

Liljana Mitkovska liljana.mitkovska@fon.edu.mk FON University, Skopje

Eleni Bužarovska eleni.buzarovska@mt.net.mk University of Ss Cyril and Methodius, Skopje

Marija Kusevska kusevska2003@yahoo.com, University of Goce Delčev, Štip

**Corpus evidence for the acquisition of modal verbs of obligation
by Macedonian Learners of English**

Abstract

This paper examines the acquisition process of the English modal verbs for expressing obligation (*must, have to, should* and *need to*) by Macedonian learners across several proficiency levels (A1, A2, B1 and B2) using data from the Macedonian English Learner Corpus. We examine how they are used in the writings of learners at each level from formal, semantic and pragmatic perspective. The frequencies will show the order of occurrence of the uses of these modal verbs and the comparison with data from native speaker corpora will determine how close or different the use of Macedonian learners is. Error types at each level and L1 influence are also identified and explained and suggestions for dealing with them in the classroom are discussed.

Key words: learner corpora, interlanguage, modal verbs, epistemic and deontic modality, language acquisition

Introduction

This paper presents the research on the use of English modal verbs of obligation by Macedonian learners of English at level A1, A2, B1 and B2 regarding their frequency of occurrence in the data base of the Macedonian English Learner Corpus (MELC). Learner corpora have become very common in the last 10-15 years. One of the first and best known is the *International Corpus of Learner English* ICLE, based in Louvain, which contains argumentative essays written by higher intermediate to advanced learners of Eng from several mother tongue backgrounds (Bulgarian, Chinese, Czech, Dutch, Finnish, French, German, Italian, Japanese, Norwegian, Polish, Russian, Spanish, Swedish, Tswana, Turkish).¹

The CECL coordinator Sylviane Granger and her team have been developing the methodology of compiling learner corpora since the 1980s and still lead in this field. They have also developed the corpus-based contrastive interlanguage analysis. This research methodology involves studies of the learner interlanguage as a system in itself. It is compared both with the TL and with the NL. Where there are corpora available, such comparisons reveal levels of overuse and underuse of certain structures or lexical items and provide evidence for possible NL influence. It has also been pointed out that findings from one IL should be compared with findings from other ILs. If a feature is common to ILs of learners with different NLs then it can be due to some developmental property. That is why using various learner corpora may be useful.²

Methodology

The present research was conducted on data from the Macedonian English Learner Corpus (MELC) which is an electronic data base of written material collected from learners of Eng in Macedonia at various ages and proficiency levels. We would like to thank all those who took part in its creation: teachers at about 20 public primary and secondary schools as well as language schools around Macedonia, faculty members, graduate and undergraduate students at FON University, the Faculty of Philology “Blaže Koneski” and the Faculty of Philology at the University “Goce Delčev”, Shtip. The corpus now contains around 500 000 words written by around 2000 participants. It is designed to be continually fed with data and we hope that it will grow.³ This will give more opportunities for further research.

MELC contains data from learners at different levels of proficiency, because we need to see how the language of the learners develops from beginners through intermediate to advance levels. In our corpus you can generate a sub-corpus according to your needs by choosing parameters of

¹ The Centre for English Corpus Linguistics (CECL) <http://www.uclouvain.be/en-cecl.html> (directed by Sylviane Granger) specializes in collection and use of corpora for linguistic and pedagogical purposes.

² See the CECL's database Learner corpora around the world <http://www.uclouvain.be/en-cecl-lcworld.html>.

³ The corpus is not freely available at the web, but access could be allowed for research purposes.

level, age, mother tongue and task.⁴ The uploaded corpus has to be analyzed manually. There are various tools you can use for automatic and half automatic analysis. We used AntConc for extracting the needed research items, but had to go through them for further analysis.

The research presented in this paper is an example of a corpus-based contrastive interlanguage study of the use of the modal verbs of obligation by learners of Eng with L1 Mac. There are many studies on the acquisition of the Eng modal verbs by speakers of various languages, but they mainly refer to intermediate and advanced learners (see for e.g. Aijmer, 2002 and McDouall, 2012). Our aim was to examine how the use of these modals evolves from beginners through intermediate and upper intermediate level. Before presenting the corpus results we give a short overview of the modal verbs of obligation in Eng and in Mac.

