

May or might?

A corpus-based study of *may* and *might* in
the interlanguage of Norwegian and
Japanese learners of English

Camilla Jørgensen Ruud



A Master's Thesis presented to the Department of Literature,
Area Studies and European Languages

Faculty of Humanities

In Partial Fulfillment of the Requirements for the Master's Degree

UNIVERSITY OF OSLO

MAY 2014

May or might?

A corpus-based study of *may* and *might* in the interlanguage of Norwegian and Japanese learners of English



A Master's Thesis presented to the
Department of Literature, Area Studies and
European Languages

Faculty of Humanities

In Partial Fulfillment of the
Requirements for the Master's Degree

UNIVERSITY OF OSLO

MAY 2014



© Camilla Jørgensen Ruud

2014

May or might? A corpus-based study of *may* and *might* in the interlanguage of Norwegian and Japanese learners of English

Camilla Jørgensen Ruud

<http://www.duo.uio.no/>

Print: Reprosentralen, Universitetet i Oslo

IV

Abstract

This study makes use of the Contrastive Interlanguage Analysis (Granger 1996) to examine the differences in use of modal auxiliaries *may* and *might* in the interlanguage produced by Norwegian and Japanese learners of English, and compares it to the use in language produced by native speakers. The material is extracted from the Norwegian and Japanese components of the International Corpus of Learner English, ICLE-NO and ICLE-JP respectively, and the Louvain Corpus of Native English Essays, and analysed according to a set framework for modal meaning based on Palmer (1990) and Collins (2002). To attempt to explain the differences between each learner group, references will be made to Contrastive Analysis and the Integrated Contrastive Model (Gilquin, 2000/2001).

The main goal of this thesis is to compare and discuss how Norwegian and Japanese learners of English use *may* and *might* compared to native speakers, and if there are any indications of overuse or underuse. The assumption that there could be an overuse among Norwegian learners is mainly based on Aijmer's (2002) study on modal auxiliaries in Swedish. As for Japanese, modal auxiliaries are used in such a different manner that one would assume this would have some kind of effect on the interlanguage produced by Japanese learners. The main findings prove that while the use differs in terms of frequency, i.e. occurrences per 10,000, the distribution of each modal according to modal meaning is proportionally similar in all groups. However, the ways in which modals are used within each category also differs. Example sentences showing similarities and differences are provided to illustrate these differences.

The thesis concludes that both Norwegian and Japanese learners of English seem to have a good understanding of how to use *may* and *might*, although there are a few prominent differences. Norwegian students overuse *might* to an almost extreme extent, while some Japanese learners seem to have problems with the syntactic features of modal auxiliaries.

Acknowledgements

To my supervisor, Associate Professor Signe Oksefjell Ebeling, whose patience, helpful feedback, corrections and support have been invaluable to me.

To my friends, who all have supported me throughout the writing process, and to Yvonne, who kindly helped me with proofreading on an extremely short notice.

To my family, whose love and support helped me through the harsher times.

To Hiep, whose patience with me has been incredible.

To my mother, who always supported and believed in me, no matter what.

And finally, to my grandmother, who always encouraged me whenever I wanted to give up. I wish I would have been able to share this moment with you.

'Learning is a treasure that will follow its owner everywhere.' - Chinese Proverb

Table of Contents

1	Introduction	1
1.1	Background and scope.....	1
1.2	Aim and research questions	2
1.3	Outline of the Thesis.....	3
2	Theoretical background.....	4
2.1	Modal Auxiliaries	4
2.1.1	Syntactic function.....	4
2.1.2	Semantic function.....	5
2.2	<i>May</i> and <i>might</i> in the English language.....	7
2.2.1	<i>May</i>	7
2.2.2	<i>Might</i>	8
2.3	Modal auxiliaries in Norwegian and possible English correspondences	10
2.3.1	Syntactic function.....	10
2.3.2	Semantic function.....	12
2.4	Modal auxiliaries in Japanese and possible English correspondences	13
2.4.1	Syntactic function.....	14
2.4.2	Semantic function.....	15
3	CIA and Learner Language	17
3.1	Defining interlanguage	17
3.2	Contrastive Interlanguage Analysis.....	17
3.3	The Integrated Contrastive Model.....	20
4	Methods and material	22
4.1	Corpora	22
4.2	Extracting data.....	22
4.3	Framework of classification	23
4.4	Overview of extracted material	24
5	Analysis.....	25
5.1	Introduction	25
5.2	Analysis of <i>may</i>	26
5.2.1	Native speakers' use of <i>may</i>	26
5.2.2	Norwegian learners' use of <i>may</i>	28

5.2.3	Japanese learners' use of <i>may</i>	30
5.2.4	Comparisons.....	32
5.3	Analysis of <i>might</i>	35
5.3.1	Native speakers' use of <i>might</i>	35
5.3.2	Norwegian learners' use of <i>might</i>	37
5.3.3	Japanese learners' use of <i>might</i>	40
5.3.4	Comparisons.....	42
6	Conclusions	46
6.1	Suggestions for further studies	47
6.2	Final comments.....	48
	References	50

List of Figures and Tables

Figures

Figure 1: Contrastive Interlanguage Analysys (Granger, 1996 : 44)	19
Figure 2: Integrated Contrastive Model (Gilquin, 2000/2001: 100, Granger 1996: 47).....	20
Figure 3: Distribution of <i>may</i> according to modal type (percent).....	33
Figure 4: Distribution of <i>might</i> according to modal type (percent)	42

Tables

Table 1: Total number of occurrences of <i>may</i>	24
Table 2: Total number of occurrences of <i>might</i>	24
Table 3: Distribution of <i>may</i> according to modal type, LOCNESS.....	26
Table 4: Distribution of <i>may</i> according to modal type, ICLE-NO.....	29
Table 5: Distribution of <i>may</i> according to modal type, ICLE-JP.....	30
Table 6: Distribution of <i>might</i> according to modal type, LOCNESS	35
Table 7: Distribution of <i>might</i> according to modal type, ICLE-NO	38
Table 8: Distribution of <i>might</i> according to modal type, ICLE-JP	40
Table 9: Occurrences of <i>may not</i> and <i>might not</i> per 10,000 words.	44

Abbreviations

ACC	Accusative
ADJ	Adjective
ADV	Adverbial
AUX	Auxiliary
BNC	British National Corpus
CA	Contrastive Analysis
CIA	Contrastive Interlanguage Analysis
COMP	Comparative
COP	Copula
DEF	Definite article
DO	Direct object
GEN	Genitive
ICLE	International Corpus of Learner English
ICM	Integrated Contrastive Model
INF	Infinitive
IL	Interlanguage
L1	First language
L2	Second language / foreign language
LOC	Locative
LOCNESS	Louvain Corpus of Native English Essays
NL	Native language
NOM	Nominaliser
PASS	Passive
QES	Question marker
S	Subject
TOP	Topic
LOC	Locative
V	Verb

1 Introduction

1.1 Background and scope

During the years I have spent as both a student of English and in the teacher education programme at the University of Oslo, I have become familiar with the concept of second language acquisition and learner language research. Various lectures and seminars in corpus linguistics have all shown that Norwegian learners tend to have certain problems with modality, for example in the case of using modal auxiliaries. In many cases these auxiliaries seem overused, as found with Swedish learners in Aijmer's study of modality in advanced Swedish learners' written interlanguage (2002). Having completed a year of studies in Japanese recently, I also became curious as to whether or not Japanese learners, whom all speak and write in a language very different from both Norwegian and English, would have the same tendencies to overuse modal auxiliaries as other learner groups, or if there would be any other differences in use.

The use of modal auxiliaries is perhaps one of the more difficult concepts for a learner of English to grasp, and there are a few indications that this is indeed the case. While Aijmer (2002) has briefly touched upon how modal auxiliaries are used in Swedish interlanguage, there have been few studies discussing modal auxiliaries used by Norwegian learners in depth. Perhaps even fewer have discussed modals in the interlanguage produced by Japanese learners. However, while there are few studies done on modal auxiliaries in learner language, there are quite a few other studies done on modality, which all proved to be invaluable resources when writing this thesis. The main source for many of the theories and assumptions I have concerning Norwegian learners, for example, will be Løken's (1996) contrastive analysis of *can*, *could*, *may*, *might* and the Norwegian *kunne*. The work of Palmer (1990) and Collins (2002) done on modal auxiliaries have proved to be invaluable, providing a great overview of their uses in the English language. For the uses of modals in Norwegian, Eide (2005) has done a comprehensive study on the topic, collecting and connecting the many views on modals of scholars before her. For the use of modal auxiliaries in Japanese, the detailed study done by Johnson (2003) proved to be extremely helpful, and so did Narrog's (2009) study on modality in the Japanese language.

This study is corpus-based, i.e. it will use ‘corpus data in order to explore a theory or hypothesis’ (McEnery & Hardie 2011: 5 – 6). In order to allow a certain degree of depth in the discussion and analysis of modals carried out in this study, its scope is limited to the two modal auxiliaries *may* and *might*. There are in particular two reasons why these two modals were chosen. Firstly, Aijmer’s study showed that both *may* and *might* were heavily overused by Swedish learners (2002: 61), and since Swedish and Norwegian are fairly similar languages, one could suspect the same to be true for Norwegian learners. As for Japanese, *may* and *might* prove interesting as there are, in fact, corresponding modal auxiliaries in Japanese, although they are used quite differently. Secondly, these two modals are found to express a wide range of modal meanings, and thus it would be interesting to examine how the distribution between these are in the interlanguage produced by Norwegian and Japanese learners as opposed to the language produced by native speakers.

1.2 Aim and research questions

Despite its limited scope, the aim of this thesis will be to contribute to the studies done on modality, more specifically modal auxiliaries, in learner language. I will throughout this thesis investigate the usage of the modal auxiliaries *may* and *might* in the interlanguage produced by Norwegian and Japanese learners of English by using computerised corpora. The ultimate goal of this thesis is to compare the use of *may* and *might* in both learner groups to each other, as well as to language produced by native speakers. By doing so, I hope that this thesis will be able to shed some light on how *may* and *might* are used by the learner groups in question, and perhaps be used to predict learner ‘mistakes’ or diagnose possible challenges learners might have with these two modal auxiliaries. This ability to predict challenges is quite useful, especially for learners of English. If challenges concerning modal auxiliaries can be identified, possible errors can be prevented. Hence, exploring the use of *may* and *might* among Norwegian and Japanese learners may provide us with valuable insights for further studies as well as language learning in general.

Based on the hypothesis that there are differences in use with regard to these modal auxiliaries, as presented in section 1.1, there are three research questions which I would like to examine further:

1. How do Norwegian and Japanese learners of English use *may* and *might* compared to native speakers?
2. Is there any indication of overuse or underuse, and, if so, why?
3. Is there any difference in the meaning conveyed by the modals as used by the three groups?

These three research questions will be central to the analysis carried out in chapter 5.

1.3 Outline of the Thesis

Following this introductory chapter, Chapter 2 will focus on the theoretical background for this study; *may* and *might* as modal auxiliaries in English, and their correspondences in Norwegian and Japanese. Chapter 3 will deal with Contrastive Interlanguage Analysis and learner language, explaining the background of the method used for this study. This subject is addressed in a separate chapter due to CIA being vital as a framework for analysing learner language, and hence I did not feel I would be able to stress this point enough other than by doing so. Chapter 4 will continue on describing the methodology employed in the study, such as framework of classification, and various methodological issues that arose. The material will also be presented in the final section of Chapter 4, before being thoroughly analysed in Chapter 5. The analysis will be split into two parts, as each auxiliary will be discussed separately; one for *may* and one for *might*. Finally, Chapter 6 will sum the main findings of this study, answers the research questions, and suggests ideas for further work.

2 Theoretical background

This chapter deals with how *may* and *might*, and modal auxiliaries in general, are used in English, Norwegian and Japanese. It will also look at possible correspondences for *may* and *might* found in Norwegian and Japanese, and touch briefly upon how modal auxiliaries are used in these two languages. As this study will not attempt to do a full contrastive analysis of *may* and *might* and instead focus on the contrastive interlanguage analysis, the background information on how modals are used in each language provides vital for discussion, and thus will be included in this section.

The first section deals with modal auxiliaries in English in general, and how it can be possible to classify them. The second section looks at *may* and *might*, and explore their functions in the English language. The final two sections deal with how modals are used in Norwegian and Japanese respectively, and possible correspondences to *may* and *might*.

2.1 Modal Auxiliaries

Studies of modality in the English language are quite numerous, and the study of modal auxiliaries, which are central to this field, is no exception. Due to the limited scope of this study, its primary focus will be on Palmer's (1990) and Collins' (2009) work on modal auxiliaries in English. Their classifications of modal auxiliaries, both in terms of syntax and function, will be the starting point for the learner language analysis in chapter 5. Section 3.1.2 concerning function will be using Palmer (1990) as its primary source for discussion.

