The use of epistemic markers as a means of hedging and boosting in the discourse of L1 and L2 speakers of Modern Greek: A corpus-based study in informal letter-writing

Lia Efstathiadi
economid@enl.auth.gr
School of English, Aristotle University of Thessaloniki

Abstract

The paper investigates the semantic area of Epistemic Modality in Modern Greek, by means of a corpus-based research. A comparative, quantitative study was performed between written corpora (informal letter-writing) of non-native informants with various language backgrounds and Greek native speakers. A number of epistemic markers were selected for further qualitative investigation on the grounds of their high frequency. The qualitative study revealed the ways epistemic markers (grammatical and lexical) are used in order to express the speaker's stance while they perform a number of discourse-pragmatic functions without violating the societal norms of politeness. The present study made use of the literature on Epistemic Modality, the face-management theory of politeness and the interpersonal metadiscoursal features known as hedges and boosters.

Introduction

The paper explores the semantic area of Epistemic Modality (EM, see Appendix 1 for list of abbreviations) by means of a corpus-based research. A comparative, quantitative study was performed between written corpora (informal letter-writing) of advanced non-native speakers (NNS) with various language backgrounds (see Appendix 2) and Greek native speakers (NS), the control group. The reader should have in mind that although this study concerns written corpora, the terms 'speaker-hearer' will be used throughout the paper in a broad sense to include the terms 'writer-addressee'. Furthermore, the speaker is assumed to bear the female identity.

A number of epistemic markers were selected for further qualitative investigation, on the grounds of their high frequency. Their contextualisation revealed the ways (grammatical and lexical) the two groups use epistemic markers to express their stance (Biber & Finegan, 1989) without violating the societal norms of politeness. The research conducted investigates the pragmatic functions these modal forms serve within the L2 discourse and evaluates their role as hedges, boosters, and face-protection devices. The present study focuses on EM (Coates, 1983; Nuyts, 2006; Palmer, 1986 & 2001; Perkins, 1983; Traugott, 2006), the face-management theory of politeness of Brown and Levinson (1987) and the metadiscoursal features known as *hedges* and *boosters* (Hyland, 1998).

The epistemic markers under investigation

The items of the study should satisfy three conditions. They should be: a) single-word markers that epistemically modify an utterance (grammatically or lexically), b) found in both corpora to facilitate quantitative and qualitative comparisons between them, c) relevant to the discussion of hedging, boosting and *face*.

Hence, the paper focuses on the following EM markers: a) the modal verbs *μπορεί/bori* (=may) and *πρέπει/prepi* (=must). An obvious exception was made for *prepi* which, although totally avoided by NS, was considered too prototypical a category to be left out of the study, b) the lexical verbs *γνωρίζω/ɣnorizo* (=I come to know), *θεωρώ/theoro* (=I presume), *νομίζω/nomizo* (=I think), *ξέρω/ksero* (=I know), *πιστεύω/pistevo* (=I believe), c) the modal adverbs *βέβαια/vevea* (=surely), *ίσως/isos* (=perhaps), *μάλλον/malon* (=rather, more), *σίγουρα/siɣura* (=certainly). Although the modal uses of *θα*, namely the epistemic *θα (θα+*E), *θα* followed by an imperfective past (*θα*+IMP) or a perfective non-past verb (*θα*+D, see Appendix 4) apparently violate the first condition, these will be investigated due to their direct association to EM, hedging and boosting respectively.

The research hypotheses

Despite the advanced level of their proficiency, the L2 informants are expected to:

- epistemically modalise their utterances to a lesser degree than NS,
- favour the use of lexical rather than grammatical exponents of EM, in order to be as transparent as possible to avoid miscomprehensions,
- show a preference towards hedging. Although earlier studies (Hyland, 2000; Hyland & Milton, 1997; Low, 1996) report a general trend towards boosting this

preference was attested on a different genre (academic writing). Given the fact that the study concerns informal letter-writing and the delicacy of the situations described in the letters (see Appendix 3), the L2 learners are expected to resort to hedges mostly.

Before going any further, it would be wise to very briefly define EM, introduce the topics of the letters as well as the notions of hedging, boosting and face and finally display the semantic properties of the selected items.

EM, hedges, boosters and face

The primary subject of this semantic field is the speaker's opinion or knowledge rather than fact, relevant to the truth-value of the expressed event (Lyons, 1977; Palmer, 1986; Philippaki-Warburton & Spyropoulos, 2006; Sweetser, 1990). EM is viewed within the framework set by modal logic, where epistemic necessity and possibility are two central notions in the speaker's reasoning, (Coates, 1983; Lyons, 1977; Palmer, 1986 & 2001).

The situations described in the letters automatically foreground the issue of speaker attitude. The topics relate to situations that are either delicate (request for money/donation) or controversial (gambling), which may potentially damage the participants' face (Brown & Levinson, 1987). Although the speaker and the hearer share a friendly relationship, both tasks are difficult and ask for the speaker's diplomatic moves in order to perform face-threatening acts (FTAs) without the danger of sounding inappropriately assertive or impolite.

The divisive nature of the topics foregrounds the use of *hedges* like *it is my belief, maybe, I think* that express uncertainty and *boosters* like *I know, I firmly believe, surely* that force the strength of one's arguments. Their use enables the speaker to keep a balanced attitude while performing FTAs such as requesting or advice-giving without violating the L1 norms of politeness.

Successful L2 writing, in the sense of being pragmatically appropriate, brings into the foreground the issue of cultural variation. It has already been mentioned that educational as well as societal differences do exist among the NNS of the study as they come from different L1 backgrounds. Cross-cultural rhetoric suggests that the rhetorical preferences of different languages and cultures tend to manifest themselves in the L2 writing (Hofstede, 1986; Hyland & Milton, 1997; Koutsantoni, 2005a, b). Very often the L2 learners violate the communicative norms of the L1 society, by being too direct and dogmatic or too tentative and even naïve. To avoid such cross-cultural communication problems, the L2 learners must explicitly be

taught the different linguistic conventions that express the same meaning (in this case EM) and the particular L2 rhetorical strategies and politeness norms.

To this end, the contribution of electronic corpora is invaluable. Computer Learner Corpus (CLC) research burgeoned as a discipline in the late 1980s to facilitate researchers in focusing upon the description of real language data. Unlike other scholars (Owen, 1996; Tan, 2005) who view corpora as inauthentic and normative bodies of text, Granger (2004), Granger and Tyson (1996) see them as being primarily descriptive rather than prescriptive and believe that their contribution to Second Language Acquisition (SLA) studies is invaluable.

On the one hand, NS oral or written corpora provide a valuable source of information for the L2 learner, as she becomes better acquainted with collocations, idiomatic expressions, chunks of language in the L1 through exposure to authentic texts or recordings from a number of different genres. On the other hand, learner corpora facilitate an in-depth comparative investigation in the language production of NNS. They help the language teacher gain an insight into the learner's interlanguage and thus locate areas of particular difficulty. They focus on the L2 performance and offer a means of evaluating the effect of variables such as the learners' age, sex, L1 background, task type, learning situation on the learner output.

