Usage of Modal Auxiliaries in Chinese English Varieties
A Comparative Analysis

Betreuer: Prof. Dr. Josef Schmied

E-mail: 
Anschrift: 
Matrikelnummer: 
Geburtsdatum: 
Studiengang: Bachelor Anglistik/Amerikanistik
Abgabetermin: 10.12.2013
# Table of Contents

1. Introduction 1

2. Semantic Discussion 4
   2.1 Modality 4
      2.1.1 Epistemic Modality 6
      2.1.2 Deontic Modality 7
      2.1.3 Dynamic Modality 8
   2.2 Modality and Hedging 9
   2.3 Modal Auxiliaries 10
      2.3.1 Central Modals 11
      2.3.2 Can and Could 11
      2.3.3 May and Might 13
      2.3.4 Shall and Should 14
      2.3.5 Will and Would 15
      2.3.6 Must 16
      2.3.7 Other Categories of Modals 17
   2.4 Other Issues – Indeterminacy 17
   2.5 Other Issues – Categories of Modality 18

3. Aim of Research 19

4. Methodology 20
   4.1 Data 20
      4.1.1 CLEC 20
      4.1.2 ICLE 21
      4.1.3 ChemCorp 22
   4.2 Filtering the Data 22
   4.3 Classification 24
Table of Contents – cont.

5. Results 25
   5.1 Overall Distribution of Modals 25
   5.2 Distribution of Types of Modality 27

6. Discussion 29
   6.1 Discussion of the Overall Distribution of Modals 29
   6.2 Discussion of the Distribution of Types of Modality 33
   6.3 Summary of the Debate on the Chinese Culture of Politeness 35
   6.4 Summary of the Debate on Specific vs. Popular Academic Writing 36
   6.5 Limitations of this Study 37
   6.6 Recommendations for Further Research 38

7. Conclusion 39

References 41

Appendix 42

Appendix 1: Table 5.1a Overall distribution of individual modals in % 42
Appendix 2: Figure 5.1a Overall distribution of individual modals in % 42
Appendix 3: Table 5.2 Relative distribution of types of modality in % 43
Appendix 4: Table 5.1b Overall distribution of the central modals in % 43
Appendix 5: Figure 5.2a CLEC – Relative distribution of types of modality 44
Appendix 6: Figure 5.2b ICLE – Relative distribution of types of modality 44
Appendix 7: Figure 5.2c ChemCorp – Relative distribution of types of modality 45
List of Tables

2.1 The auxiliary verb – main verb scale 6
4.1.1 Composition of CLEC 22
4.2 Sample 25
4.3 Classification 25
5.1a Overall distribution of modals in % 26
5.1b Overall distribution of the central modals in % 27
5.2 Relative distribution of types of modality 28

List of Images

4.2 AntConc 24
1. Introduction

Modals in English have been discussed among scholars very intently for the last few decades, and there are still many layers of unexplored terrain regarding the matter. There are many reasons for the ongoing discussions, which can be correlated to flexibility in their usage, semantic changes over time, ambiguity of meanings and certain grammatical phenomena surrounding modals, just to name a few. Various aspects of the semantics of modals have been discussed by Quirk (1985) and Palmer in *Modality and the English Modals 2nd Edition* (1990), one of the first works to give a detailed analysis on the topic. Much of this semantic discussion comes down to a few specific areas, which have many of such ambiguous features. There are the discussions concerning the categories of modals, epistemic, deontic and dynamic modality, but also discussions regarding 'degree' of modality – possibility and necessity. Issues such as indeterminacy of certain modals, mood, and also futurity raise various questions.

Certainly, these issues are of great importance to the discussion of modals, but there is another context in which modality can be discussed. The perhaps not immediately obvious cultural aspect should not be disregarded either. In the 21st century, English is considered the most important lingua franca in the world. Thus, more groups of people are also involved in the process of language change. Kachru's model (1997) estimates 1,5 billion speakers of English, of which 750 million speak English as a second language. Their influence on English is very unique in a way, as their native language would certainly have an effect on their English - and the further their mother tongue diverges from English, the more potential exists for very unique effects on their language usage.

In this paper, the connection between this cultural aspect and the usage of modals will be examined on the example of Chinese ESL students and their usage of modals. Modal auxiliaries have the ability to 'mark' politeness, since some of them are capable of weakening statements. On some occasions, modals can also be a marker of directness or, in a certain sense, impoliteness. Now, it has to be noted that the Chinese culture is very much one of politeness, an attribute very much so characteristic of South-east Asian cultures. Yin Lu (2009: 154) states the following
about politeness:

(It is) a universal phenomenon in all societies, politeness is what people of different cultural backgrounds all try to observe for the purpose of increasing communicative competence. Being an important element in intercultural communication, politeness helps establish, maintain or consolidate harmonious interpersonal relationships, and reduces conflicts and misunderstandings.

China would be considered as a country from the expanding circle if we go by Kachru's model of the three concentric circles. (1997) This suggests that the Chinese use English as a second language, perhaps for business and intercultural communication; English is not native to them. It also means that they have no influence on existing English conventions unlike inner and outer circle countries. That does not necessarily mean, that their writing cannot be different from the norm. Whether Chinese student writing is influenced by their culture and how it manifests itself in the usage of modal auxiliaries will be the main question in of this paper.

This question can, however, not be examined without the aid of reference material, which in this study, will be German student writing. Germany is an expanding circle (Kachru 1997) country as well, but one with vastly different properties than China. English is much more established as a second language and English language teaching is at quite a high level compared to China. It would be expected, that a German's use of English is much closer to native level, compared to a Chinese person. To find out how Chinese and German students compare against each other, two corpora of student essays have been selected for analysis. The usage of modal auxiliaries in the chosen corpora by Chinese and German ESL students will be compared against each other, and the consequent results will be interpreted accordingly, while keeping in mind the cultural aspect at all times.

In addition to comparing Chinese and German student writing, there will also be some discussion regarding specific and popular academic writing. In this paper, these two categories refer to writing styles that go along with different academic sectors and are targeted to different readerships – popular academic writing is aimed at a general audience while specific academic writing is aimed at experts. Student essays, of which there are two full corpora available for this paper, provide examples of popular academic writing. The third corpus, that will be utilized for this analysis
of academic writing styles, is composed of Magister theses, which would be considered specific academic writing. This makes for an interesting comparison. The reason why it is so interesting, is that the two categories may vary in terms of the usage of modal auxiliaries, as specific academic writing uses more vague language and hedging, while popular academic writing uses more absolute terms.

This paper consists of this first introductory chapter, the second chapter will provide an overview of the aforementioned discussion on the semantics of modals. The concept of modality and the kinds of modality will also be reviewed in chapter two. The central modals of English: *can, could, may, might, shall, should, will, would* and *must* will be reviewed, just as there will be a short explanation on other modals that exist in English. Some controversial issues, regarding modals will also be discussed, namely indeterminacy and futurity. The third chapter will lay out the research methodology that was used in this study, with a focus on the three utilized corpora. The later chapters then have a focus on the results and the respective interpretations that these results bring about.