Modal verbs of obligation in English and in Macedonian

Both Eng and Mac have a set of modal verbs used to mark the semantic concept of modality, which can be defined as some kind of speaker's attitude or opinion towards the situation or state of affairs described. Two basic types of modality are usually distinguished: deontic (or root) and epistemic modality. Deontic modality is "concerned with action by others and the speaker himself" (Palmer, 1998: 96) and comprises meanings that refer to real world obligation, permission and ability (*The children must be in bed by nine o'clock.*). Yule (1998: 88) equates deontic with root modality: such modals indicate 'what is socially determined'. Epistemic modality refers to "the degree of commitment by the speaker to what he says" (Palmer, 1998: 51), i.e. the speaker's assessment of the truth of the situation expressed in the main predicate. Several degrees of epistemicity are distinguished: certainty, probability or possibility (*It's past nine, the children must/should/might be in bed now*). Cross-linguistically the two types of modality are usually expressed with the same modal verbs. The deontic meaning is ontologically primary so the epistemic, which concerns the internal (personal) domain, is derived from it. Sweetser (1990) argues that this extension is due to a metaphorical transfer. It has also been shown that in child language acquisition deontic meanings develop before epistemic ones. SLA research has shown that "the use of modal auxiliaries in epistemically modalized utterances is a late achievement" (Stephany 1995: 112). Thus we initially hypothesized that at the beginner and intermediate levels modal verbs of obligation will be used predominantly in their deontic functions.

In English obligation and necessity (or requirement) is expressed by the central modal auxiliaries *must*, *shall*, *should*, *ought to* and the semi-modals *have to* and *need (to)*. Since *shall* is quite rare in this function and occurs only in formal styles, we decided not to include it in our

⁴ Task 1 contains interactive correspondence, Task 2 are mainly narrative and descriptive texts and Task 3 are guided dialogues.

research. The semi modal *ought to* was left out for the same reason. The four investigated modals differ in the way they present the obligation, but their use also depends on region and style.

The deontic *must* is used primarily in directives where speaker exerts authority over the subject participant in 2nd person. In third person *must* may express circumstantial obligation, but the speaker's opinion is strongly felt. Referring to first person (I/we must) *must* indicates self-obligation (Leech, 1989: 77). In such use “*have to* is better for expressing habitual activities and *must* for activities that seem urgent or important to the speaker” (Thomson & Martinet, 1986: 141). In its epistemic use *must* expresses strong deduction, which is rather common, especially in spoken register (Biber et al., 1999: 494). The semi modal *have to* similarly expresses requirement and necessity, but the authority of the speaker is not involved. While *must* is rather subjective as it implies what the speaker thinks, *have to* is more objective (Leech, 1989: 83).

Should is used “to express the subject's obligation or duty ... or to indicate correct or sensible action” (Thomson & Martinet, 1986: 138). No authority is imposed, but it is a matter of consciousness and good sense and there is no expectation that the obligation will be fulfilled, as with *must*. As such, *should* is often used for advice and suggestion, which is neutral, unlike the emphatic *must* in which the speaker is emotionally involved. This modal also has an epistemic meaning of logical necessity but its occurrence is rare in all registers.

According to Leech (1989: 101) *need to* is half way between *must* and *should*: it asserts obligation and necessity, but without the certainty of *must*. However, it is different from both modals in the sense that it expresses ‘internal compulsion’. Sweetser (1990: 54) also states that “*need* implies that the obligation is imposed by something internal in the doer.”

English modals and semi-modals are not equally distributed in different registers. Most of them are more common in spoken than in written discourse. According to Biber et al. (1999: 487-490), *must* and *should* are relatively common in academic prose, where *have to*, being considered more colloquial, is quite rare. In American Eng the use of *must* has considerably declined since 1990s (according to the Corpus of Contemporary American by Mark Davis “COCA”: from 433 per million to 285 PM), while *need to* shows a significant rise (COCA: from 180 PM to 280 PM).

In Mac there are two modal verbs for expressing obligation: *mora* and *treba* and both can express both deontic (root) and epistemic modality. The meaning and use of these modal verbs have not been thoroughly described in Macedonian grammar books. The best account is given in Čašule (1989). Relying on K. Koneski (1977: 112) he states that *mora* expresses imposition usually from some external circumstances often against the will of the subject participant. The source of obligation can also be the speaker and the reason for the requirement is usually implied. The modal *treba* expresses similar deontic necessity, but it is often based on some internal awareness of the subject rather than external factors such as the speaker or social circumstances. On the scale of

intensity, *mora* is more categorical than *treba*. An additional means for expressing obligation in Mac is the periphrastic verb *ima da* which has an inflected and uninflected variant, but its functions do not have close equivalents in Eng.⁵