The semantic properties of the selected grammatical and lexical exponents of Greek EM

The modal verbs prepi and bori

The epistemic *prepi* is used to express the speaker's strong conviction in the truth of what she says, based on prior knowledge and experience of the world (Coates, 1983; Kallergi, 2004). When the following lexical verb is marked as perfective past, then only the epistemic reading is possible (Mackridge, 1987; Palmer, 1986; Tsangalidis, 2004):

> *Ο Γιάννης πρέπει να έφυγε (γιατί άκουσα το αμάξι του να φεύγει)*
>
> *o Janis prepi na efije (jiati akusa to amaksi tu na fevji)*
>
> *John probably/must have left (because I heard his car leaving)*

Bori, on the other hand, may vary in meaning according to whether it is used personally or impersonally. Ambiguous or interchangeable meanings between the two readings arise when there is an overt subject and the 3rd person singular *bori*

agrees with the ensuing lexical verb: *Ο Γιάννης μπορεί να φύγει / o Janis bori na fiji.* Depending on the context, it is possible to assume that a) John is capable of leaving, b) John is granted permission to leave, c) John will probably leave. The contribution of the context is crucial in resolving the ambiguity.

As with *prepi*, a past tense verb in the subordinate clause forces an epistemic reading (Palmer, 1986), as in   *Μπορεί να έβρεξε / bori na evrekse /It may have rained. Bori* may also be found in the imperfective past and still receive an epistemic interpretation:

> *(θα) Μπορούσε να είχε καθυστερήσει το τρένο αλλά ευτυχώς ήρθε στην ώρα του*
>
> *(θα) boruse na ixe kaθisterisi to treno ala eftixos irθe stin ora tu*
>
> *The train could (might) have been delayed but fortunately it arrived on time*

<div align="right">(taken from Holton, Mackridge, & Philippaki-Warburton, 1997, p. 210)</div>

To sum up, the two verbs differ in terms of degrees of polarity: *bori* is situated on the 'weak' side of the epistemic scale denoting possibility, while *prepi* holds the 'strong' side of certainty.

The interplay of Tense-Aspect-Mood (TAM) in the expression of Greek EM

In the expression of the epistemic sense, the Greek verbal syntagm is organised around the modal particles *να/na*, *ας/as* and *θα* and the grammatical categories of TAM. This study, however, focuses only upon the verbal syntagms of *θα*.

Depending on the context, *θα* can be a marker of futurity or an epistemic marker. The alternative readings are basically determined by the tense and aspect of the ensuing verb: "with non-past forms, the futurity value is more in the foreground, whereas when *θα* combines with past or perfect verb forms it is the modality value which is prevalent" (Joseph & Philippaki-Warburton, 1987, p. 173).

An unambiguously epistemic interpretation is foregrounded when *θα* combines with the perfective past (*θα*+E), as in *Θα της μίλησε / θa tis milise / He must have talked to her* (Mackridge, 1987, p. 275)*.* The epistemic interpretation is also to be preferred with the future perfect [*θα*+E: *θα*+perfect]: *Θα έχει φύγει για να μην απαντάει στο τηλέφωνο / θa exi fiji jia na min apantai sto tilefono / He must have left, since he doesn't answer on the phone* (Kallergi, 2004, p. 20)*.* Furthermore, the epistemic present (*θα*+E: *θα* + [-perf] [-past]) can also mark an epistemic reading, serving an

inferential function, as in *θα είμαστε περίπου πενήντα άτομα εδώ μέσα / θα imaste peripu peninta atoma εδο mesa / We must be approximately fifty people in here* (Kallergi, pp. 19-20).

Taking into account the controversial nature of the topics, this paper claims that *θα* may interact with aspect and tense to facilitate the accomplishment of the speaker's goals, yielding interesting pragmatic effects each time that relate to the communicative strategies of hedging and boosting. In particular, it is claimed that the interaction of *θα*+D, apart from its 'pure' future time reference, may also enable the speaker to boost her arguments and convey absolute certainty as to the materialisation of an event. This claim is further supported by Joseph and Philippaki-Warburton's scale of certainty (1987, p. 184), from the strongest to the weakest: *πρέπει να βρεθούν prepi na vreθun / they must be found → **θα βρεθούν / θα vreθun / they will be found** → ίσως να βρεθούν / isos na vreθun / they may (possibly) be found → μπορεί να βρεθούν / bori na vreθun / they may (might) be found*

It is also claimed that the interaction of *θα*+IMP, apart from expressing conditionality and/or counterfactuality, may yield a pragmatic effect linked to hedging, as it may a) signal lack of commitment or metaphoric detachment from the reality of the situation described (Sakellariou, 2001), b) mark a polite way of 'dressing up' a directive (De Haan, 2006; Searle, 1979), c) allow the speaker to 'pass the message' along to the hearer that the realisation of an event is left upon her/his good will (Fleischman, 1995).

The lexical verbs and adverbs of the study

The lexical verbs of the study are mental state verbs found in the 1st person. *Ksero* and *ɣnorizo* are strong assertive predicators (Perkins, 1983); *pistevo* and *θeoro* are placed on the certainty end of the continuum, whereas *nomizo* is rather weak and expresses doubt (Politis, 2001). Their meaning is context-sensitive and may assume an emphatic or a weaker sense, depending on their position in the sentence (initial, medial, final), and the actual setting (Holmes, 1984; Politis, 2001).

Epistemic adverbs refer to the content of the proposition, and not to the event or the participant(s) within it, and assign a degree of likelihood this content is actual (Palmer, 2001; Swan, 1988, as cited in Kallergi, 2004). *Isos* can express possibility as well as uncertainty, for an event may or may not materialise after all (Kallergi, 2004). *Malon* is relatively opaque in its semantics, assuming meanings closer to confidence or serving a comparative function, thus entailing the sense of 'rather'. *Vevea* is

essentially emphatic in effect, although it can hedge the pragmatic force of the speaker's utterances (Kallergi, 2004). Finally, *siɣura* is a marker of considerable epistemic strength, conceptually closer to absolute confidence.

The methodology

The study employs a corpus-based approach. The data was drawn from two written corpora. The items under examination were retrieved with *Monoconc Pro 2.2*, a concordancer that provides raw frequencies of particular words and strings of words, including misspellings and other morphological variants.

The material

The material was collected with the permission of The Centre for the Greek Language (CGL), the academic institution responsible for the fostering and further promotion of MG within and outside Greece. The Division for the Support and Promotion of the Greek Language (DSPGL), exclusively organises, plans, and administers the examinations for the Certification of Attainment in Modern Greek (CAG), "which is the sole title of proficiency in Modern Greek that is valid worldwide" (visit http://www.greeklanguage.gr/eng/aims.html for further information on CGL's organisation and aims).

According to DSPGL's official website, the certificate "serves as proof of the successful candidate's level of attainment in Greek in the work-market". Level C "allows foreigners to register at a Greek institution of higher education" whereas level D "allows citizens of European Union member states to prove complete knowledge and fluent use of the Greek language and thus be employed in a Greek civil service position" (Retrieved on 6 February 2009, from http://www.greek-language.gr /greekLang/en/ certification/01.html).

