviewed in February, 2012
Revised version received in March, 2012