2.1.1 Syntactic function

According to Palmer (1990: 14-21), modals in the English language can be identified by using four criteria, all related to their syntactic functions. Collins (2009: 12) also mentions these criteria as being what separates modals from lexical verbs. The 'NICE' criteria, which they are called, all describe four different constructions in which modals can be used.

Negation. Auxiliaries can be used with the negative particle *not*, thus having negative forms.

- (1) It may not come out right. (BNC BMS 3238)

Inversion. The word order of subject and auxiliary can be inverted so that the auxiliary comes before the subject in a sentence, for example in the case of asking questions.

(2) May I come with you? (BNC AT7 2254)

Code. Post-verbal ellipsis dependent for its interpretation upon previous context (Collins 2002: 12).

(3) Let's see keep keep it on the same line if you can. Can you? (BNC K LW 1394)

Emphasis. Emphatic polarity involving the use of contrastive stress (ibid).

(4) They think he won't study, but he will (ibid).

Palmer also adds three other criteria that can help us distinguish modals from other auxiliaries (1990: 4).

- No *-s* form in the third person singular. Modal auxiliaries are not inflected in the 3rd person singular of the present tense.
- No non-finite forms. Modals can only occur in finite form, as the first element of a verb phrase.
- No co-occurrence. Two modal auxiliaries cannot occur in the same clause.

However, as also pointed out by Løken (1996), these last three criteria overlap to some extent, as no non-finite forms would also mean there would be no co-occurrence. These three criteria fit well with those mentioned by Quirk et al. (1985: 137), who also say that all modal auxiliaries are followed by the infinitive.

2.1.2 Semantic function

In terms of semantic function, modal auxiliaries are often classified according to what type of modality they express. Von Wright first introduced three types of modality in 1951; epistemic, deontic and dynamic modality (Palmer, 1990: 6-7, 35-38). As will be discussed in this chapter, these types are widely used not only to describe the types of modality in the English language. However, they may also be applied to other European languages such as Norwegian, and, to some extent, even Asian languages such as Japanese, as explored in section 2.3 and 2.4.

Epistemic modality, according to Palmer, is the simplest type of modality to deal with (1990: 50). It expresses possibility, marked by *may* and *might*, and necessity, which is marked by the usage of *must*. The function of epistemic modals is to express the speaker's judgment of whether something is possible or not, i.e. how the speaker judges the truth of propositions.

(5) This may be true. (BNC CEU 203)

(6) You must be tired after your journey. (BNC APM 2473)

Example 5 denotes epistemic possibility in the sense that the speaker is expressing his or her attitude towards whether something is true or not, accepting that there is a possibility that this something may be the case, without being certain. Example 6, on the other hand, is an example of how epistemic modality can express necessity. Necessity, as opposed to possibility, is a much more complex form of epistemic modality. As this paper concerns itself with *may* and *might* only, which are used only when dealing with epistemic possibility (Ibid: 50-53), we will not venture further into how epistemic modals are used to express necessity.

While the distinction between epistemic modality and other kinds of modality is fairly clear, the same cannot be said for the distinction between deontic and dynamic modality (ibid: 69). However, Palmer also states that the two types both seem to have different semantic and syntactic features, which would justify their treatment as two different types.

Deontic modality concerns itself with the act of giving permission, obligations, and making promises and threats. As such, it is often described as performative, since the speaker grants or asks for permission to act out the proposition of the clause in question. As with epistemic modality, deontic modality can be subdivided into possibility and necessity (ibid). Deontic possibility is often marked by either *may* or *can*, *must* or *shall*. Example 7 below is an example of *may* being used to express deontic modality in the form of giving permission.

(7) You may leave us. (BNC CK0 393)

Dynamic modality, on the other hand, is often expressed by the use of *can* and *will* (Ibid: 83-91). It deals with ability, i.e. describing the abilities of the subject to perform the action stated in the proposition of the utterance, as well as volition and non-deontic root meanings.

Therefore, while deontic modality is considered to be 'speaker-oriented', i.e. the speaker appears to be the deontic source, dynamic modality is concerned with the ability or volition of the sentence's subject, thus being 'subject-oriented' (Ibid: 36). As explained by Huddleston

and Pullum (2002: 179). ‘dynamic modality is less central to modality than deontic permission in that it does not involve the speaker’s attitude to the factuality or actualization of situation’, which they consider to be central to modal meaning. Example 8 below shows how *can* may be used when describing an ability of the subject in the sentence, which is one feature of dynamic modality.

(8) That guy can play the wind. (BNC ASA 805)

In addition, deontic modality is performative or ‘discourse-oriented’, something which dynamic modality is not. In this respect, it differs from the discourse oriented modalities mentioned above, i.e. modalities that involve both the speaker and the addressee (Palmer 1990: 83-88).

2.2 *May* and *might* in the English language

Historically, *may* and *might* have been treated as one lexeme, something that is also done by Palmer (1990), Mindt (1995) and Høye (1997). It is worth noting however, that linguistics in more recent times, such as Collins (2009) and Huddleston and Pullum (2002), suggest that for many speakers today, *may* and *might* are forms of different lexemes rather than forms of the same lexeme (Collins, 2009: 117). Both *may* and *might*, being present and preterite forms respectively, are used when expressing possibility (Oxford Dictionaries, 2014). While some insist that the distinction should be made when describing current and past situations, they also argue that this distinction is rarely made today. Whether or not these verbs are separate lexemes or not will not be dealt with in this paper, but they will be treated as independent forms due to their differences in usage.

2.2.1 *May*

The studies done on *may*, both on its meaning and use, are many in number, and opinions seem to be quite divided. According to Mindt (1995: 101), there are two basic modal meanings of *may*; possibility and permission. As we saw in section 3.1.2, *may* is mainly used as an epistemic and deontic modal, and this supports Mindt’s view of meanings for *may*. Examples 9 and 10 below show *may* used as an epistemic expressing possibility, and a

deontic dealing with permission, both being within the range of Mindt's (1995) and Palmer's (1990) definitions for *may* as a modal auxiliary.

- (9) The organisation making the most economic use of nurses may expect to have a good proportion of its staff on part-time contracts. (BNC EVY 594)
(epistemic, possibility)
- (10) 'May I see your left hand, please sir?' (BNC CDA 21)
(deontic, permission)

Collins (2009: 95-96), however, argues that there also are ways in which *may* can function as a dynamic modal. He mentions two possible dynamic uses of *may*; 'theoretical possibility' and 'dynamic implication'. The former 'involves a potentiality for action that resides in the external situation', while the latter 'involves a potentiality for action that is the basis for an implied directive speech act' (ibid: 96). In his study, Collins found that while most occurrences of *may* found were examples of epistemic modality, *may* in its dynamic meaning actually appeared more frequently than *may* in its deontic forms (ibid: 92). In fact, *may* is sometimes used as a substitution for *can* in its dynamic sense, especially when the situation requires a greater degree of formality, which is illustrated in example 11 below.

- (11) Nevertheless it may be observed that in the earliest phases of the life of the Cowdery's Down settlement when all the buildings were closely associated with fenced enclosures the majority of bone and cereal recovered came from buildings straddling the fence (BNC CFK 277).

It would therefore be safe to assume that *may* in modern English can express all three modal meanings, depending on the context.

2.2.2 Might

As with *may*, the opinions on the use of *might* as a modal auxiliary are divided. While *may* has two main modal meanings, Mindt (1995: 108) argues that *might* only has one, namely possibility. This view is also shared by Palmer (1990: 58), who states that *might* has all the same functions as *may* when denoting epistemic possibility, albeit *might* indicates a little less certainty about the possibility from the speaker's point of view. In addition, it is solely

epistemic when used in certain structures, such as when *may* could be interpreted as either epistemic or deontic depending on the context given by previous clauses or sentences.

(12) He might go to the bathroom, or come downstairs even. (BNC ABX 3476)
(epistemic, possibility)

(13) Those who won't pay may go to court. (BNC AM5 1450)
(epistemic, possibility or deontic, permission)

Collins (2009: 111-112), however, argues that this particular view may be outdated, as recent data show that they can in fact be used interchangeably. He clarifies that *might* is far more common in spoken than written language, while *may* has maintained its position as the most used modal for marking epistemic possibility in written language.

As with *may*, *might* may be used as a dynamic modal, as well as a deontic one (Collins, 2009: 113-117). *Might* is also used to describe theoretical possibility, shown in example 14, as well as dynamic implication such as suggestions, as seen in example 15.

(14) The verderers were closely examined at the Forest Eyre as to the facts to which they deposed in their rolls of presentments, and even minor contradictions in their evidence might result in their committal to prison and subsequent amercement. (BNC AE9 386)

(15) And I think you might remember that my 'Jottings' have been in the paper since long before your grandmother had her first G.I. in the War. (BNC FB9 479)

When expressing deontic possibility or permission, *might* was, according to Collins, frequently used in questions, as seen in the above example 11, and in conditionals with the addressee as the deontic source, as in example 16 (2009: 117).

(16) If I might be so bold, for one who missed out on so much (and not just in a cricket sense), I would offer a few words of advice. (BNC EB3 8)

Conditionals, such as the example above, makes use of deontic *might* to ask for or give permission, with a higher degree of speaker involvement.

2.3 Modal auxiliaries in Norwegian and possible English correspondences

Modality in Norwegian can be expressed through various means. According to Lie (1976: 60), one is the use of modal auxiliaries, which will be the focus of this study. However, studies done in the field of modal auxiliaries in the Norwegian language are rather scarce. In fact, there is no definition as to which auxiliaries belong to this category in neither Lie (1976) nor Rødhovd (1993). This is a problem also addressed by Løken (1996) in her study of Norwegian *kunne* as a correspondence to *may*, *might*, *can* and *could*, where she mentions that there is no agreement as to which verbs should be labelled as modal auxiliaries in Norwegian. She mentions that there seems to be an agreement that *kunne*, *skulle*, *ville*, *måtte* and *burde* are central modals in Norwegian. These modals are also confirmed by Eide (2005: 16) and Faarlund et al. (1997: 527) to be the five central ones in Norwegian, as they all have deontic, dynamic and epistemic readings. Their characteristics will be discussed further in the sections below.

The Norwegian auxiliaries *kunne*, *ville*, *må* and *måtte* are often mentioned as possible translations for *may*, and *kunne*, *hadde* and *ville* as possible translations for *might* (Stor Norsk-Engelsk Ordbok, 2014). Løken (1996: 128), in her contrastive analysis of *may* and *might* and Norwegian *kunne*, identifies the structure *kunne* and *kunne* + ADV as the most common correspondence to *might* when denoting both epistemic, deontic and dynamic possibility. In the case of *may*, there was a lot more variation. The most common structure, however, involved the use of Norwegian *kanskje* in combination with verbs, either lexical ones or auxiliaries. In all these cases, *may* denoted epistemic possibility. *Kan* was also shown to have a high frequency of correspondence with *may* (Løken, 1996: 114 - 118), and was found to correspond to all meanings except 'wish'. Hence, *kan* and *kunne* would seem like good translations for *may* and *might*, respectively.

2.3.1 Syntactic function

Syntactically, Norwegian modal auxiliaries differ from their English counterparts in many ways. While there is no co-occurrence of modals in the standard English verb phrase (Palmer 1990: 4), Norwegian, and other Scandinavian languages, allows co-occurrence of modals,

albeit with certain restrictions (Eide 2005: 55). In addition, Norwegian modals can occur in non-finite forms, whereas they cannot in English. Consider the following examples, taken from Eide (Ibid).

- (17) Det må kunne finnes en løsning.
There must canINF findPASS a solution.
It must be possible to find a solution.

In this example, Norwegian has two modals serving as auxiliaries, *må* and *kunne*. While this sentence is perfectly fine in Norwegian, the co-occurrence of two modals is unnatural, and ungrammatical, in Standard English.

- (18) Jon antas å måtte være morderen.
Jon presumePASS to mustINF be killerDEF.
Jon is presumed to have to be the killer.

If we were to discuss the NICE criteria mentioned by Palmer (1990) and Collins (2009), and how they apply to the Norwegian modal auxiliaries, only ‘code’ would be relevant when examining the differences between modal auxiliaries and lexical verbs (Eide, 2005: 65). Norwegian modals are used just like regular lexical verbs with regard to negation, inversion and emphatic affirmation. However, as explained by Eide, lexical verbs in Norwegian are replaced by the light verb *gjøre*, meaning ‘to do’, in elliptical constructions, as seen in example 19 below. In this sense, Norwegian and English modals seem to follow a similar pattern. Modals, on the other hand, are repeated, as shown in example 20. The following examples are taken from Eide (Ibid).

- (19) Marit svømmer, og det *svømmer / gjør Jon også.
Marit swims, and *that swims / does he too.
Marit swims, and so does Jon.

- (20) Marit skal svømme, og det skal / *gjør han også.
Marit shall swim, and that shall / *does he too.
Marit is going to swim, and so is Jon.

Hence, while the NICE criteria is perhaps not that relevant when discussing Norwegian modal auxiliaries, there are other syntactic features that have to be taken into consideration.