Under this perspective, the L2 informants were selected on the basis of their advanced level of proficiency in MG. The NNS data was drawn from the exam papers of the candidates who succeeded in the 2003 CAG examinations. CAG requires that each candidate must pass all four language skills, i.e. speaking, listening, reading and writing. The candidates' written production consists of two pieces of letter-writing, one of which is usually more formal than the other. The object of this study concerns the informal letter, which ranges from 200 words (level C) to 300 words (level D).

The informants

The NNS corpus consists of the writings of 143 adults, advanced L2 learners of MG, all holders of CAG. 78 of them hold level C and another 65 are holders of level D. On the other hand, the 114 informants of the NS corpus, are mostly 2nd and 3rd year students of the School of English, Aristotle University of Thessaloniki. The NS corpus consists of their written production, i.e. a letter written as an in-class timed (30') assignment. In order to ensure absolute comparability of the data, the NS were randomly divided into two groups and each was given one of the two topics the NNS wrote on.

The data and the compilation of the corpora

The original letters were typed and further transferred into an electronic database. The original format of the letter, i.e. misspellings, bad orthography, grammatical errors, was kept intact. The corpora consist of the main body of the letter. The date and the initial greeting (i.e. *Dear X*) were excluded as they were provided by the examination booklet at all times with the candidates filling in only the name of the friend the letter was addressed to. The named signatures at the closing part were not included in the corpora because this would only make a difference to their word-capacity without any further contribution to the purposes of this study. Table 1 presents in detail the size of the corpora. It should be kept in mind that the term *level* relates to learner corpora whereas the corresponding *topic* concerns L1 corpora.

Table 1. The size of the two corpora

NS corpus	Tokens	Informants	NNS corpus	Tokens	Informants
Topic C	16.040	60	Level C	19.429	78
Topic D	16.918	54	Level D	21.762	65
Total	**32.958**	**114**	**Total**	**41.191**	**143**

Procedure

Although the data collected were analysed both quantitatively and qualitatively, due to the paper's limited scope only the quantitative analysis is extensively discussed here. The literature on Greek EM was thoroughly investigated to locate as many grammatical and lexical exponents as possible. Previous research findings on Greek EM (Clairis & Babiniotis, 1999; 2001; Iakovou, 1999; Kallergi, 2004; Politis, 2001;

Tsangalidis, 2009) offered a substantial help, for they readily provided the study with an array of linguistic devices that convey the epistemic meaning.

Initially, a number of EM markers were retrieved from the four sub-corpora (NS-C / NS-D; NNS-C / NNS-D) with the help of *Monoconc Pro 2.2.* concordancer. This software provides raw frequencies of particular words and strings of words (including misspellings and other morphological variants) and thus allows for their in-depth contextual analysis. The next time-consuming step was that of disambiguation between deontic (mainly) and EM markers. Finally, the EM markers that marked the highest frequencies in all corpora were selected for further investigation. Figure 1 shows the outcome of the search conducted in the NNS corpus (Level C) for *pistevo*.
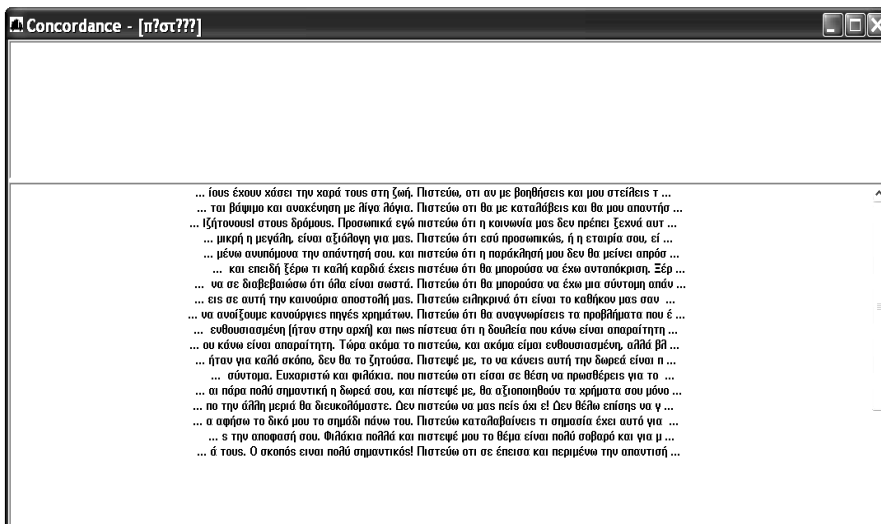

**Figure 1. The outcome of the search conducted
with Monoconc Pro 2.2. in NNS-C for *pistevo***

The markers were further grouped into five categories: modal verbs, lexical verbs, adverbs, hedges, and boosters. Multiple statistical analyses were performed, in particular the Pearson's chi-square test, for every single marker as well as for each of the five categories. The Fisher's Exact test was conducted only for those items whose normalised frequencies were less than five (Brace, Kemp, & Snelgar, 2000). Finally, a contextual analysis was performed to better explore the pragmatic function of the items under investigation.

The quantitative analysis

Table 2 presents the observed and normalised frequencies of the selected items. As the two corpora were not of equal size, their raw frequencies were normalised (per 10.000 words) to facilitate comparative analyses between the corpora. Figure 2 attempts a schematic illustration of their normalised frequencies.

Table 2 and Figure 2 reveal some interesting findings concerning the distribution of the selected epistemic markers in the corpora. First of all, most of the reasoning is primarily coded lexically with the frequent use of lexical verbs and adverbs, and secondarily with the use of the modal verbs *prepi* and *bori* and $\theta\alpha$+E. Secondly, an extensive use of the verbal syntagms of $\theta\alpha$ ($\theta\alpha$+IMP, $\theta\alpha$+D) is attested in the corpora.

Table 2. Raw and normalised frequencies of the selected items

Epistemic marker	NS	f/10.000 words	NNS	f/10.000 words
bori	65	38,64	77	35,98
prepi	0	0	6	2,85
θα+E	14	8,52	10	4,91
θα+IMP	129	78,76	99	48,85
θα+D	189	114,91	271	131,08
ɣnorizo	14	8,59	2	0,96
θeoro	13	7,87	5	2,34
nomizo	16	9,65	43	20,9
ksero	32	19,59	75	37,1
pistevo	29	17,55	30	14,5
vevea	19	11,35	38	18,38
isos	45	27,2	28	13,63
malon	6	3,6	10	4,75
siɣura	20	11,94	18	8,87
Total	**591**	**358,17**	**712**	**345,1**