With the comparative aspect in mind, this paper seeks not to present absolute results, but rather a starting point to discuss the usage of modals in non-native English varieties and the discussion on specific and popular academic writing in the context of English as a lingua franca and Englishes around the world, particularly in countries where English is used as a second language. Such discussions are going to be held in the last conclusive chapter that will mark the end of this paper.
2. Semantic Discussion

In this chapter, I will discuss some of the established concepts and terms that surround the topic of modality. Modality, like some other grammatical categories, is not something that is unique to English, but rather a concept that is mutual among most languages. The focus of this chapter will, however, remain within the boundaries of the English language. Historically, there have been different approaches to defining modality and its different subcategories. F.R. Palmer states, that modality is not a “simple, clearly definable semantic category such as time and enumeration” (Palmer 1990: 2) yet he also points out, that certain other definitions do exist, which also seem to be appropriate. As stated before in the introductory chapter, this chapter will now review the key concepts surrounding modality, different kinds of modality, the English modals and some controversies.

2.1 Modality

Modality is one of the grammatical categories in English, albeit not one of the most frequently encountered ones such as tense, number or gender. However, since modality is a category, that bears a great deal of semantic meaning, it is of great significance to interpersonal interaction and language on many levels. In popular science, modality is often discussed as a means to express possibility and necessity – permission and obligation.

Quirk gives a more general definition of modality, which reflects certain other aspects of modality rather well. “Modality could be defined as the manner in which the meaning of a clause is qualified so as to reflect the speaker’s judgement of the likelihood of the proposition it expresses being true”. (1985: 219) Other scholars note the difference between epistemic and non-epistemic modalities (root modality).(Coates: 1983) Quirks definition mostly puts emphasis on epistemic modality. However, other categories such as deontic modality do not reflect judgement of a speaker towards truthfulness of a statement. Rather, deontic modality
is used to indicate obligation or permission and its relative 'strength'.

Another way of defining modality is presented by Huddleston and Pullum: "Modality. A kind of meaning involving non-factuality or non-assertion." (2005: 302) Interestingly, this rather brief definition covers most bases well, but it does omit some details. At this point, we have a few definitions, of which none is completely satisfactory. However, this last definition by Huddleston and Pullum, only with slightly altered wording, will still be used as the working definition for this paper: “Modality. A kind of meaning usually involving non-factuality or non-assertion.”

The insertion of 'usually' may seem slightly curious, but this is easily explained by the different kinds of modality that exist in English grammar. We have epistemic, deontic and dynamic modality, of which only the epistemic and deontic kind (with some exceptions) are properly addressed in the definitions. Dynamic modality has some other characteristics that are not addressed with 'non-actuality' or 'non-assertion'. This will be discussed in the later parts of this second chapter.

One last thing, that needs to be addressed, is how modality can be expressed in a sentence. According to Quirk, we have six categories of modals in English:

| (one verb phr.) | (a) central modals | can, could, may, might, shall, should, will, would, must |
| (two verb phr.) | (b) marginal modals | dare, need, ought to, used to |
| | (c) modal idioms | had better, would rather/sooner, have got to, etc. |
| | (d) semi-auxiliaries | have to, be about to, be able to, be bound to, be going to, be obliged to, be supposed to, etc. |
| | (e) catenatives | appear to, happen to, seem to, get + -ed participle, keep + -ing participle, etc. |
| | (f) main verb + non finite clause | hope + to-infinitive, begin + -ing participle, etc. |

Table 2.1 The auxiliary verb – main verb scale (abridged from Quirk et al 1985: 137)

The English modals are perhaps the most significant means of expressing

---

1 'strength' may refer to how absolute a modal is – modals are usually related to the judgement of a speaker, so 'strength' indicates how absolute such a judgement is
modality, but they are far from being the only option available. Modality may also be expressed through adjectives, nouns, adverbs and more:

1) *He might leave.*
2) *It is possible for him to leave.*
3) *There is a potential for him to leave.*
4) *He might possibly have left.*

Here, we have four similar sentences, in which some of the means of expressing modality can be found. In 1) we find the modal *might* as means of expressing modality. In 2), the adjective *possible* expresses modality. In 3) a noun, *potential*, indicates modality. In 4), it is the adverb *possibly* in conjunction with *might* that is used as an indicator for modality.

As already mentioned, modality can take different forms, such as epistemic, deontic and dynamic modality. I would like to delve into this distinction a little bit in the upcoming chapters and provide some examples for each category. As these categories can take on very different meanings, it is quite important to make this distinction.

2.1.1 Epistemic Modality

With this category, we also have the potentially most important one, at least for this study. According to Huddleston and Pullum's brief definition: “Epistemic modality expresses meanings relating primarily to what is necessary or possible given what we know (or believe)...” (2005: 54) This means, in other words, that epistemic modality is related to judgement and expressing judgement on whether something is or is not the case. (Palmer 1990: 50) There is a variety of modals that can express epistemic modality, but the focus in this paper shall be on the central modals, all of which are capable of expressing epistemic modality:

5) *It must have happened.*
6) *Could this be the place?*

7) *He may still be around.*

These three examples all express epistemic modality. One thing to note, is that different kinds of modals can express higher or lesser degrees of possibility or necessity, as the different modals have different “strength”. In 5), the modal *must* is used and this is actually a rather strong modal. It expresses that the speaker is relatively certain of the statement he is making. In this example, it would be that the event *It* is certain to have taken place. In 6) we find the modal *could*, the past of *can*. In this case, it expresses that something has possibly happened/can possibly happen or is possibly true, but there is some doubt involved. Here, *could* is used to seek confirmation, whether or not the speaker is at a certain location. In 7), we have the modal *may*. This modal is used, if a statement has the potential to be true, but some doubt is involved. However, there is a lesser degree of doubt in the statement than in example 6). In this example, *may* is used to express, that somebody is potentially still in close proximity.