2.3.2 Semantic function

According to Eide (2005: 46), *may* and *might* in the Norwegian language, when used in either the epistemic sense or the deontic sense, correspond to the verb *kunne*. This is further supported by Løken (1996: 141), as she concludes that *kunne* is the Norwegian verb corresponding to most meanings of *may* and *might*. Løken (Ibid: 108) lists these meanings, or usages, of *kunne* in the Norwegian language as epistemic possibility, ability, permission, wish, suggestion and root possibility, which corresponds to deontic possibility in the studies by Eide (2005: 46).

Since *kunne* is also connected to dynamic modality as a possible translation of the English modal *can*, *kunne* is seemingly far more versatile than *may*. Consider the following example, taken from Eide (2005: 42).

(21) Marit kan være uforsiktig.

There are several possible ways of reading this sentence, as it, in Norwegian can possess all of the three following meanings:

It is possible that Marit is careless. (epistemic)

Marit is allowed to be careless. (deontic)

Marit is capable of being careless. (dynamic)

One possible translation could be ‘Marit may be careless’, in the epistemic or deontic sense. ‘Marit can be careless’ denotes not only dynamic modality, but also deontic in the sense that she is, in some way, permitted to be careless. Hence, *kan* in Norwegian would correspond to all three modal readings. In English however, there is a slight difference as to what modal reading *may* and *can* would denote.

In fact, most of the English verbs and constructions corresponding to *kunne*, according to Løken (1996: 151-153), denote deontic possibility, especially when *can* is used as a corresponding verb. Her study also shows that the majority of the instances where *kunne*, and other forms of this verb such as *kan*, are used, corresponds to the epistemic use of *may* and *might* (ibid: 108-142). There are also instances where *may* and *might* are used in the sense of giving permission and suggestions, although they are less common. Eide (2005) further comments on the epistemic use of *kunne* as having only one reading in contemporary

Norwegian; ‘it is possible that’ (ibid: 43-44). She adds that *kunne* also has one deontic reading, and one dynamic reading, where it corresponds to the English *can*. Hence, it would seem that *kunne* is more prominent as a Norwegian correspondence to English *may* and *might* when the latter two are used in their epistemic sense.

Based on the fact that *kunne* can correspond to both *may*, *might*, *can* and *could* in English, one would perhaps expect Norwegian learners to have trouble distinguishing the use of these four modals, and thus use some more than others leading to overuse or underuse. In addition, while Norwegian allows co-occurrence of modal auxiliaries, this is not allowed in English, and this could also be problematic for learners.

2.4 Modal auxiliaries in Japanese and possible English correspondences

Modal auxiliaries are used in a different way in the Japanese language than in Norwegian and English, and it has long been a common conception that modality in Japanese is not comparable to the conceptions found in Western studies. As Johnson explains it, ‘modal content in Japanese involves a much wider variety of subcategories, such as auxiliary verbs and sentence final particles, and therefore corresponds neither to the idea of mood nor modality present in European languages’ (Johnson, 2003: 4). In the case of Japanese and Norwegian, however, both languages share the usage of modal particles, while these types of particles do not exist in English (Eide, 2005). Despite the differences between these three languages, she points out that Japanese modal auxiliaries do have some common features with their English counterparts, such as the concepts of epistemic, deontic and dynamic modality and the notions of necessity and possibility, points that are often discussed in English language studies of modality.

Another challenge of comparing the use of modal auxiliaries is the fact that Japanese sentence structure is fundamentally different from that of Norwegian and English. While both English and Norwegian are verb-object languages, Japanese is an object-verb based language (Tomlin, 1988). Consider the following sentence, taken from Makino & Tsutsui (1986: 18).

- (22) Mr. Smith (S) is studying (V) Japanese (DO) in Japan (Adv).
Herr Smith (S) studerer (V) japansk (DO) i Japan (Adv).

Sumisu-san wa (S / TOP) Nihon de (Adv) nihongo wo (DO) benkyoushiteiru (V).
(Mr. Smith (S) in Japan (Adv) Japanese (DO) is studying (V))

Therefore, using modal auxiliaries when forming clauses in Japanese is also done differently. As noted above, modal auxiliaries in Japanese do share some of the functions found to be applicable to English modals, which will be further explored in the following sections. All examples in this chapter, and its subsections, are taken from Makino & Tsutsui (1986).

2.4.1 Syntactic function

In Japanese, auxiliaries follow immediately after the stem form of a verb, and thus changes the verb's form and meaning (Johnson, 2003: 17-23). While the Japanese auxiliaries are similar to their English counterparts in terms of meaning and function, the way in which sentences containing these auxiliaries are formed is quite different. Observe the following example.

- (23) Ano hito wa kyō no pāti no koto wo wasureta.
That person (TOP) today (GEN) party (NOM) about (ACC) forget-PAST.
That person forgot about today's party.

By adding the modal auxiliary *kamoshirenai*, to the main verb *wasureta*, or in the case of this particular example the polite form *kamoshiremasen*, the content of this sentence is modified by the auxiliary.

- (24) Ano hito wa kyō no pāti no koto wo wasureta kamoshiremasen.
That person (TOP) today (GEN) party (NOM) about (ACC) forget-PAST (V) might (AUX).
That person might have forgotten about today's party.

This particular morphological feature of agglutination, as explained by Johnson (2003: 18), means that most of the Japanese auxiliaries cannot be viewed as a morphological entity on their own. They are attached to the main verb, as seen in example 23, and follow strict rules in terms of order and what kind of verb stem they follow.

Using the NICE criteria mentioned by Palmer (1990) and Collins (2009) to classify these modals would also be rather problematic due to the vast differences in language structure between the two languages, and perhaps not necessarily appropriate.

2.4.2 Semantic function

There are quite a number of possible sentence structures in Japanese used to express modality, but only a few involve modal auxiliaries that are possible correspondences to *may* and *might*. When *may* and *might* are used in the epistemic sense, the auxiliary *kamoshirenai* is generally considered to be the best correspondence. When expressing permission, *~te mo ii* is usually used (Johnson, 2003: 112-114), which fits well with the notion of deontic modality. Dynamic modality, on the other hand, is often expressed by using the ‘potential form’ of a verb, which is a special form of conjugating a verb expressing ability. It is also common to use *~koto ga dekiru*, a structure containing the potential form of the verb *suru*, meaning ‘to do’. These claims can of course be debated further, since there are other structures in Japanese which also might deal with these types of possibility. However, for the sake of simplicity, the focus of this paper will be on these three auxiliaries as possible correspondences for *may* and *might* in English.

Kamoshirenai is generally considered a correspondence of both *may* and *might* when used in its epistemic sense. It expresses uncertainty and probability, which is evident when considering its literal translation, ‘it cannot be known’ (Makino & Tsutsui, 1986: 172-175). When used in a sentence, *kamoshirenai* expresses a lower degree of certainty than sentences using *darō*, or *deshō* in polite writing and speech, meaning ‘probably’. As we saw in example 24, *kamoshirenai* attaches itself to main verb. However, it may also follow directly after an adjective.

(25) Kyōto no sakura wa mada kirei kamoshirenai.

Kyoto (GEN) cherry blossoms (TOP) still (AVD) beautiful (ADJ) might (AUX).

The cherry blossoms in Kyoto might still be beautiful.

In terms of asking and giving permission, the phrase *~te mo ii* is used, expressing either permission or concession. *~te mo ii* is attached directly to the stem of a verb, and carries the meaning of ‘is it alright if’ if translated into English (Makino & Tsutsui, 1986: 471-473). It

can be used for either asking permission, as seen in example 26, or giving permission, as in example 27.

- (26) Koko de tabako wo sutte mo ii desu ka?
Here (LOC) tobacco (ACC) smoke-even-if OK (COP) (QES)
Is it alright if I smoke here? / May I smoke here?

- (27) Hai, ii desu.
Yes, eat-even if OK (COP)
Yes, it is alright / Yes, you may.

Another interesting observation is that the potential form of a verb may also be used to express permission, as noted by Johnson (2003: 114). Instead of saying *tabete mo ii* when giving permission, you might in some cases be able to just say *taberareru*, which is the potential form of *taberu*, meaning ‘to eat’.

In order to express dynamic possibility in Japanese, the potential form of a verb may be used. Another option is to use a structure involving the usage of the potential form of the verb *suru*, *~koto ga dekiru* (Narrog, 2009: 96-97), which can be observed in example 28. This structure describes that someone or something can do something, or is able to do something (Makino & Tsutsui, 1986: 200-201). This structure is formed by using the potential form of the verb *suru*, which means ‘to do’, and basically translates directly into ‘thing can do’. *Koto*, meaning ‘thing’, is used as a nominaliser, hence making the preceding clause a substantive. Therefore, using *~koto ga dekiru* is a way of nominalising clauses so that they may be used as subjects. The structure itself is attached after a regular verb in the infinitive.

- (28) Taguchi-san wa chūgokugo wo hanasu koto ga dekiru.
Mr Taguchi (TOP) Chinese (ACC) speak (COMP) (NOM) am able to.
Mr. Taguchi can speak Chinese.

Hence, Japanese does have good correspondences in modal meaning for *may* and *might*. But as Japanese differs from English syntactically, one would expect Japanese learners to have problems with the use of modal auxiliaries in English. While Japanese is, in general, an agglutinative language, the same cannot be said for English. Whether or not this will have some consequences for the use of modal auxiliaries by Japanese learners will be explored in chapter 5.

3 CIA and Learner Language

3.1 Defining interlanguage

A key term in this chapter, and in this thesis as a whole, is ‘interlanguage’. The notion of interlanguage is central to research done in the field of second language acquisition and learner language research, as studying interlanguage aims to explain the differences between a learner’s produced second language to that of native speakers. As explained by Tarone (2006: 747),

The central object of interlanguage research is to explain this difference – essentially, to describe and explain the development of interlanguages and also to explain the ultimate failure of interlanguages to reach a state of identity with the target language.

In this study, interlanguage will be understood as a linguistic system that has been developed by a learner of a foreign language, L2, who has yet to become fully proficient in his or her target language (Ibid: 748-749). The learner’s interlanguage differs from the target language in the sense that the learner might use different strategies with roots in their first language, such as transfer and overgeneralization, to develop linguistic structures in the target language specific to that learner group. It is this theory of interlanguage that will form the basis for discussion and analysis in the following chapters of this thesis.

3.2 Contrastive Interlanguage Analysis

While interlanguage studies before the age of computerised learner corpora were limited in both scale and range, one could say that the introduction of technology reformed the way interlanguage studies were looked upon by scholars, as well as how these studies were conducted. Computerised corpora made it possible to increase not only the size of material available for research, but also the variety of texts available. This enabled researchers to adopt a more scientific and empirically-based approach, which scholars believed was previously lacking (Granger, 1996). The introduction of computerised corpora also opened up for new types of studies to be performed, as well as for new research methods to be used, as computer corpora give easier access to numbers such as frequency of occurrence and patterns of usage (Hasselgård & Johansson, 2011: 37).

The popularity of computerised learner corpora, and the increased interest in them which developed through the 1990s, are mainly to be credited to Sylviane Granger and her team at the Université catholique de Lovain. Granger was the main force behind the extremely successful and widely used International Corpus of Learner English, or ICLE (see also section 4.1). ICLE contains argumentative essays written by higher intermediate to advanced learners of English. The corpus is divided into several comparable sub-corpora based on the writer's mother tongue background (Granger 1996). The fact that these are all comparable enables us to study the relationship between interlanguages in further detail, for example in relation to whether a specific feature of learner language only belongs to one mother tongue group, or if it is common among learners in general. This of course can be quite useful for teachers and students alike, as it would perhaps predict some of the features, and perhaps, errors, common to learner groups.

Alongside the development of the computerised learner corpus itself, Granger also developed a framework for analysing learner language called Contrastive Interlanguage Analysis, also known as CIA. CIA, as opposed to the more traditional Contrastive Analysis, 'does not establish comparisons between two different languages but between native and learner varieties of the same language' (Granger, 1996: 43). For this purpose, the comparable sub-corpora of ICLE are perfect, but to compare with native speakers, one would need a compatible corpus of native speaker texts. For this purpose, the Louvain Corpus of Native English Essays, or LOCNESS for short, was compiled (Granger, 1995: 45). LOCNESS contains essays written by British A level pupils (further explained in section 4.1), British university students and American university students, and even with minor differences from ICLE in terms of essay topics and contributors, remains the best comparable corpus available (Hasselgård & Johansson, 2011: 38).