Figure 3 groups the items into four categories: modal verbs (MODVBS), $\theta\alpha$ verbal constructions ($\theta\alpha$+verb), lexical verbs (LEXVBS), and adverbs.
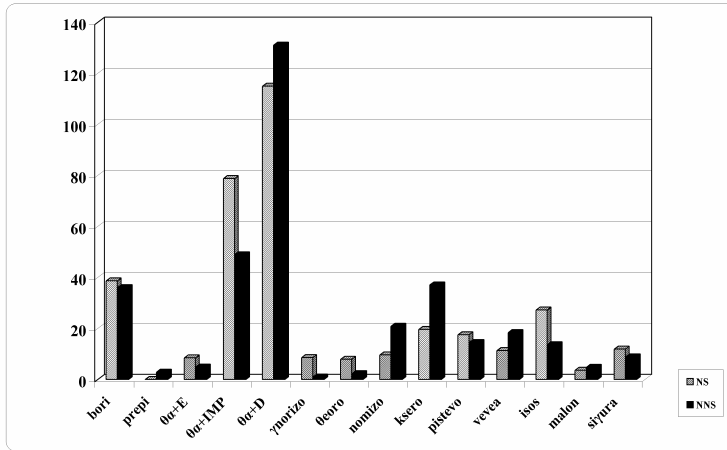
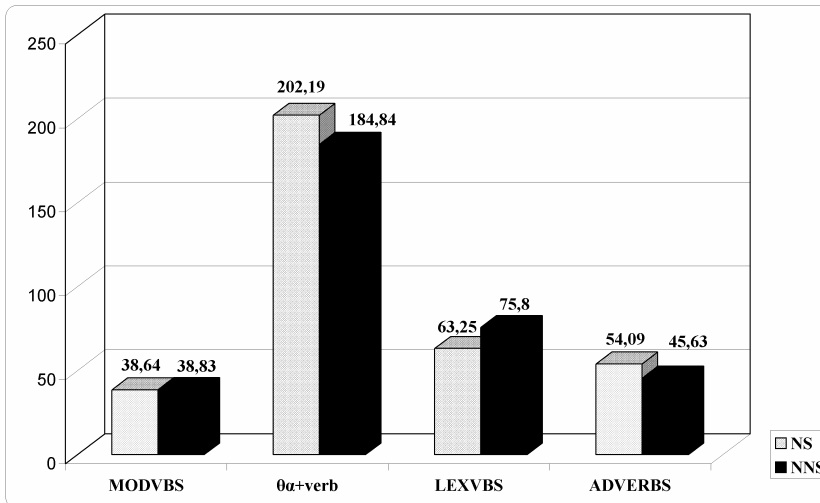**Figure 2. The normalised distribution of the markers of the study**

Figure 3. The distribution of the four categories of EM markers

Looking at the grammatical exponents of EM, one clearly sees a) a moderate use of modal verbs in both corpora, with the NNS corpus yielding slightly higher values (38,83), and b) a very high frequency of $\theta\alpha$ verbal syntagms, that receive their highest frequency counts in the NS corpus (202,19).

As for the lexical means of expressing EM, it is clear that NNS use lexical verbs more frequently (75,8) than NS (63,25), whereas the reverse picture holds for adverbs that are more commonly found in the NS corpus (54,09). Figure 4 depicts the normalised distribution of *bori* and *prepi*. It is evident that the epistemic necessity *prepi* marks a very infrequent use in the L2 corpus (2,85) and a total absence in the NS one, whereas the epistemic possibility *bori* shows a balanced distribution in both corpora, receiving its higher values in the NS corpus (38,64).
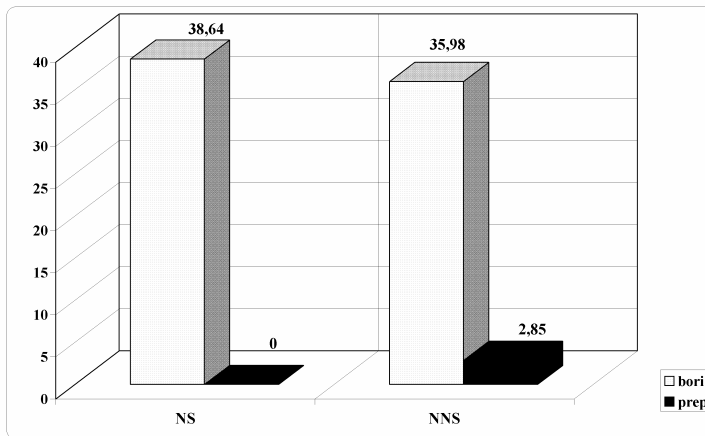

Figure 4. The distribution of bori and prepi

Figure 5 schematically presents the distribution of $\theta\alpha$ verbal syntagms. It is clear that $\theta\alpha$+E is not the preferred choice of the two groups; still, NS make a more frequent use of it (8,52) than NNS (4,91). Unlike $\theta\alpha$+E, both groups use more frequently $\theta\alpha$+D (NNS 131,08>NS 114,91). A statistically significant difference was found across the corpora, with NNS showing a strong tendency towards using $\theta\alpha$+D: x^2= 4.790, *DF* = 1, *p*=0,029. In addition, NS use $\theta\alpha$+IMP significantly more frequently than NNS (NS 78,76>NNS 48,85): x^2= 7.296, *DF* = 1, *p*=0,007.

As for the lexical exponents of EM, much fluctuation is attested in the use of some lexical verbs at the expense of others. Figure 6 illustrates their distribution. First of all, the L2 informants seem to rely mainly on *nomizo* (20,9)*, ksero* (37,1) and *pistevo* (14,5), whereas the picture in the NS corpus looks more balanced. Secondly, *ksero* is by far the first choice (56,69) of all the informants, with *pistevo* coming second (32,05) and *nomizo* third (30,55). Thirdly, both groups make a very infrequent use of $\theta eoro$ (NS 7,87>NNS 2,34) and $\gamma norizo$ (NS 8,59>NNS 0,96). A random search in the

Hellenic National Corpus (developed by the Institute for Language and Speech Processing) showed that these verbs are equally infrequent in the L1 use. Regarding the use of adverbs, Figure 7 illustrates their distribution in the corpora.
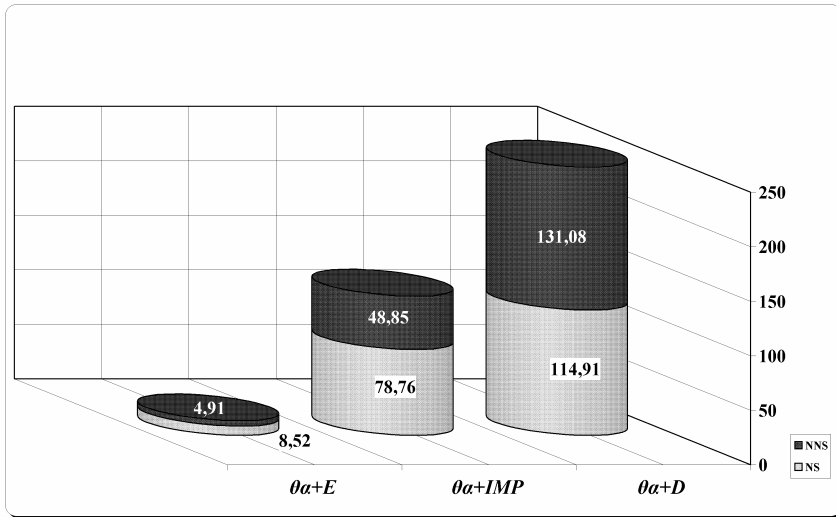