2.1.2 Deontic Modality

With deontic modality, we have the second category of modality. It is similarly to epistemic modality, quite common, but its use differs from epistemic modality quite a bit. The meaning of deontic modals is related to permission or obligation, rather than possibility and necessity. Palmer uses the following distinction between epistemic and deontic modality: “These two uses of the modals are distinguished as 'epistemic' and 'deontic' respectively, one of them essentially making a judgement about the truth of the proposition, the other being concerned with influencing actions, states or events and expressing what Searle calls 'directives'.” (1990: 6) These kinds of 'influencing actions', states or events have to do with permission and obligation that is being laid upon the referred individual or thing. Similarly to epistemic modality, there are certain degrees of 'strength' accompanying the deontic modals:
8) *It must be done by tomorrow.*

9) *You may proceed.*

10) *They should work hard.*

In the first example 8), we find the modal *must*, which is quite 'strong' in that sense, that it indicates that something is absolutely necessary to be carried out. So in this example, the *must* indicates, that whatever *It* refers to is obligatory to be finished by tomorrow. In 9), we find the modal *may*, which is not as 'strong' as the modal from 8), so in this case, it indicates that something is permitted. Here, it is the permission for someone to continue a certain activity. In 10) we have the modal *should*, which indicates obligation, albeit not with the same 'strength' as *must* would. In the example, it is used as somewhat of a recommendation with a certain amount of vigour behind it. Interestingly enough, example 10) is slightly ambiguous, as it may also be interpreted epistemically. This would take shape in such a way, that the *should* would imply that it is relatively certain that the individuals referred to as *They* are working hard on something at this present moment. However, this kind of ambiguity can usually be deciphered from context, which is, of course, not given with this short example. Still, for an epistemic interpretation to be likely, the example would probably be: *They should be working hard.*

2.1.3 Dynamic Modality

This third form of modality is, according to Huddleston and Pullum, “[...]
concerned with properties or dispositions of persons or other entities involved in the situation.” (2005: 55) Dynamic interpretations of modals are somewhat different from epistemic and deontic interpretations, as they do not concern themselves with possibility or obligation but rather with 'states' of persons and things. Dynamic modality factors in, whenever something related to ability, power, volition, or something to that respect is expressed. This type of modality may not necessarily concern politeness directly, but it is important to this study to make this distinction.
The following examples indicate dynamic modality:

11) He can walk.

12) They were asked to help, but they wouldn't.

The first example 11) shows one of such interpretations that are concerned with properties. Here, can is used to indicate that he is capable of walking. The example, however, is unique in such a way, that it may also be interpreted deontically, so that can would indicate a permission for the person to walk. Can may often result in ambiguity, which can usually be solved by looking at the context. The second example 12) shows wouldn't as the modal, the negative form of would. In this case, it is used to indicate volition of a group of individuals. In the example, they are unwilling to help despite having been asked.

2.2 Modality and Hedging

Hedging, first recognized by Lakoff who explains them as a tool to blur the standpoint of an author toward a statement (1973), is an important tool to research articles and academic writing to prevent “absolute” statements. "Academic discourse is [...] a world of uncertainties, indirectness, and non-finality - in brief, a world where it is natural to cultivate hedges." (Mauranen 1997: 115) Importantly, modal auxiliaries such as may, might, should or must are very much capable of being a hedging device.

13) Water might be the most important resource in the world.

In example 13), we find might as a hedging device. The sentence merely suggests that water is the most important resource in the world, rather than saying that water would definitely be the most important resource in the world. Such examples show, how hedging and epistemic modality are related to one another.

2.3 Modal Auxiliaries
Modal auxiliaries, being the main concern of this paper, are of great importance as a means of expressing modality. A comprehensive list has already been given in 2.1, Fig. 3.40a, taken from Quirk's *A Comprehensive Grammar of the English Language*. The focus of this paper shall be on what he lists as 'central modals', thus other categories of modals may not be discussed in detail.

Modal auxiliaries have a distinctive set of grammatical features, which sets them apart from other verbs and more specifically auxiliaries. First of all, modal auxiliaries are capable of being used in conjunction with bare infinitives.

\[14) \text{I can go.}\]
\[15) \text{*I want go.}\]

In 14), we find the bare infinitive *go* with the auxiliary *can*, and while in this example, we have a properly grammatical sentence, we can see that 15) is not grammatical, as *want* would require a to-infinitive.

Secondly, modal auxiliaries have no non-finite forms. *Canning, *maying, or *shalling do not exist. In contrast, auxiliaries such as *be, go or have* possess non-finite forms (e.g. *being, going, having*), so modal auxiliaries are unique in that respect. Thirdly, modal auxiliaries have no s-form. *Cans, *mays or *shalls are ungrammatical. Again, in contrast to this, the s-forms of auxiliary verbs such as *be, go or have* are perfectly valid in grammatical structure.

At last, modal auxiliaries have some unique properties in terms of tense. *Could, might, should and would* are all preterite forms of *can, may, shall* and *will*, yet they are not necessarily used to indicate past tense. The phenomenon becomes clear, once we look at these two examples.

\[16) \text{You could show me the neighbourhood.}\]
\[17) \text{He tried to show me the neighbourhood.}\]

Example 16) is not past tense, even though *could* is a preterite form. Rather than that,
we find that this sentence refers to an action that may take place in the future. Example 17), though, is clearly in the past tense – the action of trying to show the neighbourhood has already been completed.

2.3.1 Central Modals

This chapter will give a brief summary on the meanings of the central modals according to Quirk. (1985: 137) These modals have some of their own special uses and abnormalities which make them a relatively tricky subject.

2.3.2 Can and Could

*Can* is a widely used modal auxiliary and, interestingly enough, is actually quite versatile in its usage. This nature is a consequence of its capability to be interpreted epistemically, deontically and dynamically, the last of which is its primary usage. Three examples for each should suffice to show this characteristic.

18) *It can happen to anyone.*

19) *You can leave.*

20) *The toddler can talk.*

In 18), we find an epistemic interpretation of *can*. In this case, the *can* indicates, that something has the possibility to happen to anyone. It is important to note that a dynamic interpretation is also possible for 18) and quite possibly is even the more likely variant. Example 19) contains a deontic *can*, indicating that someone is permitted to leave. Example 20) features a dynamic interpretation of *can*. In the example, *can* is used to show that the toddler has the ability to talk. While it is possible to assign a deontic interpretation to this example, it is still the less likely variant, considering the context. Indeed, *can* is a rather context-specific modal but it is often interpreted dynamically.
The negation of *can* leads to some rather curious phenomena. *Cannot* indeed, changes the characteristics of *can* in peculiar ways. This becomes clear, if we take the same examples, but switch them with the respective negated forms.

**21)** *It cannot be assured.*
**22)** *You cannot leave.*

The *cannot* in example 21) has two effects on the meaning of the sentence that need to be pointed out. Of course, first of all there is the negation of can, but another very interesting effect is the 'strengthening' of *can* due to the negation. If we look at an dynamic interpretation of *cannot*, the sentence would mean that something is incapable of being assured. If the sentence were “*It can be assured*”, we would have the meaning that something is capable of being assured, while in our example 21), something is incapable. In that way, the negation has given the statement a much more certain tone. In example 22), this kind of phenomenon also becomes apparent for a deontic interpretation. The negation gives the statement more severity as well, making *cannot* synonymous with “you are not allowed to”.

*Could* is the preterite form of *can*, but interestingly, it is not necessarily an indicator for the past tense. Instead, it might even be used to indicate the future tense, although *could* is also used to indicate past time in some cases. The negated forms, *could not* and *couldn’t* respectively, are not as peculiar as the negated forms of *can*. There are epistemic but also dynamic interpretations possible for this modal.