The above brief description suggests that there is no full correspondence between Eng and Mac modals for expressing obligation and necessity. Learners establish equivalents on the basis of their perception of closeness of functions and will rely on the distinctions that are familiar to them from their NL. The distinction between obligation from the speaker as opposed to circumstantial obligation is irrelevant for Macedonian speakers. Thus we can assume that they will not understand fully the distinction between *must* and *have to*. It is often subjective and even the context may not indicate the difference, but for native speakers there is clear distinction. The distinction between the degree of “categoricity” of the obligation (more or less categorical) is the most important parameter of differentiation between *mora* and *treba*. Applying this, they equate *must* with *mora* and *have to* with *treba*. Having this in mind, we assume that Mac learners will overuse *must*, employing it for strong obligation and necessity more often than native speakers. They are not aware of its more subjective connotation as well as the tendency to be avoided in spoken language, especially in American Eng. The modal *should* corresponds more closely to the meanings and uses of *treba* which, being weaker, implies that something is appropriate and socially acceptable to be done, but imposes no compulsion and we expect that Macedonian learners will acquire the central uses of this verb for moral obligation, advice and suggestion easily. The uses of *need* also correspond to *treba* where it points out the internal needs of the subject, but this equation may result in confusing *need* with *have to* and its use for expressing strong external obligation. We will test these hypotheses in the results of our analysis of the corpus examples.

Distribution of the modal verbs of obligation in the corpus

Table 1 shows the occurrences of the investigated modal verbs at each level in number and in percentage from the number of words in each sub-corpus. Results from the children’s corpus (age 5-15) and the adults’ corpus (age 16-60) are not given separately. For comparison, the frequencies of the same verbs in native speaker corpora are presented. COCA contains 400 million words (1990-2010) and BYU-BNC contains 100 million words of the [British National Corpus](#) (1970s-1993). Both corpora are accessible on the web. We present overall frequencies as well as results for spoken register because modals are more common in spoken language. Moreover, most functions of the modals are related to discourse functions and learners are more likely to get instruction in such contexts. Second, the written material we got from the learners is predominantly communicative (letters to friends or institutions and conversations).

⁵ See more details in Mitkovska & Buzarovska (in press).

Table 1 Distribution of modal verbs of obligation across levels

	MELC				Native corpora	
	A1	A2	B1	B2	COCA	BYU-BNC
MUST	6	95	172	157	All: 0.036%	0.057%
	0.041%	0.100%	0.107%	0.088%	Spok: 0.020%	0.042%
HAVE TO	5	79	208	183	All: 0.049%	0.040%
	0.034%	0.083%	0.130%	0.102%	Spok: 0.095	0.089%
SHOULD	1	61	153	221	All: 0.054%	0.078%
	0.006%	0.064%	0.095%	0.124%	Spok: 0.062%	0.070%
NEED TO	2	32	63	68	All: 0.023%	0.020%
	0.013%	0.033%	0.039%	0.038%	Spok: 0.030%	0.029%

At A1 we have a rather limited occurrence of modals for expressing obligation. This corpus was also relatively small and we cannot draw relevant conclusions regarding the distribution. The predominant modal verb there was *can* for permission and root possibility, while verbs for obligation occur in a few learned expressions. It is obvious though that at A2 all four modals are much more present. It becomes clear that *must* is considerably overused, even in relation to the more conservative British variant, while *have to* is slightly underused relative to the spoken registers where it is predominantly used by native speakers and would be expected in the contexts presented in learners' writings. The statistical data seem to indicate that *should* and *need to* are used appropriately by the learners, however a review of the examples shows that *need to* has not been fully acquired (this will be discussed later).

At B1 overuse of *must* continues, but the use of *have to* increases as well. These tendencies are slightly more pronounced in the adults' sub-corpus. We also noted slight overuse of *should*, while *need to* seems to be close to native frequency of use, considering that its domain is colloquial language. At B2 the use of *must* drops, but also does *have to*, while the use of *should* increases. This could be due to the types of texts written by upper intermediate students.

Analysis of the contexts of use attested in the data for each verb

As expected, the investigated modals were used predominantly in their deontic functions. In native corpora only epistemic *must* is more frequently used in less formal registers, while epistemic *should* and *have to* are only used in academic prose (Biber et al. 1999: 494). We attested only several uses of epistemic *must* in our corpus: 6 at level B1 and 13 at B2.