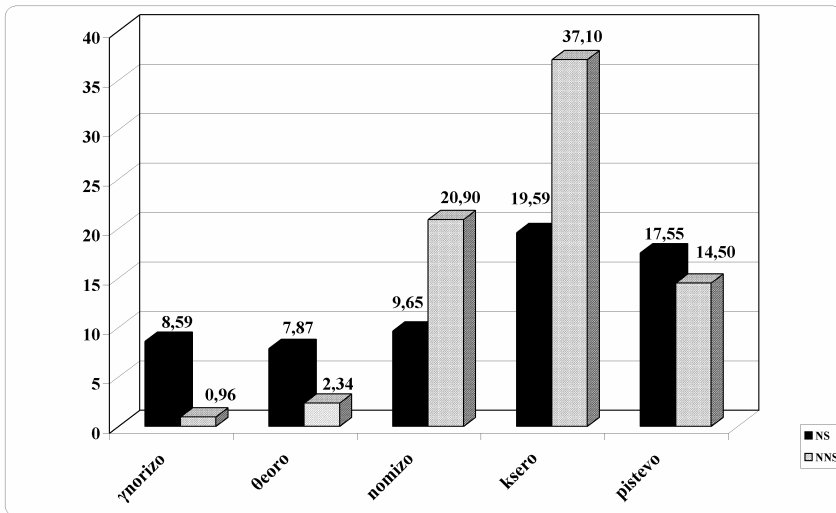**Figure 5. The distribution of θα verbal syntagms**

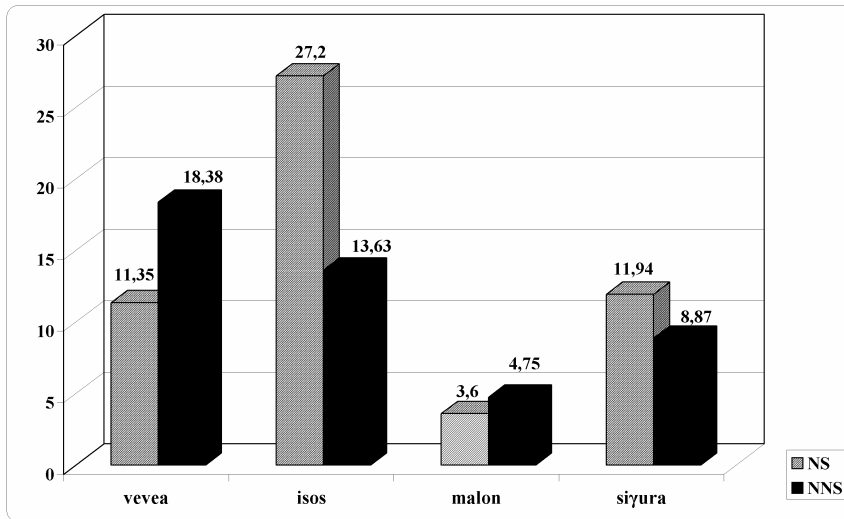**Figure 6. The distribution of the five lexical verbs**

Figure 7. The distribution of the four adverbs

Overall, *vevea* and *isos* are the most commonly used adverbs, whereas *malon* is the least frequent of all. The L2 informants show a preference for the certainty adverbs/boosters *vevea* and *siɣura* (27,25*)* over the possibility adverbs/hedges *isos* and *malon* (18,38). Conversely, NS use adverbial hedges more frequently (30,8) than adverbial boosters (23,29), which yields a statistically significant relationship between tendency towards the use of *isos* and NS corpora (NS>NNS): x^2= 3.883, *DF* = 1, *p*=0,049.

The classification of the selected markers as hedges and boosters respectively was based on earlier discussion in the paper relevant to the interaction of *θα* with tense and aspect and the brief introduction into the semantics of each marker. Tables 3 and 4 present the normalised frequencies of the epistemic hedges and boosters, whereas Figure 8 illustrates their distribution in the corpora.

It is clear that the totality of the informants opt for epistemic boosters. However, it is important to note that the distribution of boosters and hedges looks more balanced in the NS corpus as compared to the NNS one. A statistically significant relationship was found between tendency towards the use of the group of a) hedges and NS corpus (NS>NNS): x^2= 5.814, *DF* = 1, *p*=0,016, and b) boosters and NNS corpus (NNS>NS): x^2= 5.814, *DF* = 1, *p*=0,016.

Figure 8. The distribution of hedges and boosters

Table 3. The normalised frequencies of hedges

Hedges	NS	NNS
bori	38,64	35,98
θα+E	8,52	4,91
θα+IMP	78,76	48,85
nomizo	9,65	20,9
isos	27,2	13,63
malon	3,6	4,75
Total	**166,37**	**129,02**

Table 4. The normalised frequencies of boosters

Boosters	NS	NNS
prepi	0	2,85
θα+D	114,91	131,08
γnorizo	8,59	0,96
θeoro	7,87	2,34
ksero	19,59	37,1
pistevo	17,55	14,5
vevea	11,35	18,38
siγura	11,94	8,87
Total	**191,8**	**216,08**

Qualitative analysis

By looking at frequency lists in a corpus, some items become salient due to their high frequency. To detect their pragmatic function, one needs to go beyond the level of concordance lines to that of sentences or whole stretches of discourse. The brief (due to space limitations) qualitative analysis that follows was based on a plethora of examples retrieved from the corpora. It sheds more light on the ways NNS express their stance on a controversial issue as it charts the semantic nuances involved in the L2 expression of Greek EM (doubt/possibility, conviction/necessity) and evaluates the pragmatic function of the epistemic markers in question as hedges, boosters, and face-protection devices.

The modal verbs prepi and bori

The epistemic *prepi* is mostly used in its deontic sense. The limited use of *prepi* is based on the speaker's prior knowledge and life experiences, that license her in reaching logical conclusions and presenting claims with strong conviction. On the other hand, the semantic import of *bori* is weaker and associates itself with the speaker's highly subjective evaluation of an event as being possible or probable.

The modal uses of θa

Throughout the corpora the use of $\theta\alpha$+E is indeed limited. What is of particular interest, though, is the exhaustive use the L2 informants make of their linguistic resources. They borrow characteristics that belong to different modalities (epistemic and non-epistemic) such as semantic load, aspectual marking, and right use of tenses, to hedge or boost their arguments and achieve their communicative goals. Thus, prototypically deontic or dynamic verbs like *θέλω/θelo (=want) or* personal *μπορώ/boro* (=can) are frequently found in the imperfective past, to epistemically qualify an utterance, denoting modal remoteness rather than temporal reference (Iakovou, 1999).

The L2 informants (and for that matter the L1 group) use $\theta\alpha$+IMP to mitigate the illocutionary force of a directive to make a polite request, show displeasure or make an evaluative assessment of the hearer's practices. The hedged expressions enable them to put forward controversial statements with extreme caution without raising the hearer's opposition. Finally, the use of $\theta\alpha$+D serves a definite boosting effect as it conveys the L2 speaker's strong conviction.