**23)** *I could help you.*

The example 23) shows the epistemic usage of *could*. *Could* is used to indicate, that it is possible that the person speaking is helpful. Additionally though, there is also a dynamic interpretation possible for the example. In that case, the sentence would mean, that the person speaking is capable of helping.

**2.3.3 May and Might**
May is yet another frequently occurring modal and it is quite similar to can in many ways. Indeed, many grammarians still ponder about the usage of may and can and whether they are interchangeable. Generally, may seems to be the more formal expression in comparison with can. The negated form of may is may not and it has to be noted that forms like *mayn't do not exist in English.

There are epistemic and deontic interpretations possible, however, unlike with can, there are no dynamic interpretations possible with may, so here we have one major difference between the two.

24) It may be too late.
25) You may leave.

The first example shows an epistemic interpretation, where may is used to indicate, that there is a possibility for something to be too late. Example 2) contains a deontic interpretation, where may is used to indicate that someone is permitted to leave. Compared to example 22) from 2.3.2, this is then a more formal version of this sentence.

Might is not as versatile as may as it only has an epistemic usage. Generally might indicates a higher level of uncertainty towards a statement when compared to could. Also, might as the preterite of may has similar anomalies to could. It may be a preterite, but it does not just refer to the past - it can also be used to refer to the present or the future.

26) It might be too late.
27) It might work itself out.
28) It might have worked.

These examples show the usage of might with different references of time. Example 26) shows the usage of might as a reference to the present. The meaning of the sentence is, that is probable to be too late for something. Example 27) shows the usage of might with reference to the future. In this sentence, might is used to indicate
that some kind of issue has the possibility to fix itself at some point in time. Example 28) contains the usage of *might* with reference to the past. *Might* is used to indicate, that something had the possibility to work but did not. One thing to note is, that *might* is interchangeable with *may* in these examples without changing the meaning.

### 2.3.4 Shall and Should

Compared to other modals, *shall* is one that is actually not very frequent in present day English any more. It does still have its use as a deontic modal, but it can also simply be an indicator for futurity. *Shall* is mostly used, if something is guaranteed to happen. The abbreviated negative form *shan't* is quite rare in present day English.

29) *None shall pass.*

*Shall* in example 29) is used as an indicator that nobody is allowed to pass. It may also be interpreted such that nobody is going to pass which would show the usage as an indicator of futurity. The first interpretation is still the more likely on, as this statement can be considered as a regulation, a way in which *shall* is still used as a standard form.

*Should*, the preterite of *shall*, is actually quite different in meaning. It has both an epistemic and a deontic interpretations, both of which are also very frequent in present day English. Yet again, despite being a preterite, this modal can indicate past as well as futurity. Its negative form is *should not*, with the abbreviated form being *shouldn't*.

30) *They shouldn't be in trouble.*

31) *You should leave.*

32) *He should be there too.*

Example 30) contains *shouldn't* as an epistemic modal. In this case, it indicates, that
someone is unlikely to be in trouble. In 31), its deontic interpretation applies - the person addressed is urged to leave. In 32), we find another epistemic interpretation. *Should* indicates, that someone is likely to be at a certain location.

2.3.5 Will and Would

*Will* has a bit of a special place among the modals, as it is the most common indicator for future tense in English. However, this comes due to the connection between modality and future time references. As there is no absolutely certain way to predict the future, the modal *will* is used to indicate that something is expected to happen in the future. Rather than a fact, future time references would thus be considered predictions, which would then mean that *will* is used as an epistemic modal. However, this very common usage of *will* is going to be further referred to as futurity usage in this paper, as it is makes more sense to give this phenomenon a category of its own in the context of this discussion. There is also a context-specific usage of *will* as a deontic modal, but it is much more rare than the epistemic variant.

33) *He will be here by noon.*

34) *It won't happen.*

35) *You will have them leave at once.*

In 33), we find the futurity usage of *will*, which in this case indicates that someone is predicted to be at a certain place at noon. 34) features the abbreviated negative form *won't*, which in this case indicates that a certain event is expected to not take place, so this is, again, the futurity variant. In 35), however, we have the deontic variant. This sentence can be interpreted as an order – the person addressed is obligated to carry it out, so this example is clearly deontic.

*Would* is the preterite form of *will* and is very different in many ways. It does not necessarily serve as a futurity indicator, but as an epistemic modal when it is used to indicate tentative meaning; in some cases deontic interpretations can be possible as well. Another form of *would* occurs, when it is used to refer to the future but in
past tense.

36) *It would function after some preparation.*

37) *There would be an uproar.*

38) *Would you help?*

39) *He told her he would do it later.*

In 36), we find the epistemic usage of *would*, indicating that something had a high probability to function, if there were some preparation. In 37), we have the epistemic usage yet again, indicating that there is a high possibility for an uproar under certain circumstances. In 38) we have the deontic usage. In this interrogative clause, we have what seems like the epistemic usage, but what the sentence actually means is not: “Is there a possibility for you to help?” Rather than that, it means roughly: “I request your help.” In 39) we have the ‘future in the past’ usage. *Would* is used to indicate that he was going to do something later during a report of an event that happened in the past at some point.

2.3.6 Must

*Must* is the last of the central modals that we are going to discuss. It is important for expressing both epistemic and deontic modality. It is also a quite 'strong' modal, as the epistemic interpretation implies a high amount of certainty and the deontic interpretation implies a rather strong obligation.

40) *It must be late, the sun is already going down.*

41) *You must listen.*

Example 40) features the epistemic usage. In this case, *must* indicates, that it has to be late, as it is already sundown. In example 41), we have the deontic interpretation, where *must* is used to indicate, that the addressed person is obligated to listen. These examples depict the aforementioned 'strength' of the modal rather well.
2.3.7 Other Categories of Modals

There are of course other categories of modals, a complete list of which can be found in 2.1. They will not be discussed in more detail in this paper, as the focus of the research lies with the central modals. It has to be noted, that one of the central modals occurs within the category of modal idioms, the modal would in would rather. It will not be included as its own category for the statistical analysis in the later chapters.

2.4 Other issues - Indeterminacy

There are some issues about the central modals that are worth mentioning. Some of them have already been mentioned before, like futurity which has been touched upon during the discussion of will. These issues need to be mentioned in this paper, as they will be relevant for the analysis of the corpora later on.

The problem of indeterminacy can occur with certain modals, but not all of them. The issue at hand is, that in certain sentences, it can be unclear, which type of modality occurs. According to Palmer (1990: 197), the three particularly problematic cases are:

[I] futurity and subject-oriented (dynamic volition) will
[II] neutral (epistemic possibility) and subject-oriented (dynamic ability) can
[III] deontic and neutral (epistemic necessity) must

For [I], we can look at one of the following example:

42) There is a party but they won't come.