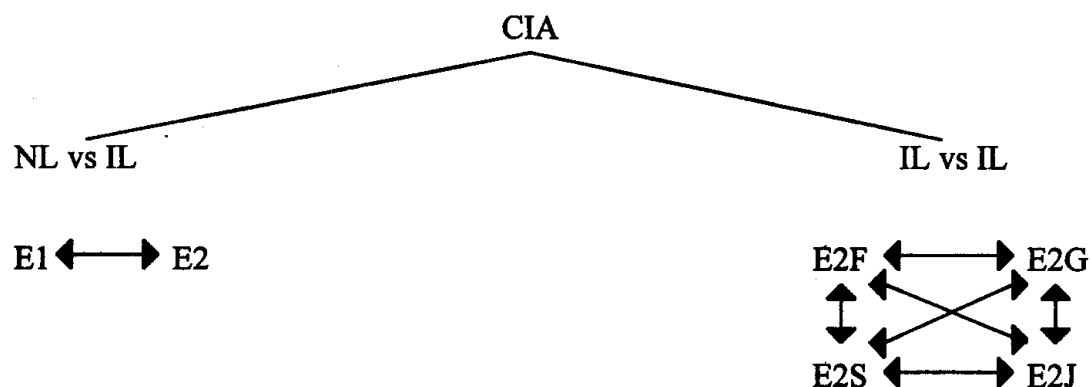


Figure 1: Contrastive Interlanguage Analysis (Granger, 1996 : 44)

With the corpora available, and the framework set, CIA involves two types of comparisons, as seen in figure 1 above (Granger, 1996: 44). The first would be comparing native language, NL, to interlanguage, IL. In other words, it is a comparison between texts that have been produced by native speakers with English as their first language, L1, and text produced by learners with English as a second language, L2. Its main goal is to document how L2 English differs from L1 English, where native speakers' texts are used as a control corpus. The second type of comparison is between different interlanguages, IL versus IL. Here, the different interlanguages of the L2 language, written by learner groups with different mother tongue backgrounds, are compared. Throughout this study I will be focusing on comparing the interlanguage produced by Japanese learners compared to that of Norwegian learners, which is an example of CIA being used for comparing interlanguages. Moreover, the other branch of CIA, in which the language of native speakers is compared to interlanguage, will also be addressed, as the interlanguage produced by both learner groups will be compared to the language of native speakers.

The advantages of having vast amounts of data stored are invaluable to CIA. Granger (1996: 45) points out that what she calls 'over- and underrepresentation' has to be done by using a quantitative-contrastive approach, which is only manageable by using corpora of significant sizes such as ICLE and LOCNESS. Previously neglected in learner language due the fact that quantitative studies were hard to conduct because of corpus size, the terms 'overuse' and 'underuse' are today widely used when comparing interlanguage to native language. Overuse happens when a word or an expression has a much higher frequency in an interlanguage than in the language of a native speaker. Similarly, there is underuse when there is a much lower frequency attested in the interlanguage than in the native language.

3.3 The Integrated Contrastive Model

As we saw in section 3.1, CIA is an invaluable framework for analysing interlanguage. However, to take full advantage of CIA, Granger (1996: 46-47) developed a method for integrating CIA and traditional Contrastive Analysis (CA), creating the Integrated Contrastive Model. By doing so, Granger devised a tool in which CIA and CA work together to explain not only certain features of interlanguage, but also predict or diagnose them. As explained by Hasselgård & Johansson (2011: 44),

this model offers a new dimension to interlanguages studies, enabling the researcher not only to differentiate general from L1-specific learner problems but also to explain and/or predict such problems on the basis of contrastive analyses of the L1 and the target language, in the spirit of the weak version of the contrastive analysis hypothesis.

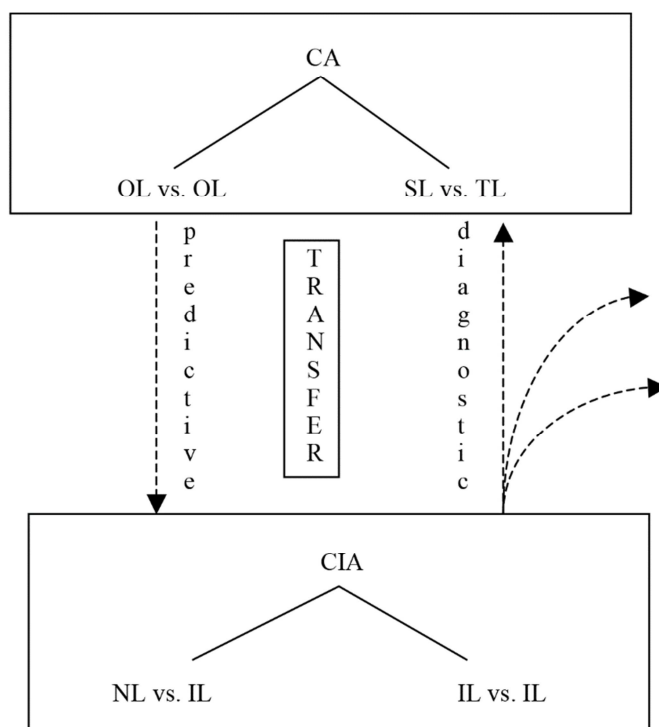


Figure 2: Integrated Contrastive Model (Gilquin, 2000/2001: 100, Granger 1996: 47)

The Integrated Contrastive Model, as illustrated in figure 2, involves a ‘constant to-ing and fro-ing between CA and CIA’, where ‘CA data helps analysts to formulate predictions about interlanguage which can be checked against CIA data’ (Granger, 1996: 46). This is shown in figure 2 by the arrow pointing from CA to CIA marked as ‘predictive’. Having checked

possible predictions using CIA data, the mismatches found between native language and learner language, or between two learner languages, can then be explained by going back to the contrastive analysis as a resource. This is shown by the arrow marked as ‘diagnostic’ in figure 2.

Central to the Integrated Contrastive Model is the notion of ‘transfer’, which means a learner group may let their mother tongue, L1, influence their acquisition of a second language, L2, either positively or negatively (Granger, 1996: 46). However, research conducted by Gilquin (2000/2001) raised questions on whether using only CA and the notion of transfer were enough to explain and diagnose all features found in interlanguage deviating from native language. Gilquin suggested some changes to the model, which in figure 2 above is represented by the arrows pointing out of the figure, to show that not all errors can be explained by contrastive analysis (Gilquin, 2000/2001: 100). In addition, the usage of broken lines between CA and CIA were also introduced in this revision, to indicate a weaker connection between the two, which is also pointed out by Hasselgård & Johansson (2011: 44-45).

This particular study’s main focus will be on that of the block marked CIA found in the Integrated Contrastive Model, and one could perhaps argue that the model is not necessarily relevant to the analysis carried out in chapter 5. I would, however, argue that it is, as the diagnostic part of the model is crucial in terms of explaining the results of the CIA carried out. In addition, thoughts on notions central to the Integrated Contrastive Model such as transfer will be addressed, although I will not be attempting a full CA due to the limitations of this thesis. Hence, the Integrated Contrastive Model will be used in discussion as a way of developing theories and thoughts on the interlanguage of Japanese and Norwegian learners compared to the language of native speakers.

4 Methods and material

4.1 Corpora

In this study, corpus methods will be used to perform the Contrastive Interlanguage Analysis (see chapter 3). The material that will be used for the analysis is drawn from the International Corpus of Learner English, ICLE, and the Louvain Corpus of Native English Essays, LOCNESS, both of which were introduced in chapter 3. This study will use both the Norwegian subcorpus of ICLE, ICLE-NO, and the Japanese subcorpus, ICLE-JP. ICLE-NO contains a total of 210.367 words, while ICLE-JP contains a total of 198.241 words (Granger, 1993, Granger, 1996). LOCNESS, on the other hand, contains a total of 324.304 words. Since the corpora differ a bit in size, normalised frequencies will be calculated to make the results comparable. By normalised frequencies, we understand it as ‘a frequency expressed relative to some other value, as a proportion of the whole – for example, frequency of a word relative to the total number of words in the corpus’ (McEnery & Hardie 2011: 247). The normalised frequency will be calculated per 10.000 words, unless otherwise stated.

As mentioned briefly in section 3.1, LOCNESS is fairly similar to both ICLE-NO and ICLE-JP in the sense that they all contain student essays. Hence, LOCNESS seems a fairly good choice for studies comparing the written English of learners and native speakers. However, one particular challenge is that the essay topics in these corpora do not necessarily always match, and it can thus be difficult to compare them as different words might be triggered by different environments and topics. For the purpose of this study, this particular problem is perhaps not very relevant, as all the texts should contain structures using *may* or *might*, regardless of their topic.

4.2 Extracting data

Using the ICLE v2 interface available on CD-rom (Granger et al., 2009) a simple search for *may* and *might* was carried out separately in both ICLE-NO and ICLE-JP, once for each verb in each learner group. Results were not filtered directly in the interface, but instead downloaded as a .pdf, printed, and gone through manually. For accessing LOCNESS, the

online version available at the University of Oslo was used.¹ Searches for *may* and *might* were performed in the web interface, and then saved as .pdf-files created from the web panel directly.

Although ICLE-NO, ICLE-JP and LOCNESS are computerized corpora, and even though they have a search interface, one is not guaranteed to find all examples of the word searched for (recall), and even if an item is found, it might not be the item one was searching for (precision) (Gries 2009: 16). Misspellings and similar lexical items are typical examples of possible problems. Unfortunately, it is almost impossible to search for misspellings, as one would have to know the exact way in which the item was misspelt to get a hit in the search interface. When seen in connection with this study, precision will probably not be a significant problem when searching for either *may* or *might*, since we will be looking at all instances where it is found, although instances of the nouns *May* and *might* will have to be discarded manually. Recall, on the other hand, could quite possibly be an issue. Since both ICLE-NO and ICLE-JP are learner language corpora, one could assume that misspellings are to be found more often in their texts compared to corpora containing works by native speakers. However, this is just an assumption. In addition, all search results will be checked manually to filter out any possible errors.

4.3 Framework of classification

This study will classify the material according to categories based on the semantic functions of *may* and *might* presented section 2.2, epistemic, deontic, and dynamic, as well as two additional categories, ‘ambiguous’ or ‘wrong’. The two latter categories were added as the three categories in section 2.2 proved to be insufficient. ‘Ambiguous’ contains examples that were too difficult to classify as belonging to one specific category, as they could be interpreted as belonging to more than one. The category ‘wrong’ contains instances where *may* and *might* are not used correctly by the learner groups. Some of these examples, which were identified as either ambiguous or wrong, will be further discussed in chapter 5.

The examples were then counted, and the total number of occurrences for each learner group, as well as occurrences per category, were registered. This was all done by hand and on paper, not using computerised databases. The normalised frequency of each verb in each learner

¹ The online version of LOCNESS available at the University of Oslo may be accessed at <http://www.tekstlab.uio.no/cgi-bin/omc/LOCNESSsearch.cgi>.

language group, as well as in that of the native group, was then calculated to make the number of occurrences comparable. An overview of these frequencies are presented in tables 1 and 2 in section 4.4 of this chapter. A further analysis was then carried out to decide which categories the verbs belonged to. Again, normalized frequencies were calculated for each category, each verb and each language group. These findings are presented in chapter 5.

4.4 Overview of extracted material

The material extracted from ICLE-NO, ICLE-JP and LOCNESS form the basis for analysis in chapter 5. Presented in the tables 1 and 2 below, are the filtered results, which means that possible noise, such as instances where *may* and *might* are used as nouns, already has been removed. These figures provide us with an interesting first impression of tendencies regarding overuse and underuse.

Table 1: Total number of occurrences of *may*

Corpus	Number of occurrences	Total word count	Occurrences per 10,000 words
ICLE-NO	268	210,367	12.7
ICLE-JP	293	198,241	14.8
LOCNESS	473	324,304	14.6

Table 2: Total number of occurrences of *might*

Corpus	Number of occurrences	Total word count	Occurrences per 10,000 words
ICLE-NO	271	210,367	12.9
ICLE-JP	130	198,241	6.6
LOCNESS	85	324,304	2.6

5 Analysis

5.1 Introduction

Tables 1 and 2 in section 4.4 provide us with a general overview of the differences in usage of *may* and *might* between Japanese and Norwegian learners of English, as well as the differences between these two learner groups and native speakers.

Interestingly, the use of *may* seems to be fairly similar in the written language of the three groups in question. While results from the LOCNESS show that native speakers are at 14.6 occurrences per 10,000 words, there is a slight underuse of *may* found in ICLE-NO where the results show a total of 12.7 occurrences per 10,000 words. Japanese learners' use of *may* is very close to that of native speakers, with 14.8 occurrences per 10,000 words.

Initial results for *might*, as seen in table 2, show vast differences in terms of frequency. Both learner groups, Japanese and Norwegian, were found to have quite a substantial overuse of *might* when compared to the use of *might* by native speakers. Although both learner groups overuse *might*, it is particularly noticeable in the results from ICLE-NO, where the frequency is twice as high as that of the frequency found in ICLE-JP. While there were only 2.6 occurrences of *might* per 10,000 words in LOCNESS, ICLE-JP had a total of 6.6 occurrences, and ICLE-NO an even higher frequency of 12.9 occurrences per 10,000 words. Hence, one might expect to find several differences in usage especially with regard to *might*, due to the differences in frequency, as opposed to *may* which seems to be more evenly distributed between the three groups.