- (1) I write to come in your country. It *must be* fantastic. (B1)
- (2) I was playing basketball and I *must have lost* the money on the playground. (B2)

Since deduction and assumptions are actually taught at upper intermediate level, we may expect that learners have not acquired these structures fully for active use. An additional reason for this avoidance could be the fact that the modals *mora* and especially *treba* are not as commonly used in epistemic sense as the English counterparts, especially for past deduction. The adverbs *najverojatno*, *sigurno* etc. are more usual markers (*Sigurno/ Najverojatno veke stignale doma* 'They have surely/most probably arrived home).

Now we turn to the deontic uses of the modals at each level and determine the functions that learners predominantly employ with each modal. At A2 the system of modals for expressing obligation starts shaping up and we look at the developments for each modal verb, discussing points where there seem to be some characteristic discrepancies with the TL use.

At A2 *MUST* is used more often for circumstantial obligation. *Must* can be used in such functions, but it implies subjective attitude. Sometimes it seems that the learner simply felt this is strong obligation and used *must* as in example 3, where it sounds very unusual.

(3) I *must* visit to your country in november end I ask for some advise. (A)

Subjective attitude is more appropriately expressed when moral necessity is implied (4) and with strong advice (5). This is expected since learners find strong equivalence in such uses of Mac *mora*.

(4) We *must* all respect our teachers and future teachers.

(5) I think that we can go at the church, you *must* see it.

At B1 and B2, as well, when *must* is used for obligation and necessity it is not clear if the writer really intended to express subjective attitude, if the necessity really comes from the speaker or it is only circumstantial. In some contexts maybe *have to* or even *should* would be more appropriate as in (6), where there are external reasons that compel the subject to act.

(6) I'm busy this weekend, so we *must* cancel your visit for some other weekend. (B1, A) That means that I *must* study all day on Saturday, (B1, Ch)

In the adults' corpus *must* is used significantly as a marker for introducing a topic in the discourse, as in (7). Although this is possible, native speakers would often prefer *have to* in such case. Adults use of *must* for subjective/internal obligation (8) increases its occurrences at B2 level.

(7) I *must* tell you that I made body piercing and I change the color of my hair. (B1, A)

(8) Oh, I *must* take you out this weekend (B2, A)

HAVE TO is used predominantly for circumstantial necessity (9) at all levels. Only at A2 we find more unclear uses in the children's corpus.

(9) During the week I *have to* get up early. (A2, A)

Both adults and children at A2 and B1 use *have to* for advice, where *should* would be more appropriate (10 & 11). Here the equation with *treba* is obvious.

(10) You *don't have to* go shopping for a month if you like to save some money and buy more expensive tinks. (A2, Ch)

(11) What is the best time to visit Croatia, what place I *have to* visit and what is the best way to get there. (A2, A); What you think, which place I *have to* visit? (B1, A)

At B1 and B2 the negative form is used for absence of obligation (12), but we also encountered inappropriate negative forms of *have to* where *shouldn't* is to be used (13).

(12) So you *don't have to* go to hotel. (B1, Ch)

(13) Some smokers associated in different organizations around the world think that they *don't have to* be treated as children and they should do whatever they want (B2, A)

At A2 *SHOULD* is used for moral obligation (appropriate action) (14). Adults have used it predominantly for advice and suggestion (15).

(14) We *shouldn't* cut the trees because that can pollute the air. (A2)

(15) I visit to Venice for two weeks. Do you think I *should* hire a car or travel by bus? (A2)

The use of *should* increases at level B1 and especially at B2. Apart from the basic use, for correct or sensible action, it is extensively used for speech acts: suggestions, giving and asking for advice and arrangements (16). There are examples of indirect suggestion or advice (17), where *should* does not have a direct translation equivalent in Mac.

(16) *Should* I wait for you? (B1) When and where *should* we meet? (B2)

(17) Me and my friends were discussing about where we *should* spend the weekend (B1)

The use of *NEED TO* is more problematic. At A2 learners don't seem to have made the distinction between *need to* and *have to/should* as all instances are equated with *treba*. Thus learners often use *need to* instead of *have to* or *should*, but this also continues at B1 and B2 to some extent (example 18 and 19).