From the example 42), it is not clearly definable whether the won't is one that
indicates volition or one that indicates futurity. The first interpretation would be, that there is a party but they are unwilling to come, while the second interpretation would be that there is a party but they are not going to come.

For [II], we can then look at this next example:

43) The child can walk.

In example 43), it cannot be clearly defined, whether can conveys a meaning of possibility or ability. One interpretation could be, that the child addressed has the possibility to walk. The other interpretation would be, that the child has the ability to walk, perhaps because it just recently learned it.

For [III], we can look at another example:

44) You must be quick or it will not work.

Must can either be interpreted as deontic or as means to express epistemic necessity. One interpretation can be, that the person addressed is urged to be quick for the action to work. The other interpretation is, that it is necessary for the person addressed to be quick to fulfil the criteria for the action to work.

2.5 Other Issues – Categories of Modality

With modality, there are not always clear cut explanations as to what is epistemic, deontic or dynamic. For example, there are certain aspects of dynamic modality, that could be attributed to epistemic modality. Futurity is one such category, that could be considered both epistemic or dynamic. As we can never be certain if something will happen in the future, it would make sense to say, that futurity belongs to epistemic modality, the category that is concerned with possibility. On the other hand, if the speaker of a sentence is absolutely certain that something does occur in the future and there is no doubt in his mind, then futurity
would fit in the category of dynamic modality. In this paper, we try to circumvent this issue by giving futurity its own category, avoiding confusion about this issue in the process.

3. Aim of Research

There are three major questions that will be discussed in this paper. They relate back to the topic of language usage and cultural background, which have been touched upon in the introductory chapter.

*Research Question I:* What differences occur in the usage of modal auxiliaries between Chinese and German ESL students?

*Research Question II:* How do the usage of modal auxiliaries and cultural background of Chinese ESL students affect each other?

*Research Question III:* Is the Chinese culture of politeness visible through the usage of modal auxiliaries in English?

*Research Question IV:* What kind of difference is there between popular and specific academic writing in regard to modality?

4. Methodology
This chapter will be used to present the methodology that was applied to conduct this study. The three corpora will be briefly introduced and a few samples will be presented from the data. Then, there will be a brief explanation on how the data was analysed, as there are some difficulties regarding the classification of the modals into epistemic, deontic and dynamic modality with such a large amount of samples.

4.1 Data

The data for this analysis has been taken from three corpora of learner English, which will be introduced in this chapter. The first and primary corpus for this comparative analysis is going to be the Chinese Learner English Corpus (CLEC).

4.1.1 CLEC

CLEC has been compiled during the course of an ongoing study based in Shanghai. The corpus is a collection of texts (mostly essays) from various universities across the cities of Guangzhou, Shanghai and Xinxiang. The result was a corpus of almost 1.2 million tokens, which are made up from middle school students, university minors and university majors.

For this study, we will only use the data from university students, so we will only end up with about 980 000 tokens. The distribution of the other categories of students is, relatively even, so we still have a fair distribution of different students as we can see from this table, which has been abridged from the website of the Hong Kong center for language education.
<table>
<thead>
<tr>
<th>Middle School</th>
<th>st2</th>
<th>251 558</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-majors, Band 4</td>
<td>st3</td>
<td>231 436</td>
</tr>
<tr>
<td>Non-majors, Band 6</td>
<td>st4</td>
<td>206 906</td>
</tr>
<tr>
<td>Majors, junior</td>
<td>st5</td>
<td>237 978</td>
</tr>
<tr>
<td>Majors, senior</td>
<td>st6</td>
<td>258 099</td>
</tr>
</tbody>
</table>

Table 4.1.1 – Composition of CLEC


4.1.2 ICLE

The second corpus that is used is the International Corpus of Learner English (ICLE). The updated v2 version has roughly 3,7 million tokens from ESL students with various backgrounds. (Bulgarian, Chinese, Czech, Dutch, Finnish, French, German, Italian, Japanese, Norwegian, Polish, Russian, Spanish, Swedish, Turkish and Tswana). (Granger et al.: https://www.uclouvain.be/en-277586.html: 13/10/2013)

The “interesting” data from this corpus concerns the German tokens, which ends us up with roughly about 240 000 tokens. The ICLE data comes from speakers with a relatively similar learning background to the CLEC speakers – high intermediate and advanced learners of English. The primary genre of the texts, similarly to CLEC, is also the student essay. Thus, ICLE would be a fitting choice as the corpus to which the CLEC data will be compared to, which of course is the purpose of this paper in the first place.

4.1.3 ChemCorp
ChemCorp can be considered as the reference corpus for the data from ICLE, as ChemCorp is also made up from texts by German ESL students. In this case, however, the texts are final papers by English majors, who can be considered as near native in their proficiency of English. This data will help to showcase the differences between specific and popular academic writing, as thesis papers should be of a much higher level than an essay. The corpus itself has about 100,000 tokens.

4.2 Filtering the Data

The basic approach to the study is to simply examine the corpora for the amount of modal auxiliaries. AntConc will be used to streamline the data output, as the program detects all the occurrences of a specified word within the corpus. AntConc not only gives us numbers but also displays the results in their respective context, which is particularly important for this study on modal auxiliaries. With the search results lined up in such a way, we can now examine the modal auxiliaries and which type of modality they belong to. The specification in this study only goes as far as epistemic, deontic and dynamic usage and does not make further distinctions.

As presented in 4.1, there is quite a large amount of data to go through, the CLEC corpus being the biggest hurdle with almost 1 million tokens. If we look at the results for *can* in CLEC as an example, we realize that the problem is the sheer number of results that needs to be looked at. According to AntConc, *can* has 8725 occurrences in the st3, st4, st5 and st6 files. With so much data, some restrictions have to be made, as it is near impossible to manually detect the kind of modality for every single token. Thus, only 400 results for each modal will be analysed - the first 100 results from st3, st4, st5 and st6 respectively. For the other corpora, a similar approach will be taken – the first 400 tokens for each modal will be looked over. In case there are less than 400 tokens overall, we will simply look at all the results that come up.
Image 4.2 displays a sample of the data and its view in *AntConc*. The tokens showcased in the image are from st3, which would be Chinese English minors. For the first 100 tokens of st3, we will simply look at each of them to deduct which type of modality is being communicated. For token 1, we find the usage of *can* in the following context: “We *can* get to know the world outside the [...]” In this case, *can* is used to communicate dynamic ability. We know this, because the *can* is readily substituted with “have the ability to”, which is an indication of dynamic ability. The rest of the 400 results have been analysed the same way and have then been organized in a way such as in table 4.2.
<table>
<thead>
<tr>
<th>Can/cannot</th>
<th>CLEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>8725</td>
</tr>
<tr>
<td>Epistemic</td>
<td>1/400</td>
</tr>
<tr>
<td>Deontic</td>
<td>4/400</td>
</tr>
<tr>
<td>Dynamic</td>
<td>393/400</td>
</tr>
<tr>
<td>Unclear</td>
<td>3/400</td>
</tr>
<tr>
<td>Futurity</td>
<td>0/400</td>
</tr>
</tbody>
</table>

*Table 4.2 Sample*

### 4.3 Classification

In table 4.2, we can see the categories into which the modals have been ordered. However, there are some more minor distinctions which have to be made during the classification process. In other words, we have to utilize a clear system as to how the tokens from the data can be classified. There are different approaches to certain kinds of modal meanings, like futurity as has been discussed in 2.4. Similar issues occur with meanings of volition which may be interpreted both as deontic (personal will as one's obligation) or dynamic. (personal will as an absolute constant)

A short overview of the classification categories has been given below in table 4.3.