The lexical verbs *ynorizo, θeoro, nomizo, ksero, pistevo*

There is adequate evidence in the data to suggest that the L2 speakers perceive the differences involved in the semantic import of the five verbs. They are aware of their positional variation in the sentence and the impact this has upon the semantics of the whole utterance. The sentence-initial use of these verbs conveys primarily the speaker's confidence, whereas a parenthetical use is mostly associated with matters of linguistic politeness (Coates, 2003; Holmes, 1984).

The adverbs *vevea, isos, malon, siyura*

The qualitative analysis corroborates other studies (Altenberg, 2006; Goutsos, 2007) as to the mobility of the adverbs, although both groups favour the initial and medial position to boost or hedge their arguments respectively.

NS are naturally expected to cope with such subtle manoeuvres within the epistemic modal meaning since they write in their L1. What is striking is the skillful production of NNS. The L2 learners create "epistemic clusters" (Hyland & Milton, 1997, p. 199) by combining together modal markers that belong to the same or to different degrees of EM, yielding different interpretations each time and reflecting a wide variety of pragmatic functions.

The hypotheses of the study revisited

Regarding the study's first hypothesis, the figures in Table 3 demonstrate that it is marginally valid. Although the L2 informants modalise epistemically their utterances to a lesser degree than NS, the respective values fall very close, due perhaps to the advanced level of the L2 learners. The study confirms the second hypothesis. Regarding the lexical exponents of the epistemic stance, clarity is the characteristic feature in the discourse of both groups, who choose to express their arguments explicitly to avoid miscomprehensions.

Finally, the corpus-based analysis disproves the third hypothesis. The L2 informants clearly prefer the use of boosters. This preference was also attested in the NS discourse, with a more balanced distribution, though, of the two features. The prevalence of boosters over hedges has been attested in other studies too (Hyland & Milton, 1997; Koutsantoni, 2005 a & b), and seems to hold cross-culturally as boosters are found to be more visible than hedges (Low, 1996).

The limitations of the study

No study is without limitations. The major limitation of this research is that it relies on a limited sample of informants. Although lack of generalisability is an inherent problem of small-scale studies, by no means should it serve to negate the investigative value of this study's findings. Future work on larger corpora needs to be conducted in order to reach firmer conclusions.

The findings of this study concern written resources only. A systematic investigation of the oral production (the interview part) of the L2 informants might have given rise to a variety of other markers of EM or to different patterns of their pragmatic use. In a similar vein, it would be interesting to examine the range of markers that would have arisen had this letter been a formal rather than an informal one.

Also, one should bear in mind that the L2 learners of MG are holders of the two more advanced levels of CAG. Equally interesting if not more, would be the investigation of the learners' performance who failed the exams to see the degree to which they express EM, the way they cope with politeness norms, the strategies they use in order to protect the participants' face.

Conclusion

The study yields a number of interesting findings, all relative to EM and the communicative strategies of hedging and boosting. There is enough evidence to support the gradient model of EM (Coates, 1983; Halliday, 1985; Nuyts, 2006; Perkins, 1983; Traugott, 2006) that ranges from certainty, via probability, to possibility.

The possibilities of combining together modal markers can be infinite, possibly because we tend to think in terms of degrees of likelihood and not in black-and-white terms (Leech & Svartvik, 2002). The selected epistemic markers display a remarkable ability to diffuse their meaning to other modal markers within the same sentence. This spread of meaning enables the L2 speakers to express subtle semantic nuances within the epistemic meaning or, what is more, convey their stance on divisive matters without deviating from societal norms of politeness.

The results corroborate previous research findings (Hyland & Milton, 1997) which demonstrate the pragmatic importance of EM markers as a discoursal resource for the negotiation of knowledge or claims and the marking of stance towards one's propositions and the hearer.

As for the expression of EM, the results of this study are aligned with research findings on the expression of modality (Dittmar & Ahrenholz, 1995; Giacalone Ramat, 1995; Stephany, 1995), which show that: a) modal verbs play an important role in the expression of deontic/dynamic modality, b) the epistemic modification of utterances is a later achievement in the process of both L1 and L2 acquisition and is usually expressed by lexical verbs of belief in the 1st person, and c) the lexical means that convey EM are preferred by NNS to the grammatical ones.

Overall, the L2 data show that NNS prefer to be direct and straightforward in the expression of EM. Their definite preference towards boosting is in alignment with previous research conducted in the Greek context (Hatzitheodorou & Mattheoudaki, 2006). Although this is the case, it does by no means imply that the presentation of the NNS' claims is not extremely cautious. They choose to be direct only when the context invites such a straightforward expression of attitude. The only difference is that they do so to a greater extent than NS. Otherwise, they are equally able to distinguish fact from personal opinion and modify their (quite damaging) remarks in a way that is likely to be both convincing and positively accepted by the hearer. The pragmatic reading is always polite and within the bounds of acceptable public behaviour.

The L2 data suggest that the expression of EM is just the means to an end. The use of epistemic markers is not related to matters of knowledge or lack of it. EM is the vehicle towards the achievement of the speaker's communicative needs. It is the linguistic realisation of negative as well as positive politeness strategies. Depending on the context, the speaker chooses to a) 'round off' the sharp edges in her discourse to minimise the negative impact of a highly sensitive material, or b) exploit the strength of the epistemic modal forms to project other aspects in her discourse.

The corpus-based analysis reveals the polyfunctionality of the epistemic markers which cover a wide range of meanings ranging from uncertainty to reassurance. It also demonstrates that hedging is misunderstood in the sense that it is not to be associated only with doubt, uncertainty or unassertiveness as there is nothing unassertive about choosing to talk about sensitive subjects or about sharing feelings and experiences (Coates, 2003). Linguistic politeness, in the sense of showing consideration to the feelings of others (Thomas, 1995) and conforming to the L1 standards of 'acceptable' public behaviour, is the key factor that draws together the semantic nuances and pragmatic functions of the epistemic markers.

Sharp distinctions within the epistemic meaning are not always possible. This, in its turn makes difficult the formation of explicit rules or definitions. Apart from the inherent difficulty involved in the epistemic meaning, part of the students' difficulty

is caused by the fact that the significance of the whole array of devices that realise it is largely ignored, underestimated or partly presented in both the teachers' and the students' textbooks of MG as an L2 (Spyropoulos & Tsangalidis, 2005).

To this end, the contribution of electronic corpora may prove to be invaluable. High frequency items in the L1 use, retrieved with concordancers like *Monoconc Pro 2.2.* or *Wordsmith Tools*, can be displayed along with the different degrees of the epistemic meaning they express. The application of electronic corpora can contribute to a better understanding of the semantic nuances involved in the expression of EM. Extensive exposure to concordances is expected to help learners realise that modal markers do not just operate in isolation and that it is the textual or social context that actually determines the challenging interplay between semantic usage and pragmatic function.