While these statistics provide us with an interesting overview of the use of *may* and *might* as a whole, it does not take into account how these modals are used and what type of modality they are most likely to express in the language of each learner group, nor in the language of native speakers. Neither do they form any grounds of further discussion as to what may have caused the overuse of *might*, nor as to why there seems to be a slight underuse of *may* by Norwegian learners. These differences, or perhaps also similarities, in usage will be thoroughly investigated in the following sections, and, hopefully, the qualitative analysis of the material will shed some light on the differences these initial observations have shown us.

5.2 Analysis of *may*

5.2.1 Native speakers' use of *may*

As one might have expected, epistemic *may* accounts for the majority of instances found in the material. In fact, 395 out of the 473 occurrences of *may* are instances where *may* is used epistemically.

Table 3: Distribution of *may* according to modal type, LOCNESS

	Number of occurrences	Occurrences per 10,000 words	Distribution (per cent)
Epistemic	395	12.2	83,6%
Deontic	19	0.6	4,1%
Dynamic	53	1.6	11,0%
Ambiguous	6	0.2	1,4%
Wrong	-	-	-
Total	473	14.6	100%

Many of these occurrences are what Huddleston & Pullum (2002: 182) refer to as ‘concessive *may*’. Concessive *may*, as explained by Collins (2002: 27, 93), is used to reinforce the overall larger construction’s concessive meaning. It involves a type of pragmatic strengthening ‘in which the speaker concedes the truth of the proposition, rather than expressing a lack of confidence in it’ (Ibid: 93). As shown in example 29, concessive *may* is usually followed by *but* or the like, but *may* also be part of concessive constructions like example 30.

- (29) Sometimes they already have a good idea of what they would like to make of their lives, but it may be impractical, or at least others tell them that it is (LOCNESS-US-SCU-0006.1).
- (30) Rare as it may seem, there are females that are capable of the same crime (LOCNESS-US-SCU-0002.2).

Epistemic *may* also involves the use of intensifying adverbs, such as *well*, observed in example 31.

- (31) In fact the increase in Sporting facilities and training academies may well increase our sporting prowess in the long term, something that would be most welcomed and for which many would be happy to pay for.

Most of the instances, however, are related to whether or not the proposition of the clause is true or not. Hence, it passes some type of judgement. Interestingly, some of the instances found seem to use *may* as an objective epistemic modal, instead of the more common use of *may* as a subjective one, a use which Collins (2002: 92-93) mentions as being quite rare with regard to *may*, although not impossible.

- (32) PHS has some arguments against marijuana use by AIDS patients-mainly that it contains carcinogens and it may harm the immune system of AIDS patients (LOCNESS-US-PRB-0036.2).
- (33) I may advise him to read "How to win friends and Influence people", but to take it with a grain of salt and not to overdue the "lavish praise" (LOCNESS-US-PRB-0030.1).

Example 32 can be seen as an example of the objective use, whereas example 33 is interpreted as subjective. The objective usage of *may* usually indicates that the judgement passed is that of public record, and not necessarily limited to the speaker. The subjective use of *may*, on the other hand, involves the speaker's commitment to the proposition, which is seen in example 33. Although mere speculations, the fact that there are quite a few instances of objective use could be due to the material used for this study. As mentioned in chapter 4.1, LOCNESS consists primarily of texts written in an academic setting, and these do not usually have prominent writer visibility in terms of using the subject *I*, as noted by Aijmer (2002: 71), especially with regard to the phrase *I think*. This is not to say that there are few instances found in the material where *may* is subjective, merely that there are many instances of objective uses as well.

There are also a few instances of deontic *may*, and most of these deal with permission as seen in examples 34 and 35. These occurrences generally refer to what is permitted under certain circumstances, laws and regulations, and do not seem to include the speaker as the deontic source. Hence, most of the occurrences are objective, as often is the case stating rules and regulations.

- (34) 1. May be directed primarily against social or primarily against sexual forms of display; (LOCNESS-US-SCU-0009.3).

- (35) The result is that a pupil who has obtained the 'bac' may continue into higher education (LOCNESS-BR-SUR-0021.1).

Dynamic *may* is also present, and for the most part found in the form of possibility as well as suggestions with the use of *also*, which can be seen in example 36 below. There are also instances where *may* is used in combination with *as well*, forming *may as well*. The latter is classified by Collins (2002: 116) as a modal idiom that seems to indicate that the speaker is involved somehow, passing on his or her suggestions, although they may sound more like judgements, as seen in example 37. Both of the sentences below are examples of what Collins call 'dynamic implication' (Ibid: 116). Dynamic implication is, according to Collins (2002: 166) 'a pragmatic extension of the unreal hypothetical use, with the utterance being used to perform a speech act other than mere assertion'.

- (36) The financial hardship and the constant disappointment may also put a strain on what used to be an ideal marriage (LOCNESS-ALEV-0007.8),
- (37) As a natural cosequence of this, there would be a political loss of sovereignty, because if there is no way that British instructions can exercise their power, they may as well not exist (LOCNESS-BR-SUR-0003.3).

As seen in table 3, there were six instances in total which I was unable to identify as belonging to one single category. These are all labeled as ambiguous. Example 38 below could be interpreted as either deontic or epistemic. Either they are allowed, or permitted so to speak, to spend the night, which would make the use of *may* here deontic, or it might be a possibility that they will, which would mark *may* as epistemic.

- (38) Do they mind if people are there at all times of the night and that they may even spend the night? (LOCNESS-US-SCU-0002.2)

5.2.2 Norwegian learners' use of *may*

Epistemic *may* is by far the most commonly used modal meaning among Norwegian learners of English, as can be observed in table 4 below. Based on percentage, however, the distribution of each modal category is similar to that of native speakers. In fact, it seems as if Norwegian learners use *may* more or less in the same way native speakers do, with some minor exceptions.

Table 4: Distribution of *may* according to modal type, ICLE-NO

	Number of occurrences	Occurrences per 10,000 words	Distribution (per cent)
Epistemic	227	10.8	85,0%
Deontic	13	0.6	4,7%
Dynamic	25	1.2	9,5%
Ambiguous	3	0.1	0,9%
Wrong	-	-	-
Total	268	12.7	100%

In the interlanguage of Norwegian learners, as in the language of native speakers, epistemic modality is often used to pass judgment on whether the proposition of the clause is true or not. The instances found usually involve some sort of objective use with less speaker involvement, in many cases with *it* as a subject.

- (39) The motivation of the hired soldiers may turn out to be of very different nature (ICLE-NO 02/04/07 F 22 522 Army).
- (40) It may be that some people feel that a feminist is someone who is constantly "nagging" about women's issues , that feminists have a negative attitude towards men and that they appear little nuanced in their views (ICLE-NO 02/04/07 F 23 997 Feminists).

Sentences where *may* is interpreted concessively are quite prominent.

- (41) This is a situation that may change your life forever but you are not thinking of the consequences at this moment, the excitement drives you (ICLE-NO 02/04/07 F 27 549 Crime).

While the instances in LOCNESS are found to be able to appear in several different structures, concessive *may* in the interlanguage of Norwegian learners is generally found to appear in connection with *but*, with no other clear concessive constructions found in ICLE-NO.

Interestingly, the use of *may* together with strengthening adverbs such as *well*, however, is rare. In fact, there are no results found for *may well* at all in ICLE-NO, whereas there were seven found in LOCNESS.

Deontic use occurs throughout the corpus, such as when asking or giving permission. As with the results found in LOCNESS, these examples also deal with stating what is generally permitted by the public, regulations and laws, as seen in example 42 below.

- (42) Amended in the late sixties, the law now states that instead of serving under arms, men may do "civilian" service instead, working as part of the civilian defence (ICLE-NO 02/04/07 M 25 1191 Army).

Norwegian learner's use of dynamic *may* seems to be similar to that of native speakers, with one noticeable difference. There were no uses where the phrase *may as well* would denote dynamic possibility in terms of suggestions found in ICLE-NO. In fact, there are no instances of *may as well* found in all of ICLE-NO.

Ambiguous results are also present in the interlanguage of Norwegian learners, where *may* usually could be interpreted as either deontic or epistemic. Consider example 43 below, where it is rather unclear whether the speaker intended to give the prisoner some sort of permission to earn more money if he chooses to work in jail. It could, perhaps, be intended as epistemic possibility, in the sense that there is a possibility that this prisoner will earn more money if he decides to work in jail.

- (43) The prisoner is paid every day and may earn more money if he would like to work in jail (ICLE-NO 02/04/07 F 22 886 Prison).

5.2.3 Japanese learners' use of *may*

Japanese learners, too, frequently use *may* to express epistemic modality as opposed to other modal uses, as seen below in table 5. Compared to native speakers, there is only a slight overuse, which could imply that Japanese learners are able to use *may* idiomatically. The classifications of the other modal categories are also listed in table 5, most of them being similar to the frequencies found in both the native speaker material and the Norwegian learner material, with the exception that Japanese learners have instances where *may* is used wrongly, as opposed to Norwegian learners.

Table 5: Distribution of *may* according to modal type, ICLE-JP

	Number of occurrences	Occurrences per 10,000 words	Distribution (per cent)
Epistemic	250	12.6	85,1%
Deontic	13	0.7	4,7%
Dynamic	20	1.0	6,8%
Ambiguous	2	0.1	0,8%
Wrong	8	0.4	2,7%
Total	293	14.8	100%

However similar the results may seem, there is one main difference especially apparent when *may* is used in its epistemic sense by Japanese learners. As previously mentioned, *may* tended to be used rather objectively, i.e. the speaker had less involvement in the proposition. For the interlanguage of Japanese learners, however, the opposite seems to be the case. While both *it* and *this* are used by Japanese learners, there are far more instances found in ICLE-JP using first person personal pronouns *we* and sometimes *I*, and third person personal pronouns *he*, *she* and *they* in combination with *may*.

(44) I may work at some kind of laboratory, or go on to graduate school so that I can work at university (ICLE-JP-HI-0002.1).

(45) They may have a good life before they do awful crime (ICLE-JP-TF-0007.1).

The deontic use of *may* seems to be very similar to that of the Norwegian learner group, and as such will not be discussed in further detail. Example 46 below illustrates the use of deontic *may* found in ICLE-JP in regards to permission.

(46) “May I grasp your hands?” (ICLE-JP-SWU-0019.4)

May used to express dynamic modality seems to be used in the same manner by Japanese learners as native speakers, albeit with little variation. As with the Norwegian learners, there were no instances found for *may as well* in ICLE-JP either, which could perhaps be one reason why the occurrences of dynamic *may* are fewer in the two learner groups as compared to native speakers.

Perhaps one of the more intriguing things about the Japanese learner group, is that there are instances where *may* is used incorrectly, as seen in examples 47 and 48. In fact, there are more instances where *may* is used incorrectly than there are apparent ambiguous ones.

(47) If English was n't an international common language, I would may think that it is no need to master English for us (ICLE-JP-SWU-0025.1).

(48) I think it will may hate English (ICLE-JP-SWU-0036.1).

According to the criteria established by Palmer (1990: 4) which we explored in chapter 2.1.1, there can be no co-occurrence of modals. This is exactly what is the case in the above

examples. Other examples were instances where perhaps *may* would work perfectly fine, but sentence structure and perhaps confusion between *may*, *may be* and *maybe*, such as example 49.

(49) [...] suddenly an idea came to him; “fried noodle may be do.” (ICLE-JP-WAS-0012.4)

Now, the below example is actually quite interesting on its own. It is considered among the occurrences classified as ‘wrong’ in table 5 above, but it is particularly interesting because it might have been meant to be used as an epistemic one modal, but the use of multiple adverbs, as well as some confusion with *be* and sentence structure, actually make the use of *may* wrong.

(50) In fact, one who helps apparently may be actually helped by one who is helped seemingly, that is, they are completely equal (ICLE-JP-TM-0002.1).

The writer might have intended to use *may* here as an epistemic modal, describing how it is possible for the person who helps someone to be helped himself. The first noticeable feature of this particular example is the use of two adverbs, *apparently* and *actually*, which do not necessarily go well together. Nor does *apparently* seem to go well together with *may* in this particular instance, but *actually* could work if the structure is reorganised, i.e. ‘may actually be helped’. Depending on how the sentence is rewritten, and what the writer intended, *may* could very well be classified as epistemic in this case.

5.2.4 Comparisons

As shown in section 5.1, neither overuse nor underuse of *may* seemed to be a problem in either learner group, as the frequency results were fairly similar to those of the native speaker group. Even after dividing the research material and categorising *may* into modal categories, the distribution of *may* between each category seems to be fairly evenly distributed, as we may observe in figure 3 below.