<table>
<thead>
<tr>
<th>epistemic</th>
<th>possibility, necessity, hypothetical meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>deontic</td>
<td>permission, obligation</td>
</tr>
<tr>
<td>dynamic</td>
<td>ability, volition</td>
</tr>
<tr>
<td>unclear</td>
<td>undefinable due to error or ambiguity due to lack of context</td>
</tr>
<tr>
<td>futurity</td>
<td>referral to future events</td>
</tr>
</tbody>
</table>

*Table 4.3 Classification*

The reason for futurity having its own category is because of the special role of *will* which is used to refer to futurity almost all of the time. On the other hand, there is also *would*, which has the ability to refer to future events in the past. *Would* has epistemic, dynamic and futurity meanings in the data and instead of trying to include
futurity within the epistemic or dynamic category, we have a more clear picture of the classifications by giving futurity its own category.

5. Results

This chapter will serve to present the results of the statistical analysis that was carried out on the three corpora. The main role herein will be attributed particularly to the relative distribution, rather than to absolutes. There are certainly some observations to be made about the absolute numbers as well, but since the the aim of the study is to compare the CLEC (Chinese) and ICLE (German) corpora, relatives will be of greater significance.

5.1 Overall Distribution of Modals

The first observations can be made on the overall distribution of modals which has been calculated with the help of AntConc. Table 5.1a below shows the distribution of the central modal auxiliaries in the three corpora. CLEC and ICLE are this study's main focus, while ChemCorp serves as a reference corpus.

<table>
<thead>
<tr>
<th></th>
<th>#words</th>
<th>Can</th>
<th>Could</th>
<th>May</th>
<th>Might</th>
<th>Shall</th>
<th>Should</th>
<th>Will</th>
<th>Would</th>
<th>Must</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLEC</td>
<td>980k</td>
<td>.91</td>
<td>.08</td>
<td>.11</td>
<td>.02</td>
<td>.36</td>
<td>.51</td>
<td>.13</td>
<td>.20</td>
<td></td>
</tr>
<tr>
<td>ICLE</td>
<td>240k</td>
<td>.49</td>
<td>.15</td>
<td>.08</td>
<td>.08</td>
<td>.01</td>
<td>.18</td>
<td>.24</td>
<td>.28</td>
<td>.09</td>
</tr>
<tr>
<td>ChemC</td>
<td>100k</td>
<td>.54</td>
<td>.09</td>
<td>.08</td>
<td>.11</td>
<td>.01</td>
<td>.06</td>
<td>.27</td>
<td>.15</td>
<td>.04</td>
</tr>
</tbody>
</table>

Table 5.1a Overall distribution of individual modals in %

Can is the most frequently used modal auxiliary in all of the corpora. However, with 0.91%, it occurs almost twice as often in CLEC as it does in ICLE with .49%. The distribution in ChemCorp is quite similar to ICLE, so here we find a similarity between the two German corpora. With could however, it is the other way

---
2 corresponding figures for tables 5.1a, 5.1b and 5.2 can be found in the appendix
around. With 0.08%, it occurs half as frequently in CLEC than it does in ICLE were it occurs with 0.15%. Interestingly, the ChemCorp distribution differs from ICLE and is closer to the values in CLEC.

For may, we find a relatively similar distribution among all three corpora and with 0.11% a slightly higher percentage in CLEC than the 0.08% in the German corpora. With might on the other hand, we only find 0.02% in CLEC and almost four, and respectively five times as frequent occurrences in the German corpora.

Shall occurs very rarely among all the corpora, with only 0.02% and 0.01%. It has to be noted, that these numbers are rounded up to two digits, so there are slight differences between the German corpora. However, due to the low amount of shall-occurrences in the corpus, this should be negligible. Should has much more differing results, with a massive 0.36% in CLEC. This is much higher than the 0.18% in ICLE. Interestingly, ChemCorp has even fewer occurrences than that, with only 0.06%.

Will is quite frequent in CLEC yet again with 0.51%. In ICLE and ChemCorp, the distribution is only at 0.24% and 0.27% which is quite similar, but also much lower than CLEC. Would shows a similar phenomenon to could as CLEC and ChemCorp have a similar distribution of 0.13% and 0.15% while ICLE shows a distribution of 0.28%.

Must occurs with 0.20% in CLEC. With only 0.09% and 0.04% in ICLE and ChemCorp, the frequencies are again much lower there.

There are some interesting phenomena to be observed in this overall distribution already. The CLEC corpus has the highest percentages for all the present forms of the modals. Only in some of the preterite forms do the other corpora overtake CLEC in this overall distribution. For some of the modals such as can, will, should and must CLEC has almost twice as high of a percentage for these values. CLEC and ChemCorp, while generally being quite similar with their results, do have some discrepancies for some of the modals such as could, should, would and must.

<table>
<thead>
<tr>
<th>CLEC</th>
<th>ICLE</th>
<th>ChemCorp</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.34</td>
<td>1.6</td>
<td>1.35</td>
</tr>
</tbody>
</table>

Table 5.1b Overall distribution of the central modals in %
Another vital statistic is the distribution of the central modals overall in the three corpora. Here, we find that 2.34% of the whole CLEC corpus is made up of the central modals, while ICLE and ChemCorp have 1.6% and 1.35% respectively.

5.2 Distribution of Types of Modality

Now we can move on to the more detailed analysis with a look at the percental distribution of the different categories of modality. For reference on how this distinction was made, we can refer to 4.2 and 4.3. Table 5.2. Below is a depiction of all the individual types of modality and their respective occurrences.