The findings of this study have obvious pedagogical implications that directly relate the design of the L2 teaching materials and the instruction of L2 grammar (syntax, idiomaticity, phraseology, etc.) to the applications of Computer Technology. A plethora of oral activities (exposure to oral corpora) or written exercises based on concordance lines (e.g. fill-in the gap, match form and meaning, look for positive/negative connotation) (O'Keeffe, McCarthy & Carter, 2007), can be designed to serve the learners' needs and level and facilitate their better understanding of the actual usage of the target language.

In conclusion, this study is better seen as providing some indication for further research. Much work remains to be done and huge amount of data to be collected, stored and further analysed with the help of computers and software like *Monoconc Pro 2.2*. This will hopefully shed more light on the ways the L2 learners of MG exploit the infinite conventions the language provides for the realisation of the epistemic meaning.

References

Anthony, L. (1997). ESP: What does it mean? ON CUE. http://interserver.miyazaki- Altenberg, B. (2006). The function of adverbial connectors in second initial position in English and Swedish. In K. Aijmer, & A.-M. Simon-Vandenbergen (Eds.), *Pragmatic markers in contrast, Studies in Pragmatics 2* (pp. 11-37). Amsterdam, The Netherlands: Elsevier.

Biber, D., & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text,* 9(1), 93-124.

Brace, N., Kemp, R., & Snelgar, R. (2000). *SPSS for psychologists. A guide to data analysis using SPSS for Windows* (Versions 8, 9 & 10)*.* London, United Kingdom: Macmillan Press Limited.

Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage.* Cambridge: Cambridge University Press.

Clairis, C., & Babiniotis, G. (1999). *Grammatiki tis neas ellinikis: Domolitourgiki-epikinoniaki,* (in cooperation with A. Mozer, A. Bakakou-Orfanou, S. Skopetea) Tomos II: To rima: I organosi tou minimatos [*Modern Greek Grammar: Structural-Functional-Communicative*, Vol. II: The verb: The organisation of the message]. Athens, Greece: Ellinika Grammata Publications.

Clairis, C., & Babiniotis, G. (2001). *Grammatiki tis neas ellinikis: Domolitourgiki-epikinoniaki,* (in cooperation with A. Mozer, A. Bakakou-Orfanou, S. Skopetea) Tomos G' (II. 2): Ta epirimatika stixia: I eksidikefsi tou minimatos [*Modern Greek Grammar: Structural-Functional-Communicative,* Vol. C (II. 2): The adverbial elements: The specialisation of the message]. Athens, Greece: Ellinika Grammata Publications.

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

Coates, J. (2003). The role of epistemic modality in women's talk. In R. Facchinetti, M. Krug & F. Palmer (Eds.), *Modality in Contemporary English*: *Topics in English Linguistics* (pp. 331-348). Berlin, Germany: Mouton de Gruyter.

De Haan, F. (2006). Typological approaches to modality. In W. Frawley (Ed.), *The expression of modality* (pp. 27-69). Berlin, Germany: Mouton de Gruyter.

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. Giacalone Ramat & G. Crocco Galèas (Eds.), *From Pragmatics to Syntax: Modality in Second Language Acquisition* (pp. 197-232). Gunter Narr Verlag Tűbingen.

Fleischman, S. (1995). Imperfective and irrealis. In J. L. Bybee & S. Fleischman (Eds.), *Modality in grammar and discourse* (pp. 519-551). Amsterdam, The Netherlands: John Benjamins.

Giacalone Ramat, A. (1995). Function and form of modality in learner Italian. In A. Giacalone Ramat & G. Crocco Galèas (Eds.), *From Pragmatics to Syntax: Modality in Second Language Acquisition* (pp. 269-293). Gunter Narr Verlag Tűbingen.

Goutsos, D. (2007). Μόρια, δείκτες λόγου και κειμενικά επιρρήματα: Η οριοθέτηση των γλωσσικών κατηγοριών με τη χρήση Ηλεκτρονικών Σωμάτων Κειμένων. [Moria, diktes logou ke kimenika epirimata: I oriothetisi ton glossikon katigorion me ti hrisi ilektronikon somaton kimenon]. *Electronic proceedings of the 8th International Conference on Greek Linguistics,* 754-768. Retrieved October 9, 2010, from http://www.linguist-uoi.gr/cd_web/docs/greek/ 012_goutsosICGL8_OK.pdf

Granger, S. (2004). Computer learner corpus research: Current status and future prospects. In U. Connor & T. A. Upton (Eds.), *Applied Corpus Linguistics: A multidimensional perspective* (pp. 123-145). Amsterdam, The Netherlands: Rodopi.

Granger, S., & Tyson, S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes,* 15(1), 17-27.

Halliday, M. A. K. (1985). *An Introduction to Functional Grammar.* London, United Kingdom: Edward Arnold.

Hatzitheodorou, A. M., & Mattheoudakis, M. (2006). *The Greek corpus of advanced learner English (GRICLE): An electronic database of written discourse.* Paper presented at the 30th International Conference on Functional Linguistics, Cyprus.

Hofstede, G. (1986). Cultural differences in teaching and learning. *International Journal of Intercultural Relations,* 10, 301-320.

Holmes, J. (1984). Modifying illocutionary force. *Journal of Pragmatics,* 8, 345-365.

Holton, D., Mackridge, P., & Philippaki-Warburton, I. (1997). *Greek: A comprehensive grammar of the modern language.* London, United Kingdom: Routledge.

Hyland, K. (1998). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics,* 30, 437-455.

Hyland, K. (2000). Hedges, boosters and lexical invisibility: Noticing modifiers in academic texts. *Language Awareness,* 9(4), 179-197.

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

Iakovou, M. I. (1999). *Tropikes katigories sto rimatiko sistima tis neas ellinikis* [*Modal categories in the verbal system of Modern Greek*]. Unpublished doctoral dissertation, National and Kapodestrian University of Athens, Greece.

Joseph, B. D., & Philippaki-Warburton, I. (1987). *Modern Greek.* London, United Kingdom: Routledge.

Kallergi, H. (2004). *Greek modality's semantic map: Particles, auxiliaries and adverbs.* Unpublished master's thesis, University of Amsterdam, The Netherlands.

Koutsantoni, D. (2005a). Certainty across cultures: A comparison of the degree of certainty expressed by Greek and English speaking scientific authors. *Intercultural Pragmatics,* 2(2), 121-149.

Koutsantoni, D. (2005b). Greek cultural characteristics and academic writing. *Journal of Modern Greek Studies,* 23, 97-138.

Leech, G., & Svartvik, J. (2002). *A communicative grammar of English*. London, United Kingdom: Pearson Education Limited (3rd ed.).

Low, G. (1996). Intensifiers and hedges in questionnaire items and the Lexical Invisibility Hypothesis. *Applied Linguistics,* 17(1), 1-37.

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

Mackridge, P. (1987). *The Modern Greek language.* London, United Kingdom: Oxford University Press.

Nuyts, J. (2006). Modality: Overview and linguistic issues. In W. Frawley (Ed.), *The expression of modality* (pp. 1-26). Berlin, Germany: Mouton de Gruyter.

O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom. Language use and language teaching.* Cambridge, United Kingdom: Cambridge University Press.