(18) Betoven don't understand the point of this film but the director is very proud because the movie *needs to* be komedi. (A2 Ch) You *need to* show me around (B2)

(19) Every building *need to* have video surveillance especially schools. (A2 A) Can you please tell me some good places which I *need to* visit? (B1)

Most common errors in form

The most common problems with the form of the constructions of the modal verbs are connected to the form of the lexical verb. The use of *to* where bare infinitive is required (20-21) can be due to the analogy with other forms (*want to*), but also to L1 influence as in Mac modal verbs are all followed by *da + V*. This tendency is especially pronounced with *must* and occasionally persists up to B2 level, probably because of the use of *to* with the similar *have to*.

(20) We *should to* pass a great time together. (B1)

(21) It's wonderful place that you *must to* see it. (B2 A)

Use of inflected form of the main verb, -ing/-en participle, present or past tense form, is more present at A2 and B1 (22-24), but diminishes at B2.

(22) ... we *should to cleaning* a rubbish and put in a big bin. (A2)

(23) I *must said* that they are free and interesting. (B1)

(24) Oh I'm fine thanks but you *don't have to thanks* me about that. (B1)

At B1 past tense of the verb is often used for past reference while the modal remains in present.

(25) I *have to postponed* it for the next weekend. (B1)

At A2 *must* is used for past time (26), but at B1 and B2 *had to* is more consistently used.

(26) I *must to go* that day on the meeting. (A2) You *must be* here before 40 minutes (A2)

Modals with perfect infinitive occur first at B1, but still most frequently unfulfilled past is expressed with infinitive (27) or lexical verb in past tense, as well as past deductions (28). At B2 perfect infinitive is more consistently used (*you must have been very tired*, B2).

(27) You *should be* here at 10 o'clock! and now do you know how much you are late? (B1)

(28) I'm so sorry, I *must forgot* the alarm. (B1 epistemic)

Concluding remarks and implications for teachers

Comparison of the English and Macedonian systems of modal verbs for obligation suggested that learners will overuse *must* and confuse *have to*, *should* and *need to*, as their use will be equated with *treba*. The analysis confirmed that learners use *must* more than native speakers and that sometimes it may sound too strong. It seems that even at B2 they are not aware of its formal character and that native speakers avoid it in conversation. The equation of Macedonian *treba* with *have to*, *should* and *need to* leads to confusion of the use of the English modals, though the use of *should* seems to be acquired best, especially in its discourse functions.

The discrepancy of use of these modal verbs may also be due to cultural differences. It has been pointed out (Hinkel, 1995) that the use of modal verbs of obligation depends on social norms and values which differ among communities. That is why comparison with the Mac verbs may be misleading. It gives a wrong picture of when students can use the Eng modal verbs. The exposure of our students to real life input is limited and they don't have a chance to work out the meaning on their own. So they need help, which may be given in form of presentation of different situations and discussion on the use of modal verbs in such situations; namely why it is more appropriate to use *have to* than *must*, or *should* than *must*, etc. Students have to be made aware of the different attitudes and norms that these verbs express. They need to be made aware why the occurrences of *must* are dropping and the occurrence of *need to* and *should* are on the increase.

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 & Philadelphia: John Benjamins.
- Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education Limited.
- Hinkel, Eli (1995). The Use of Modal Verbs as a Reflection of Cultural Values. *TESOL Quarterly* Vol. 29, No. 2, 325-343.
- Leech, Geoffrey. N. (1989). *Meaning and the English Verb*. London: Longman Group Limited.
- McDouall, A. (2012). A Corpus Based Investigation into the Use of English Modal Auxiliaries by Adult Korean L2-learners. *KLING* Vol. 6, 33-44.
- Palmer, Frank R. (1986). *Mood and Modality*. Cambridge: Cambridge University Press.
- Stepahny, U. (1995). Function and Form of Modality in First and Second Language Acquisition. In Anna Giacaloneamat/Grazia Crocco Galeas (Eds.) *From Pragmatics to Syntax: modality in second language acquisition* (pp. 105-120). Tübingen: Narr.
- Sweetser, Eve E. (1990). *From Etymology to Pragmatics*. Cambridge: Cambridge University Press.
- Thomson A.J. & A.V. Martinet (1986). *A Practical English Grammar*, fourth edition. Oxford: Oxford University Press.
- Yule, George (1998) *Explaining English Grammar*. Oxford: Oxford University Press.
- Митковска, Лилјана & Елени Бужаровска (in press). Типови модалност во неменливите *има да*-конструкции, *Македонски јазик LVIII*.
- Чашуле, Илија (1989). Модалните глаголи во македонскиот јазик. *Прилози*, том 14/2. МАНУ: Скопје, 89-117.