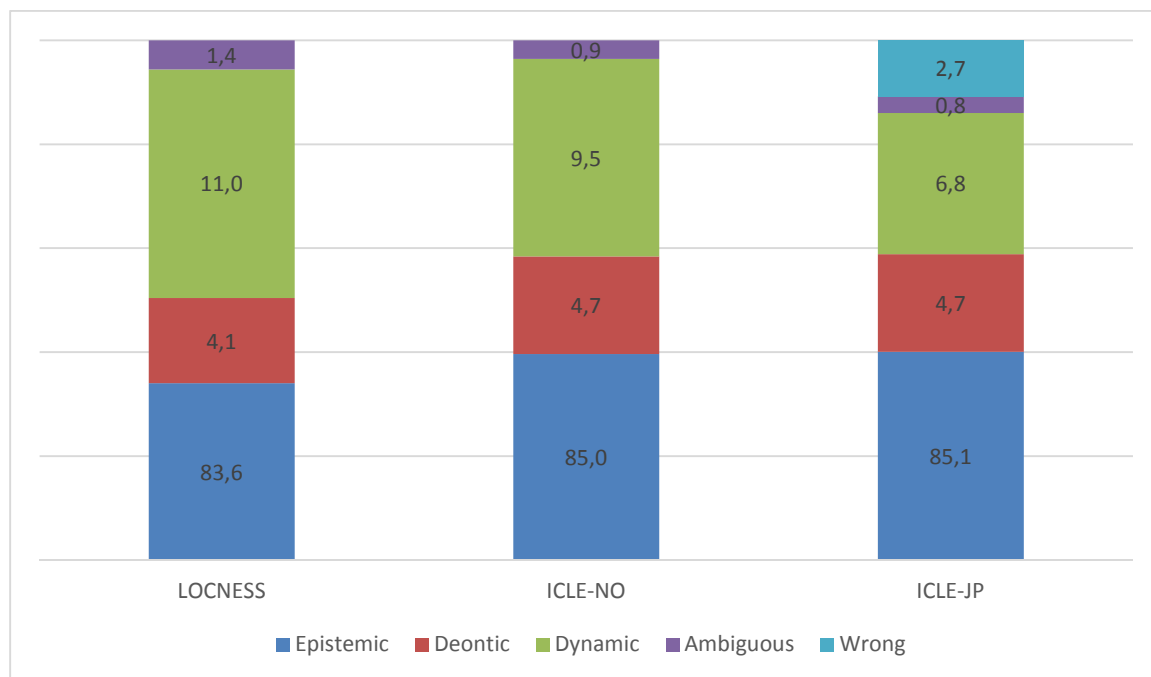


Figure 3: Distribution of *may* according to modal type (percent)

Both Norwegian and Japanese learners have a slight overuse of epistemic *may* compared to native speakers. When the percentage of use of the two learner groups is compared, however, it is surprisingly similar. I find it rather intriguing that a learner group with a mother tongue so different from English and Norwegian seems to be able to use *may* idiomatically when used in its epistemic sense, albeit with a few mishaps. As it is, in my opinion, very unlikely that there should be some kind of positive transfer in this situation considering that the use of modal auxiliaries in Japanese does not necessarily correspond to the use of English *may*, as well as Norwegian *kan*, there might be something else enabling Japanese learners to use *may* more or less as native speakers would. Whether this is teaching induced, i.e. learnt in school, or caused by other aspects of language acquisition, is perhaps impossible to say, but it is evident that there are other reasons than transfer that affect the Japanese learner group in their acquisition and usage of *may* as a modal.

One aspect of the Japanese language that might have been transferred is perhaps the subject-oriented use of epistemic *may*. The fairly equivalent Japanese auxiliary *~kamoshirenai*, which we explored in chapter 2.4.2, seems to always involve the speaker's sentiment in a much higher degree than the prominent objective use of *may* found in the interlanguage of Norwegian learners and native speakers. Since the results for epistemic *may* showed that there

are, in fact, more instances where *may* is used with personal pronouns in the interlanguage of Japanese learners compared to that of the other two groups, this could perhaps be some kind of transfer as well.

Norwegian, on the other hand, has plenty of good correspondents to *may*, all of which are thoroughly investigated by Løken (1996) in her contrastive analysis of *may* and *might*. Hence, one would probably expect that positive transfer may occur in the interlanguage of Norwegian learners of English. Indeed, there seems to be some evidence of positive transfer from the L1 in the case of Norwegian learners, it is also apparent that there could be other reasons behind Norwegian learners' ability to use *may* almost idiomatically. While Norwegian, according to *Stor norsk-engelsk ordbok* (Kunnskapsforlaget, 2014), only has one corresponding translation to *may*, namely *kan*, *kan* in Norwegian may be translated as either *may* or *can*. Hence, students have to be taught the difference, and so some specific features of *may*, such as formality, might be teaching induced. This again could lead to overgeneralization, but the results from ICLE-NO do not appear to show any bias towards certain language structures where *may* is included.

Interestingly, both learner groups also have exactly the same distribution of *may* used as a deontic modal, and there are no immediately apparent differences in use compared to that of native speakers. However, deontic *may* is slightly overused by both learner groups, as one can observe in figure 3. It would therefore seem that both learner groups have less problems with deontic *may* than the other categories. Since most of the occurrences of deontic *may* in all of the corpora do in fact deal with permission, and, as we saw in chapter 2, each language seems to have good correspondences to *may* when dealing with permission, this is perhaps expected. As such, this use of *may* could be an example of positive transfer.

It is in the dynamic use that frequencies differ the most. While Norwegian learners have a slight underuse, Japanese learners have much more of an underuse compared to the use dynamic *may* by native speakers. One striking discovery is that neither Norwegian, nor Japanese learners seem to be using *may as well* to express suggestions, whereas this is done by native speakers. Why this is the case is hard to explain. Both Japanese and Norwegian seem to have constructions that could correspond to *may as well*, *~te mo ii* (Bunt, 2006) and *kan like (så) godt* (Kunnskapsforlaget, 2014). However, the Japanese *~te mo ii* is only used when dealing with permission, and hence would probably be expected to only correspond to *may as well* when some sort of permission is given, thus expressing deontic modality, and not

as dynamic modality in the form of an idiom expressing a suggestion as seen in the native speaker material. Norwegian speakers seem to have a bias towards *might as well*, which will be further addressed in chapter 5.3.2. This bias could perhaps be one of the reasons *may as well* is not found in ICLE-NO at all.

5.3 Analysis of *might*

5.3.1 Native speakers' use of *might*

There are in total 85 occurrences of *might* found in LOCNESS, which is quite a small number compared to *may* in the same corpus. In fact, as seen in table 4, it only occurs 2.6 times per 10,000 words.

Table 6: Distribution of *might* according to modal type, LOCNESS

	Number of occurrences	Occurrences per 10,000 words	Distribution (per cent)
Epistemic	72	2.2	85,1%
Deontic	-	-	-
Dynamic	10	0.3	13,9%
Ambiguous	3	0.1	3,9%
Wrong	-	-	-
Total	85	2.6	100%

Most of these 85 occurrences of *might* are instances in which epistemic possibility is expressed.

- (51) As a British student of modern languages, having lived and worked in Paris for a year, and become acquainted with the French culture, I might be expected to have a more "European" perspective than the average British person (LOCNESS-BR-SUR-0021.3).

The above example illustrates the use of *might* as an epistemic modal expressing possibility, in the sense that the speaker in this case does not necessarily know if he will be expected to have a more 'European perspective', it is merely a possibility, a hypothesis made by the speaker.

- (52) Those people would argue that a scientist should not introduce small fragments of human DNA into a harmless bacteria, to produce hormones that some individuals are incapable of producing, because there is a slight risk of mutation or other chemicals might be produced or the technology could be used by the military or terrorists to produce new weapons (LOCNESS-ALEV-0013.8).

What seems particularly interesting is that in many cases, *might*, as an epistemic modal, alternates freely with *could*, as seen in example 52 above. This is also mentioned by Collins (2009: 113) as a possibility, as he seems to have found multiple cases in his data supporting this. While there are only a few cases in which this claim would be supported in the data in this study, it would seem that this is indeed possible, and furthermore that it is in fact done by native speakers.

However, there are also a few instances where *might* could be interpreted as expressing dynamic possibility. These are often cases in which *might* describes what is circumstantially possible, or what would sometimes be the case in terms of expressing characteristics of some sort, such as ability.

(53) When you bump other cars, a car that is hit might break down in the middle of the track and can easily get hit by another car (LOCNESS-US-SCU-0002.4).

There are also other instances of dynamic modality, for example when *might* occurs as part of a fixed expression or idiom, such as *might as well*. There are three instances in total of this use of *might* found in the LOCNESS, which are all classified as dynamic. For these three examples, the expression as a whole was being taken into consideration, as previously mentioned in connection with *may as well*. Consider the following example.

(54) [...] after all if you have to drive twenty miles to the nearest station, you might as well just drive to wherever you are going (LOCNESS-Transport 02).

It is clear that the speaker uses the whole phrase *might as well* to express a suggestion, which would mark the expression as a dynamic implication, as it would have with *may as well* explored in section 5.2.1. It would be very difficult in these cases, perhaps nearly impossible, to look at *might* as a separate constituent with a specific contribution to the clause. It could perhaps be labelled as ambiguous; it may be epistemic in the sense that the speaker informs of the possibility to drive to wherever, or deontic in the sense that permission is given for the addressee to drive wherever he or she is going. Hence, it is difficult to isolate the meaning of *might* itself as a modal and its individual contribution to the clause as a whole.

There were also other instances where *might*, often together with *also*, could be interpreted as expressing suggestions, as seen in the example below. It appears here as the speaker presents

the content of the clause to the reader as a suggestion of what might happen. This is extremely similar to *may also*, which we discussed in section 5.2.1.

(55) The young might also be chaperoned in her date (LOCNESS-US-SCU-0015.2).

Interestingly enough, there were no clear instances in which *might* was used on its own denoting deontic modality, which we saw mentioned in section 2.2.2 as a potential use. Whether or not this is evidence that the deontic use of *might* may not be very frequent in English, or if it is merely because of the material used for this study in particular, is difficult to say. However, it provides us with some interesting data when comparing native speakers' use of *might* to that of learners'.

In some cases, distinguishing between these two types of possibility turned out to be rather difficult; the 'ambiguous' category accounts for three occurrences in total, as shown in table 4. One of these occurrences can be observed in example 56 below.

(56) Should there be a fire in the house, many people might be seen furiously running about in the streets, clutching the machine tightly to their chests.

The above example could very well be an example of dynamic modality indicating a suggestion. The house is on fire, hence it would be no rare sight to have people running about, which is perhaps what the writer is suggesting. However, it may also express epistemic modality in the sense that we, or the speaker, do not know if this will be the case. It might just be a possibility due to the circumstances.

5.3.2 Norwegian learners' use of *might*

While the use of *may* among Norwegian learners of English seems to be fairly similar to that of native speakers, the use of *might* appears to be quite different. This is perhaps first observed in the differences the number of occurrences per 10,000 presented in table 2, as Norwegian learners show a great overuse of *might* compared to native speakers. The classification of modal meaning, and the occurrences of *might* as a modal auxiliary by Norwegian learners, can be seen in table 7 below. Interestingly, as we saw with *may*, the distribution of modal categories per cent is similar to that of native speakers.

Table 7: Distribution of *might* according to modal type, ICLE-NO

	Number of occurrences	Occurrences per 10,000 words	Distribution (per cent)
Epistemic	221	10.5	81,4%
Deontic	6	0.3	2,4%
Dynamic	35	1.7	13,2%
Ambiguous	7	0.3	3,0%
Wrong	2	0.1	0,8%
Total	271	12.9	100%

There is one category modal meaning for *might* that is overused in particular, namely epistemic modality. All other modal categories are relatively close to occurrences to the numbers found for native speakers, with the exception being dynamic modality for which there is a slight overuse as well.

Epistemic *might* occurs in several structures in ICLE-NO, and is used widely with both pronouns, as seen in examples 57 and 58, and existential *there*, as in examples 59.

Interestingly, the use of pronouns is far less common in the native speaker material.

- (57) She might not be educated or have any way or earning money at all (ICLE-NO 02/04/07 F 19 508 Abortion).
- (58) Maybe it is important for the people to have an army to believe in even if it is not as effective as they might think (ICLE-NO 02/04/07 F N/A 546 Army).
- (59) To a certain distinct I agree that there might be, moreover, a greater need of personal involvement in the chase for dreaming and imagination, but as one get started, I believe that there are clear advantages to be utilized (ICLE-NO 02/04/07 M N/A 471 Dreaming / imagination).

There are also quite a few instances of concessive *might*, which is used in more or less the same way as concessive *may*. The typical uses of concessive *might* can be seen in examples 58 and 59 above. There is also one case in which epistemic *may* and *might* are used in the same sentence, as in example 60 below, a use which was not found in any occurrences in ICLE-JP, and only once in LOCNESS.

- (60) Computers and machines may take over, the human kind might be enslaved by aliens, plagues may lay waste the earth but for now, having more knowledge and capability than ever- our future does not look so hopeless (ICLE-NO 02/04/07 M N/A 615 Dreaming / imagination).