Owen, C. (1996). Do concordances require to be consulted? *ELT Journal,* 50(3)*,* 219-224.

Palmer, F. R. (1986). *Mood and Modality*. Cambridge, United Kingdom: Cambridge University Press.

Palmer, F. R. (2001). *Mood and Modality*. Cambridge, United Kingdom: Cambridge University Press (2nd ed.).

Perkins, M. R. (1983). *Modal expressions in English*. London, United Kingdom: Frances Pinter.

Philippaki-Warburton, I., & Spyropoulos, V. (2006). To englitiko sistima tis Ellinikis: Singritiki theorisi me tin Tourkiki ke didaktikes protasis. [The mood system of Greek: A comparison with Turkish and teaching suggestions]. In S. Moschonas (Ed.), *Syntax in the Learning and Teaching of Greek as a foreign language.*

Politis, P. (2001). Doksastika rimata ke epistimiki anageotita. [Doxastic verbs and epistemic necessity]. *Proceedings of the 4th International Conference of Greek Linguistics* (pp. 277-284). Thessaloniki, Greece: University Studio Press.

Sakellariou, A. (2001). I ipothetiki logi tis neas ellinikis [Conditionals in Modern Greek]. *Proceedings of the 4th International Conference of Greek Linguistics* (pp. 323-330). Thessaloniki, Greece: University Studio Press.

Searle, J. R. (1979). *Expression and meaning: Studies in the theory of speech acts.* Cambridge, United Kingdom: Cambridge University Press.

Spyropoulos, V., & Tsangalidis, A. (2005). I grammatiki sti didaskalia. Sygrisi enxeiridion gia ti didaskalia tis ellinikis os ksenis glossas. [The teaching of grammar. Comparison of textbooks for the teaching of Greek as a foreign language]. Series editor: Sp. Moschonas. Athens, Greece: Patakis Publications.

Stephany, U. (1995). Function and form of modality in first and second language acquisition. In A. Giacalone Ramat & G. Crocco Galèas (Eds.), *From Pragmatics to Syntax: Modality in Second Language Acquisition* (pp. 105-120). Gunter Narr Verlag Tűbingen.

Swan, T. (1988). *Sentence adverbials in English: A synchronic and diachronic investigation.* Oslo, Norway: Novus.

Sweetser, E. E. (1990). From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure. *Cambridge Studies in Linguistics 54.* Cambridge, United Kingdom: Cambridge University Press.

Tan, M. (2005). Authentic language or language errors? Lessons from a learner corpus. *ELT Journal, 59*(2), 126-134.

Thomas, J. (1995). *Meaning in interaction: An introduction to pragmatics.* London, United Kingdom: Longman.

Traugott, E. (2006). Historical aspects of modality. In W. Frawley (Ed.), *The expression of modality* (pp. 107-139). Berlin, Germany: Mouton de Gruyter.

Tsangalidis, A. (2002). Homonymy, polysemy, category membership: The case of Greek modal particles. *Belgian Journal of Linguistics: Particles, 16,* 135-150.

Tsangalidis, A. (2004). Disambiguating modals in English and Greek. In R. Facchinetti & F. Palmer (Eds.), *English modality in perspective: Genre Analysis and contrastive studies* (pp. 231-268). Frankfurt am Main, Germany: Peter Lang.

Tsangalidis, A. (2009). Modals in Greek. In B. Hansen & F. De Haan (Eds.), *Modals in the languages of Europe: A reference work* (pp. 139-163). Berlin, Germany: Mouton de Gruyter.

Appendix 1. Glossary of abbreviations and acronyms

Abbreviation Acronym	Meaning
CAG	Certificate of Attainment in Greek
CGL	Centre for the Greek Language
CLC	Computer Learner Corpus
DSPGL	Division for the Support and Promotion of the Greek Language
EM	Epistemic modality
FTA(s)	Face-threatening act(s)
L1	Language 1 = mother tongue
L2	Language 2 = foreign language
LEXVB(s)	Lexical Verb(s)
MG	Modern Greek
MODVB(s)	Modal Verb(s)
NS-C	Native speakers (Topic C)
NS-D	Native speakers (Topic D)
NNS-C	Non-native speakers (Level C)
NNS-D	Non-native speakers (Level D)
SLA	Second Language Acquisition
TAM	Tense-Aspect-Mood

Appendix 2. The L1 backgrounds of the NNS of the study

L1 background	N	L1 background	N
Albanian	6	German	18
Arabic	3	Hungarian	1
Bulgarian	16	Italian	7
Catalan	1	Polish	4
Czech	5	Portuguese	1
Dutch	9	Romanian	5
English	7	Russian	15
Estonian	1	Serbian	6
Finnish	2	Slovak	1
French	11	Spanish	22
Georgian	1	Swedish	1
Total			**143**

Appendix 3. The topics of the two letters

Topic C / Level C: Write a letter to a rich friend of yours, who happens to be an executive director in a big firm. Your purpose is to ask for a donation for the homeless shelter you are a volunteer at. Stress the importance of such a financial contribution and explain how the money will be invested. Use any argumentation or additional information you like in order to sound more convincing (200 words).

Topic D / Devel D: It is your firm belief that gambling on a systematic basis, especially during adolescence, can be the cause of a number of financial, psychological and social problems. When it comes to your knowledge that a close friend of yours, who lives and works in another city, gambles regularly, you decide to send him/her a letter, where you argue against it and you stress the reasons why (s)he should stop doing so (300 words).

Appendix 4. The observed frequencies of epistemic markers in the corpora

Marker	NS	NNS
bori*	65	77
prepi	0	6
θα+E**	14	10
θα+IMP***	129	99
θα+D****	189	271
γnorizo	14	2
θeoro	13	5
nomizo	16	43
ksero	32	75
pistevo	29	30
ipoθeto	0	2
fantazome	3	4
ime vevei/os	1	1
ime siγuri/os	12	16
ime pepismeni/os	1	1
ine aδinato	3	2
ine apoδeδiγmeno	2	1
ine veveo	1	1
ine γeγonos	2	1
ine δinato	1	0
ine piθano	7	1
ine siγuro	1	0
araje	0	1
vevea	19	38
enδehomenos	3	1
isos	45	28
malon	6	10
oposδipote	1	10
piθanos/a/otata	5	4
praγmati	2	3
siγura	20	18
Total	**636**	**761**

*   The items in bold are the modal markers under investigation, selected on the grounds of their higher frequency in the corpora.

**   θα+E (epistemic θα):   θα + [-perf] [-past]: θα γράφει/θa γrafi (epistemic present)

θα + [+perf] [+past]: θα έγραψε/θa eγrapse (epistemic past)

θα + perfect: θα έχει γράψει/θa ehi γrapsi (epistemic perfect)

θα + pluperfect: θα είχε γράψει/θa ihe γrapsi (epist. remote past)

***   θα+IMP: θα + [-perf] [+past]: θα έγραφε/θa eγrafe

**** θα+D: θα + [+perf] [-past]: θα γράψει/θa γrapsi

(see Tsangalidis, 2002, pp. 138-139)