<table>
<thead>
<tr>
<th></th>
<th>Can</th>
<th>Could</th>
<th>May</th>
<th>Might</th>
<th>Shall</th>
<th>Should</th>
<th>Will</th>
<th>Would</th>
<th>Must</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLEC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overa.</td>
<td>8865</td>
<td>811</td>
<td>1078</td>
<td>151</td>
<td>155</td>
<td>3558</td>
<td>4991</td>
<td>1281</td>
<td>1993</td>
</tr>
<tr>
<td>Epist.</td>
<td>.25</td>
<td>25</td>
<td>97.75</td>
<td>96.6</td>
<td>0</td>
<td>.75</td>
<td>0</td>
<td>58.25</td>
<td>7.75</td>
</tr>
<tr>
<td>Deont.</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>36.1</td>
<td>99</td>
<td>0</td>
<td>.75</td>
<td>92</td>
</tr>
<tr>
<td>Dyn.</td>
<td>98.25</td>
<td>68.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>.15</td>
<td>22</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Uncl.</td>
<td>.75</td>
<td>4.5</td>
<td>.25</td>
<td>3.4</td>
<td>1.9</td>
<td>.25</td>
<td>0</td>
<td>.25</td>
<td>.25</td>
</tr>
<tr>
<td>Fut.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>61</td>
<td>0</td>
<td>98.5</td>
<td>18.75</td>
<td>0</td>
</tr>
<tr>
<td>ICLE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overa.</td>
<td>1173</td>
<td>346</td>
<td>190</td>
<td>201</td>
<td>18</td>
<td>439</td>
<td>576</td>
<td>675</td>
<td>214</td>
</tr>
<tr>
<td>Epist.</td>
<td>3</td>
<td>8.38</td>
<td>96.85</td>
<td>98</td>
<td>0</td>
<td>7.25</td>
<td>0</td>
<td>78</td>
<td>26.63</td>
</tr>
<tr>
<td>Deont.</td>
<td>.75</td>
<td>0.28</td>
<td>1.05</td>
<td>0</td>
<td>5.55</td>
<td>91</td>
<td>0</td>
<td>.25</td>
<td>71.49</td>
</tr>
<tr>
<td>Dyn.</td>
<td>96.25</td>
<td>91.32</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6.5</td>
<td>0</td>
<td>1.86</td>
</tr>
<tr>
<td>Uncl.</td>
<td>0</td>
<td>0</td>
<td>2.10</td>
<td>2</td>
<td>0</td>
<td>0.75</td>
<td>.25</td>
<td>0</td>
<td>1.86</td>
</tr>
<tr>
<td>Fut.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>95.45</td>
<td>1</td>
<td>99.75</td>
<td>15.25</td>
<td>0</td>
</tr>
<tr>
<td>ChemCorp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overa.</td>
<td>569</td>
<td>89</td>
<td>82</td>
<td>111</td>
<td>15</td>
<td>60</td>
<td>286</td>
<td>158</td>
<td>45</td>
</tr>
<tr>
<td>Epist.</td>
<td>.75</td>
<td>20.22</td>
<td>98.78</td>
<td>100</td>
<td>0</td>
<td>11.66</td>
<td>0</td>
<td>93.03</td>
<td>28.88</td>
</tr>
<tr>
<td>Deont.</td>
<td>.25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>85</td>
<td>0</td>
<td>0.63</td>
<td>66.66</td>
<td>0</td>
</tr>
<tr>
<td>Dyn.</td>
<td>98.25</td>
<td>70.78</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.69</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Uncl.</td>
<td>1</td>
<td>8.98</td>
<td>2.22</td>
<td>0</td>
<td>0</td>
<td>3.33</td>
<td>0.69</td>
<td>1.26</td>
<td>4.44</td>
</tr>
</tbody>
</table>
Table 5.2 Relative distribution of types of modality in %

| Fut. | 0   | 0   | 0   | 0   | 100 | 0   | 98.61 | 6.32 | 0   |

For *can* it seems to be unanimous among all three corpora that the dynamic usage is by far the most frequent. There are a few instances of epistemic and deontic usage which all appear to be close to 1-3% in the three corpora. *Could* presents much more varied results as can be seen by the differences between CLEC and ICLE. The dynamic use of *could* is much less prevalent in CLEC compared to ICLE. The epistemic use of *could* on the other hand makes up almost a fourth of all the analysed tokens, so there is clearly a differentiation here. ChemCorp is closer to the results of CLEC this time around, although it has to be said that only 89 tokens were found in ChemCorp overall so the percentages may be slightly misleading.

For *may* and *might* we do not have too many differences for the percentages, which lies within the nature of these modals. They are in almost all cases epistemic modals, and this is no different in our data. The tokens are almost unanimously epistemic.

It is hard to make a clear statement about *shall* since we only have respectively 18 and 15 occurrences in the German data, which makes it difficult to compare them to the CLEC data where we at least have 155 occurrences. What can be observed is, that *shall* is usually utilized to indicate futurity in the German data, while in CLEC, it is also frequently used to show deontic meaning. *Should* is mostly used to indicate deontic meaning in all of the texts, with 99% in CLEC and 91% and 85% in the German data. In the German corpora, we do however find a reasonable amount of epistemic usage of *should* as well.

*Will* is a straight-forward case, as it can be found almost solely to indicate futurity in the corpora. Very rarely does it also indicate dynamic modality. With *would* we have the most difficult set of tokens to categorize. Data in all three corpora varies greatly. The only thing all three corpora have in common is the fact that the epistemic usage is the most frequent. However, with only 58.25% in CLEC, 78% in ICLE and 93.03% in ChemCorp, there are some notable differences. The usage of *would* to indicate futurity in the past also occurs quite often in CLEC and ICLE with 18.75% and 15.25%, although not so much in Chemcorp, where it occurs 6.32% of the time. The dynamic usage of *would* is much higher in CLEC than in ICLE and
ChemCorp though. ChemCorp does, in fact, not have any occurrences of dynamic usage at all.

*Must* has a much larger distribution of deontic usage in CLEC than in ICLE and Chemcorp. Still, it is the most common for all the corpora. Sometimes it is also used to indicate epistemic modality, but this usage is much more frequent in the German corpora.

6. Discussion

This chapter is dedicated to discuss and evaluate the findings from chapter four. The discussion will take place with regard to comparing Chinese and German student writing. The discussion questions have been presented in chapter three and will very much be the underlying questions for this sixth chapter as well.

6.1 Discussion of the Overall Distribution of Modals

*Can* is the most frequently used modal in all corpora. However, it has become apparent, that there is a massive difference between Chinese and German student writing when it comes to its usage. As it occurred almost twice as frequently in CLEC than it did in the ICLE corpus, there seems to be some kind of difference between the writing styles of Chinese and German students when it comes to their usage of *can*. Interestingly, the comparison of ChemCorp and ICLE shows no major discrepancy between the frequencies of *can* in specific and popular academic writing among these two German corpora. Thus, the differences of the usage of *can* seems to be of a cultural origin.

*Could* on the other hand is not quite as frequent in all three corpora. However, we have a frequency in ICLE that is twice as high compared to its frequency in CLEC - in German student essays, *could* seems to be much more popular. If we refer back to the differences between ICLE and ChemCorp however, we see that *could* is used much less frequently in ChemCorp, so a difference between specific and
popular academic writing seems to occur through the usage of *can* as well.