Collins (2002: 111) mentions that this type of usage, where both modals express similar degrees of likelihood, is in fact a possibility. However, according to the data from all three corpora used in this study, this use seems to be extremely rare among novice writers.

While there seems to be no instances of *might* as a deontic modal in LOCNESS, there are some in ICLE-NO. The following example is one instance where *might* seems to denote deontic modality.

(61) A dealer might be arrested if he were caught (ICLE-NO 02/04/07 F 36 502 Crime).

In the example above, the speaker does not seem to pass any judgment on what ‘the dealer’ is permitted to do. Quite the contrary, the use here is based on what is generally permitted by laws and regulations, thus being rather objective.

(62) I also wish for it to stay that way, that my dream might be kept whole and untainted by disappointments (ICLE-NO 02/04/07 F 20 1775 Dreaming / Imagination).

Example 62, on the other hand, may also be seen as an example of deontic modality, as *might* here is used to express a wish. This use is generally attributed to *may*, and the use of *might* here might seem unnatural.

In the instances where *might* is classified as dynamic in table 7, the use of implied suggestions is prominent. This use is either expressed by using *might* on its own, or, more frequently found in ICLE-NO, the modal idiom *might as well*.

(63) But ,on the other hand if you actually can handle it, and find the pleasures in it, you might enjoy it (ICLE-NO 02/04/07 F 30 535 Dreaming / imagination).

(64) Breaking with all good scientific philosophy, the statement might as well claim that since so many criminals evade justice, we might as well all become criminals (ICLE-NO 02/04/07 M 25 1191 Army).

In addition to these three modal categories, there are occurrences that may be classified as either ‘ambiguous’ or ‘wrong’. The latter category might be a bit surprising, as there are no occurrences of wrong usage with *may* found in the same learner group, although there are instances found among Japanese learners. Interestingly, the rule of no co-occurrence proposed by Palmer (1990: 4), which is the main problem for Japanese learners when using *may*, is the main source of wrong uses of *might* among the Norwegian learners.

(65) What will work against one type of criminal will might encourage another possible criminal to become a criminal (ICLE-NO 02/04/07 M 27 788 Crime).

As for the ambiguous results, it is often unclear whether the speaker meant to use *might* in its epistemic or its deontic sense. Consider the following example.

(66) You might ask yourselves why (ICLE-NO 02/04/07 M 24 698 Breakfast).

As seen in example 66, it is not obvious whether the writer meant to use deontic permission, meaning that the writer gives the reader some sort of permission to ask himself why and allows the reader to do so. It may also be considered an epistemic modal, as the writer may have intended to merely state what he thinks is a possibility considering what the reader would do in a certain situation.

5.3.3 Japanese learners' use of *might*

Japanese learners do not seem to overuse *might* to the same extent as their Norwegian counterparts. There is, however, still a significant overuse of *might* compared to native speakers, with a total of 6.6 occurrences per 10,000 words in comparison to 2.6 in LOCNESS. The classification of modal meaning for these occurrences can be seen in table 8 below.

Table 8: Distribution of *might* according to modal type, ICLE-JP

	Number of occurrences	Occurrences per 10,000 words	Distribution (per cent)
Epistemic	123	6.2	93,9%
Deontic	-	-	-
Dynamic	2	0.1	1,5%
Ambiguous	3	0.2	2,3%
Wrong	2	0.1	1,5%
Total	130	6.6	100%

As was also the case with *may*, Japanese speakers frequently use *might* as an epistemic modal. In fact, almost all of the occurrences of *might* in ICLE-JP are of this type, in contrast to the at least slightly more varied use by Norwegian learners and native speakers. The cases where *might* is used as an epistemic modal are often instances where personal pronouns are used as subjects, as seen in example 67 and 68 respectively.

(67) If I couldn't use English then, I might be at a loss for an answer (ICLE-JP-SWU-0021.2).

(68) It might be safe for kids, but the problem is how to use the cell phone (ICLE-JP-SWU-004.4).

The concessive use of *might* is also present among the occurrences found in ICLE-JP, and as with Norwegian learners, this type of usage is quite common, one such instance being found in the example below.

(69) It is so difficult but I feel it might be interesting want to talk fluently (ICLE-JP-SWU-0020.4).

Clear deontic uses of *might* seem to be missing, albeit there are a few ambiguous occurrences which could be interpreted as such. It might be unclear in the following example whether or not the writer meant that it is possible for the couple to divorce as they are permitted to do so, which would be a deontic use of *might*, or if it is something that is hypothetically possible for them to do, and as such could be interpreted as epistemic.

(70) If things do not go well, they might divorce (ICLE-JP-WAS-0009.3)

Dynamic *might* is almost non-existent in the material retrieved from ICLE-JP. The two instances in which it does occur are both examples in which *might* is used in the modal idiom *might as well*, expressing some sort of suggestion.

(71) Therefore, we might as well have no pornographic pictures on the internet at all (ICLE-JP-WAS-0010.1)

(72) To answer this question, we might as well begin by deciding who should not (ICLE-JP-WAS-0007.3).

Instances classified as 'wrong' are not as high in number with *might* as they were for the same learner group with *may*. It would seem, however, that the problem of co-occurrence is in fact not present in the occurrences of *might* at all. Examples 71, 72 and 73 are therefore instances where it would seem that the writer simply chose the wrong modal auxiliary, or it seems wrong due to the structure of the sentence or lack of proper references.

- (71) Other school might did like this (ICLE-JP-SWU-0008.1).
- (72) If you break the rule, you might pay money (ICLE-JP-SWU-0004.0).
- (73) Consequently, student do not study as hard as they might (ICLE-JP-WAS-0016.3).

5.3.4 Comparisons

While *may* seems to be used fairly similar to native speakers by both the Norwegian and the Japanese learners, the same cannot be said for *might*. In addition to being overused in general, it would seem that there is an overuse of epistemic *might* in both learner groups as well, and this is seemingly most prominent in the Japanese learner group. As seen in figure 4 below, 93.9 percent of the instances where *might* is found in ICLE-JP are epistemic.

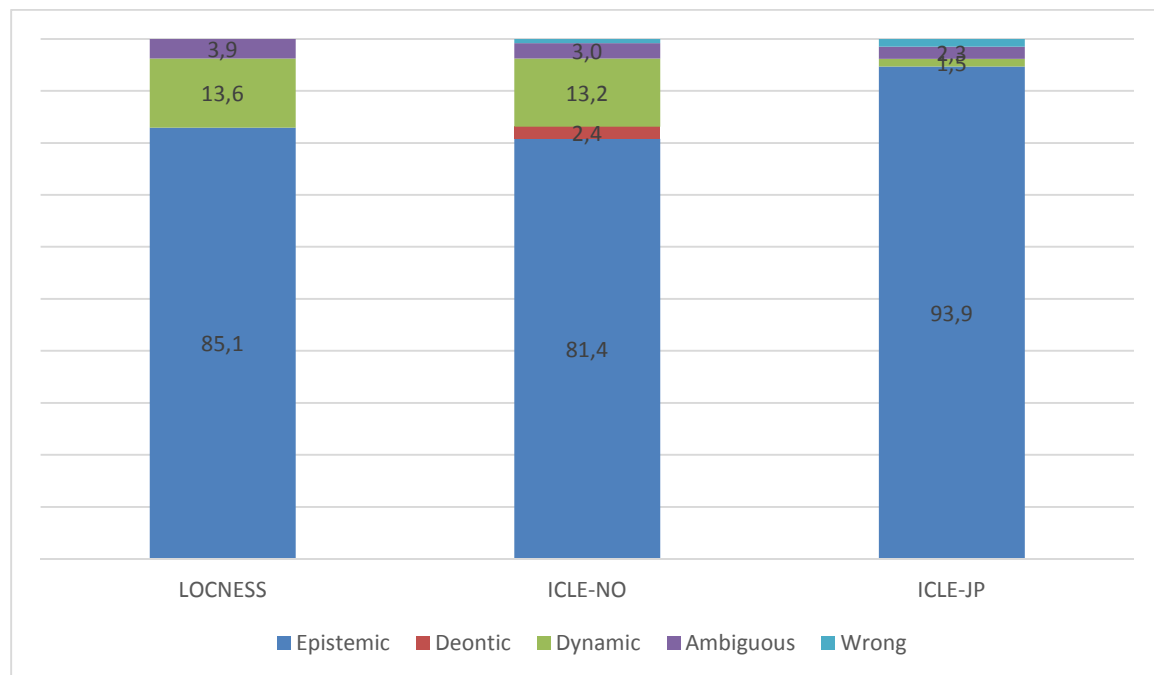


Figure 4: Distribution of *might* according to modal type (percent)

As to why this might be, I would expect the main cause to be transfer, particularly from the Japanese expression *kamoshirenai*, which we explored in section 2.4.2. *Kamoshirenai* is considered very similar to *might* in the sense that both auxiliaries express a low degree of certainty (Makino & Tsutsui, 1986: 172-175, Collins, 2002: 113). Other expressions found in Japanese used to express deontic and dynamic modality do not necessarily seem to correspond to the degree of modality *might* has in the English language, nor do they seem to

be a particularly good match. This may be one reason why the use of epistemic *might* is exceptionally high among Japanese learners, while the other uses of *might* are almost obsolete.

Norwegian learners, on the other hand, have a surprisingly high overuse of *might* overall compared to native speakers, considering the total frequency of *might* per 10,000 words found in ICLE-NO. Despite this, they seem to use *might* as an epistemic modal as much as native speakers when considering percentage of the total occurrences found, as seen in figure 4. There are, however, far more variation in the uses of *might* as an epistemic modal found in ICLE-NO than in LOCNESS and ICLE-JP. As to why there is an overuse of *might* in the first place, there may be several possible explanations. According to Løken, the most frequent correspondent to *might* is *kunne*, followed by constructions such as *kunne nok* and *kunne kanskje* (Løken, 1996: 127). She also points out that *kunne* is often translated back into English not using *might*, but instead by means of other modals or constructions. While it can, to some extent, be a question of transfer since *might* and *kunne* seem to correspond well with each other, there has to be other reasons why *might* is overused to the extent that it is.

It may also have to do with the degree of certainty expressed. In many cases where *may* could have been used, Norwegian learners opted for *might* instead. *Might*, as we know, carries an even stronger epistemic meaning than *may*. Although this is based on my personal experiences as a teacher, and is by no means representative for the learner group as a whole, Norwegian learners tend to think that the stronger the modality, the better, as one should not be too decisive in one's arguments. The reason why ICLE-NO seems to contain the only instances of *might* as a deontic modal seems to be connected to what degree of modality one would like to express as well. While native speakers seem to prefer using *may*, Norwegian learners again opt for *might* in some cases where it would perhaps have been better to use *may*. This belief might be one of the reasons as to why there is an overuse of *might*, but this claim would of course have to be investigated further, as there is no data available to support this claim.

What also surprised me was the lack of deontic uses of *might* in both LOCNESS and ICLE-JP. While it is certainly rare in ICLE-NO, it does appear, as we saw in section 5.3.2. Example 62 also illustrated *might* used when expressing a wish, a use which is generally, according to the literature consulted in this paper, listed as a possibility for deontic *may*, and not necessarily for *might*. Whether or not this is also related to the degree of modality expressed

is difficult to say. It does not, however, seem to be related to transfer, as the Norwegian *kan* is usually used when expressing wishes, which again is more equivalent to *may* than *might* (Løken, 1996: 113).

As for the dynamic use of *might*, the most striking similarity between both learner groups and native speakers is the use of *might as well* as a modal idiom to express suggestions.

Occurrences where *might* is used to make suggestions in a dynamic sense is only found in ICLE-NO. This type of dynamic use was far more common among the occurrences of *may* by native speakers, and not common at all among Japanese speakers. While Japanese learners seem to have a very limited use of dynamic *might*, there are still instances found of *might as well* in ICLE-JP. Going back to the findings in section 5.2.4, there were no instances of *may as well* found in either ICLE-NO or ICLE-JP. In LOCNESS however, *may as well* and *might as well* seem to be evenly distributed with three hits each. Hence, both Norwegian and Japanese learners seem to be biased towards using *might as well* when wanting to create sentences and modal meanings similar to that of native speakers. The bias among Norwegian learners might be explained by transfer, using the expression *kan like (så) godt*, which we briefly touched upon when examining *may as well*, as a possible source. The problem for the Japanese learners, however, would be the same here as it would be for *may as well*; there is no good correspondent available in the Japanese language containing the same sense of dynamic modality as *might as well*.

Another interesting observation made was that negation with *might not* was far more common among Norwegian learners than Japanese learners or native speakers. Likewise, the use of *may not* was higher in the two latter groups than the former. Table 9 illustrates the occurrences of *may not* and *might not* in each corpus.

Table 9: Occurrences of *may not* and *might not* per 10,000 words.

	Number of occurrences	Occurrences per 10,000 words
<i>May not</i> , LOCNESS	56	1.2
<i>May not</i> , ICLE-NO	20	0.9
<i>May not</i> , ICLE-JP	24	1.2
<i>Might not</i> , LOCNESS	4	0.1
<i>Might not</i> , ICLE-NO	26	1.2
<i>Might not</i> , ICLE-JP	11	0.6

Again, Norwegian learners seem to prefer using *might* instead of *may*, which could be due to

the apparent preference in choosing the modal with a lesser degree of certainty. It is worth noting, however, that the results for *may not* only has a slight underuse. Japanese learners, on the other hand, seem to use *may not* in the exact same frequency as that of native speakers. They do have a slight overuse of *might not* as well, but not as prominent as that of the Norwegian learner group. Again, whether or not this is due to transfer from the L1 of the two learner groups is difficult to say. Interestingly, the Japanese learners construct negation in quite a different way than that of native speakers. *Kamoshirenai* itself is not negated, but the verb preceding the auxiliary is, as it is put in the *~nai*-form to express negation. As an example, 'kare wa isshou ni ikanai no kamoshirenai' would carry the meaning of 'he may not go together (with us)' (Tanimori, 1994: 82). *Ikanai* would here mean 'not go', and *kamoshirenai*, as we have seen earlier, would correspond to *might*, but it can also, according to Tanimori, correspond to *may*, even though Makino & Tsutsui (1986: 173-175) do not list *may* as a possible translation. Hence, choosing between *may not* and *might not* could be attributed to personal preference of the students and not necessarily transfer. This, of course, is difficult to prove, but it is apparent that there are aspects other than transfer that have to be taken into consideration here as well.

6 Conclusions

This thesis has compared the use of the modal auxiliaries *may* and *might* between Norwegian and Japanese groups, as well as between learners and native speakers. It has offered a qualitative analysis of *may* and *might* with regard to which modal categories these belong to. The framework for classifying modals, as well as the main analysis of the use *may* and *might* as modal auxiliaries, is mainly based on literature written on modals and modality by Palmer (1990) and Collins (2002). The analysis was carried out using a corpus-based approach, using ICLE and LOCNESS as the sources of material. The methodical framework followed is that of CIA, originally outlined by Granger (1996), and the revised Integrated Contrastive Model by Gilquin (2000/2001). Focusing on the two modal auxiliaries, both the interlanguage produced by Norwegian and Japanese learners and the produced language by native speakers have been examined thoroughly.

Revisiting the main research questions of this study, we can safely say that the use of the modal auxiliaries *may* and *might* among Norwegian and Japanese learners of English is surprisingly similar in some respects, a fact which I did not expect when first starting to analyse the occurrences. Rather I would have expected an underuse in the case of Japanese learners, as much as I would have expected an overuse by the Norwegian learner group, as the Japanese use modal auxiliaries in a very different way syntactically than Norwegian and English, although the latter two have their differences as well. These assumptions proved to be both wrong and right. There seems to be no underuse of *may* among Japanese learners compared to that of native speakers. In fact, both groups are extremely similar in terms of frequency. There is even an overuse of *might* in the same learner group.

Furthermore, the Japanese learners' use of *may* according to the modal categories used in this thesis was extremely similar to that of native speakers. I would have expected much more variation, as Norwegian and Japanese are quite different languages compared to each other, as well as to English. However, the fact that these two learner groups seem to use *may* more or less idiomatically is actually interesting on its own, as there either has to be some kind of positive transfer or perhaps other aspects of language acquisition that enables them to do so.

As we have seen, Japanese learners may show signs of positive transfer due to the auxiliary *kamoshirenai*. However, this can be translated as both *may* and *might*, depending on individual preference perhaps. There seems to be, however, some sort of concordance

between the degree of modality one would want to express and the choice of *may* or *might*, although I cannot be absolutely certain. Another aspect I would like to mention is that the Japanese learner group is the only group in which Palmer's (1990) rule of no co-occurrence between modals is broken. This might be a problem to be aware of when dealing with Japanese learners' second language acquisition, as one could, based on the results of this study, expect these sorts of things to happen.

As for Norwegian learners, there was indeed overuse, but not for both modals. There is, in fact, a slight underuse of *may* compared to the use of native speakers and the Japanese learner group. Perhaps the most interesting part of the analysis is, however, the Norwegian learners' overuse of *might*, which is extremely high. This might also be due to transfer, which we discussed in section 5.3.4, as the use of Norwegian correspondents *kan* and *kunne* might influence the interlanguage produced by the learner. There *may*, however, also be other reasons, one of them being the possibility that the Norwegian learners prefer expressing a lesser degree of certainty in their sentences, although this is merely a hunch I personally have based on personal experiences from teaching English to Norwegians.

This overuse by Norwegian learners also means that there is a much broader use of *might* found among the occurrences in ICLE-NO. Deontic uses of *might*, which were not found in neither LOCNESS nor ICLE-JP, were present in ICLE-NO in the form of permission and expressing a wish. Even in its epistemic use, *might* seems to have a wider spectre of use among Norwegian learners than that of native speakers. These observations could be valuable for further studies of Norwegian learners and second language acquisition, as well as something to consider, perhaps in a classroom situation, where attention could be drawn to these issues.

6.1 Suggestions for further studies

Since this study was of a rather limited scope, one might consider doing a full contrastive analysis of both *may* and *might* in parallel corpora, such as the English-Norwegian Parallel Corpus or the Tanaka Corpus of Parallel Japanese-English Sentences to take full advantage of the Integrated Contrastive Model developed by Granger (1996) and Gilquin (2000/2001). While I have only shared my thoughts on the types of use that might be due to transfer from the L1, and what might be attributed to other aspects of learner language output, these

thoughts would still have to be ‘proven’. One step in uncovering the reason behind the findings of this study further could therefore be to expand it, to include a contrastive analysis for Norwegian and English, as well as Japanese and English.

Furthermore, I have only touched upon a small portion of the modal auxiliaries. I suspect that there are still many interesting aspects of both Norwegian and Japanese learner language waiting to be explored, perhaps especially with regard to Japanese, as the use of modal auxiliaries is entirely different from that of native speakers. Due to the limited time available to me, I chose to focus on *may* and *might* as my representatives of the modals, which gave me the opportunity to investigate their use to a certain depth. I have not, however, been able to give any conclusive answers, as further investigations would be needed to do so. Instead, I have illustrated the use of *may* and *might* by Norwegian and Japanese learners of English, and attempted to explain the differences in use by using the Integrated Contrastive Model as a reference. As the limited scope of this study does not allow time for a full contrastive analysis to exploit the full potential of the ICM, I would like to include this in further research.

However, I do believe that the tendencies in use explored by this study to be important and highly interesting, for both learners of English and further research. A start would perhaps be to examine *can* and *could*, since in many cases these modals can overlap with *may* and *might*. This would perhaps be an interesting study to carry out based on Norwegian learner language as well, seen in connection with Løken's (1996) study of *can*, *could*, *may* and *might*, and the Norwegian *kunne*.

6.2 Final comments

To sum up, both Norwegian and Japanese learners seem to have few problems with the use of modal auxiliaries *may* and *might*. The Norwegian learners have a good understanding of the syntactic and semantic features of these two modals, despite the fact that there are prominent differences in the use of modals in Norwegian and English. However, they were found to struggle with overuse of *might*, even though they manage to use the modal in the way native speakers do. In the case of the Japanese learners, modals were not necessarily overused nor underused, but they seem to struggle with the syntactic features, i.e. how modal auxiliaries are used in English. A few of the instances found in ICLE-JP contained co-occurrence of two modals, which is not permitted in English. These findings are well suited for further studies

on other modal auxiliaries, as well as in-depth contrastive analysis on the same modals in, for example, Japanese. It seems to me that these findings could also prove to be useful for learners of English in general, as modality and the use of modal auxiliaries is perhaps one of the more difficult aspects of language to master. I, for one, have become more aware, both as a learner of English and as a teacher of English myself, as to how *may* and *might* as modal auxiliaries are used by native speakers, compared to Norwegian and Japanese learners. Therefore, my hope is that this study will have contributed to the awareness of how modals are used by learners and native speakers alike, and that it has raised some interesting points, and perhaps questions, regarding similarities and differences between these groups.

References

Books, articles and online resources

- Aijmer, K. (2002). Modality in advanced Swedish learners' written interlanguage. In S. Granger, J. Hung, & S. Petch-Tyson, *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 55-76). Amsterdam/Philadelphia: John Benjamins.
- Bunt, J. (2006). *Oxford Beginner's Japanese Dictionary*. Oxford: Oxford University Press.
- Collins, P. (2009). *Modals and Quasi-Modals in English*. Amsterdam/New York: Rodopi.
- Eide, K. M. (2005). *Norwegian Modals (Studies in Generative Grammar)*. Berlin/New York: Mouton de Gruyter.
- Faarlund, J. T. (1991). *Morfologi: Bøyingssystemet i nynorsk og bokmål*. Oslo: Det Norske Samlaget.
- Faarlund, J. T., Lie, S., & Vannebo, K. I. (1997). *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Gilquin, G. (2000/2001). The Integrated Contrastive Model. Spicing up your data. *Languages in Contrast* 3(1), pp. 95–123.
- Granger, S. (1993). The International Corpus of Learner English. In J. Aarts, P. de Haan, & N. Oostdijk, *English Language Corpora: Design, Analysis and Exploitation* (pp. 57-69). Amsterdam: Rodopi.
- Granger, S. (1996). From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. In K. Aijmer, B. Altenberg, & M. Johansson, *Languages in Contrast. Text-based cross-linguistic studies*. Lund: Lund University Press.
- Gries, S. T. (2009). The three central corpus-linguistic methods. In *Quantitative Corpus Linguistics with R. A Practical Introduction* (pp. 7-18). London/New York: Routledge.
- Hasselgård, H., & Johansson, S. (2011). Learner corpora and contrastive interlanguage analysis. In F. Meunier, D. Sylvie, G. Gaëtanelle, & P. Magali, *A Taste for Corpora. In honour of Sylviane Granger*. (pp. 33–62). Amsterdam / Philadelphia: John Benjamins.
- Hoye, L. (1997). *Adverbs and modality in English*. London / New York: Longman.
- Huddleston, R., & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

- Ishikawa, S. (2010, September 21). Modality Expression in Interlanguage: A Study Based on Learner Corpus. *Corpus, ICT, and Language Education*, pp. 19-30.
- Johnson, Y. (2003). *Modality and the Japanese Language*. University of Michigan: The Center for Japanese Studies.
- Kunnskapsforlaget. (2014, Februar 13). *Stor engelsk ordbok*. Retrieved from Ordnett.no: <http://ordnett.no>
- Lie, S. (1976). *Innføring i norsk syntaks*. Oslo: Universitetsforlaget.
- Lie, S. (2011). *Norsk morfologi*. Oslo: Ling forlag.
- Løken, B. (1996). *Expressing possibility in English and Norwegian. Can, could, may, might and kunne and their correspondences*. Oslo: Department of British and American Studies.
- Makino, S., & Tsutsui, M. (1986). *A Dictionary of Basic Japanese Grammar*. Tokyo: The Japan Times.
- McEnery, A., & Hardie, A. (2011). *Corpus Linguistics. Method, Theory and Practice*. Cambridge: Cambridge University Press.
- Mindt, D. (1995). *An Empirical Grammar of the English Verb: Modal Verbs*. Berlin: Cornelsen.
- Narrog, H. (2009). *Modality in Japanese: The Layered Structure of the Clause and Hierarchies of Functional Categories*. Amsterdam: John Benjamins.
- Oxford University Press. (2014, Februar 7). 'May' or 'might'? Retrieved from Oxford Dictionaries: <http://www.oxforddictionaries.com/words/may-or-might>
- Palmer, F. K. (1990). *Modality and the English Modals* (2. ed.). London: Longman.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- Rønhovd, J. (1993). *Norsk morfologi*. Oslo: Gyldendal.
- Tanimori, M. (1994). *Handbook of Japanese Grammar*. Tokyo / Singapore: Tuttle.
- Tarone, E. (2006). Interlanguage. In K. Brown, *Encyclopedia of Language and Linguistics* (pp. 747–751). Boston: Elsevier.
- Tomlin, R. S. (1988, March). Basic Word Order: Functional Principles. (M. Noonan, Ed.) *Language*(64), pp. 196-197.

Corpora

British National Corpus. Davies, M. (2004). BYU-BNC (Based on the British National Corpus from Oxford University Press). Retrieved online from <http://corpus.byu.edu/bnc/>, accessed February 2014.

International Corpus of Learner English v2. Granger, S., Dagneaux, E., Meunier, F. & Paquot, M. (2009). Louvain-la-Neuve: Presses universitaires de Louvain. Available through CD-rom, accessed February 2014.

The Louvain Corpus of Native English Essays.

Centre for English Corpus Linguistics. (2005). Université catholique de Louvain. Retrieved online from <http://www.tekstlab.uio.no/cgi-bin/omc/LOCNESSsearch.cgi>, accessed February 2014.