*May* has its highest frequency in CLEC with 0.11%, but it is not much lower in both German corpora where it is at about 0.08%. Still, we have another case, where Chinese and German student writing seems to be different in the frequency of modals. From the assumption based on the cultural perspective, that Chinese have an innate tendency to use more polite forms, this would make sense since *may* is a uniquely epistemic modal, expressing uncertainty of the speaker towards a statement. What is also interesting is the fact, that the German corpora have such similar distributions of this modal. Essays and Magister theses are very different in their writing styles, yet the frequency of this modal is almost the same, so this suggests, that the underlying reason for differences between Chinese and German student writing here is not necessarily related to genre, but rather to some kind of cultural phenomenon.

*Might* is quite curious, because it has an incredibly low frequency in CLEC with only 0.02% compared to the 0.08% in ICLE. ChemCorp has an even higher frequency than ICLE with 0.11%. An important observation has to be made for both *may* and *might* at this point. Both these modals are essentially epistemic in their usage and are also very important hedging devices. It would appear that Chinese students clearly prefer the use of *may* over *might* in their writing. In German student writing, we do not find a big margin between the frequencies of *may* and *might* so this seems to show a difference in the preference of certain modals as hedging devices in Chinese and German texts.

However, there is another phenomenon that is becoming apparent through the example of *may* and *might*. If we add up the frequencies of these two modals for all three corpora, we find that CLEC contains both with 0.13%, while the German corpora contain them with 0.16% and 0.19% respectively - a substantially higher percentage. In the introductory chapter, it was assumed that there would be some kind of correlation between the Chinese culture of politeness and Chinese student writing. The usage of certain hedging devices such as *may* and *might*, which arguably relate to politeness, do not seem to fully support this theory according to these data sets, despite a higher frequency of *may* in CLEC. The sum of the percentages of both paints a slightly ambiguous picture.
Moving on to *shall* we find that it is very rare overall. It occurs with 0.02% in CLEC and both German corpora contain *shall* only 0.01% of the time. The difference here is relatively marginal due to the rarity of this modal, but it still has to be noted that it occurs more frequently in Chinese student writing. This, again, is interesting, as *shall* is another modal that is used to indicate things that are relatively certain. One would expect that Chinese students would use such a modal less frequently as it is not necessarily one that could be related to this culture of politeness. *Shall* is one of the modals that, in general, seems to have become less and less frequent in native writing as well. On *shall* and other modal verbs, Neil Millar notes in an analysis of the TIME corpora that “shall, ought and must show a considerable decline in frequency.” (Millar 2009: 199) Thus, it is interesting that the Chinese have a higher frequency of *shall* in their English than German students seem to have – potentially there is room to argue that the Chinese use a more outdated form of English, as they use *shall* more often than other non-native speakers. Analysing the reasons for such a difference would, however, exceed the limits of this paper as it would require more than just the analysis of a single modal auxiliary.

With *should* we find another modal that is most regularly used in CLEC. With 0.36%, it is twice as frequent as in ICLE. In ChemCorp, we find that is used even less frequently with only 0.06%, so there is once again a visible difference between specific and popular academic writing in terms of the usage of *should*. Some interesting points for this modal regarding the cultural aspect and the debate of specific vs. popular academic texts will be discussed in 6.2, where the results on the analysis of the types of modality will be discussed.

As was mentioned in 2.3.5, *will* is a unique modal, because it is used to indicate futurity, a category that presents more of an absolute and is considerably less 'modal' than for example epistemic modals like *may* or *might*. This fact may help to shed some light on the discussion on China's culture of politeness and its reflection in student writing. CLEC contains *will* with a frequency of 0.51% while in ICLE and ChemCorp, it is represented with frequencies of 0.24% and 0.27%. In other words, this modal *will*, which is used to indicate futurity, something that is very certain to happen, is used almost twice as frequently by Chinese students than by German students. Here, we have yet another modal, that seems to go against the theory that
the Chinese culture of politeness is reflected in their English writing.

*Would* is more frequent in ICLE than in CLEC overall, with more than twice as high of a percentage of 0.28%. This would seem like another example, where this aforementioned cultural aspect is not mirrored in the usage of modals. What is also quite striking is the fact, that ChemCorp has merely half as many occurrences of this modal. This seems to signify a difference between specific and popular academic writing in regard to this modal. Some more detailed discussions for *would* will take place in 6.2.

With *must* we have another modal that is most frequent in CLEC, where it occurs with a rate of 0.20%. In ICLE it occurs with only 0.09% and in ChemCorp with just 0.04%, so there seem to be differences not only between Chinese and German student writing, but also between essays and Magister theses in German student writing. Again, more detailed analysis for this modal will take place in 6.2, as the spread of the modal meanings should provide some more interesting phenomena.

A final observation in this 6.1 part concerns Table 5.1b with the distribution of the central modals among all three corpora. It becomes quite apparent, that these central modals seem to be used much more frequently in Chinese student writing in general. This could indicate that the Chinese, generally, use the central modals much more often than German students. This proves to be another interesting factor, since studies on native Englishes have shown that many modals have been getting less and less frequent over time. This raises the question, whether the Chinese, perhaps, use an older style of writing than German students.

After the discussions in this part, we are yet to find a clear correlation of how the usage of modals in Chinese student writing reflects a Chinese culture of politeness. Rather than that, we have a slightly ambiguous picture as the analysis of *may, might, will, would* and *shall*. With the analysis of the different categories of modality in 6.2, we will try to shed some more light on these phenomena.
6.2 Discussion on the Distribution of Types of Modality

A discussion on the different types of modality and the distribution for the different modal meanings will be held in this chapter. Chapter 6.1 has left some questions open which hopefully will be answered with this more detailed analysis.

For *can*, we have seen that it occurs almost twice as frequently overall in CLEC, but the way it is used is almost the same for all three corpora. The more popular dynamic *can* is also the one with the highest frequency in the three corpora where it is in the high 90% ranges. There are some marginal differences between the German ICLE, where 3% of all the occurrences of *can* are as an epistemic modal, and the other corpora where not even 1% of *can* are epistemic. However, these 3% are still too low a value to justify any conclusions.

For *could*, we see that the dynamic meanings prevail in all corpora, but the distributions are actually quite different. In CLEC, the dynamic interpretations make up 68.5% of all of its occurrences, while 25% are allotted to the epistemic occurrences. In ICLE, these values look quite different – only approximately 8% of the tokens have epistemic meanings. This phenomenon would support the initial idea, that the Chinese politeness becomes visible in their usage of modals, as the epistemic usage of *could* usually constitutes hedging and weakening of a statement. So here, we have very different results for the usage of *could* in the same genre of writing, in this case the student essay. On the other hand, ChemCorp also has quite similar percentage values to those of China. In Magister theses, using modals as hedges is quite important, so it should come as no surprise that the epistemic usage of *could* is quite frequent. The frequency of epistemic *could* in Chinese student essays is, in this case, actually almost as high as in specific academic writing.

*May*, similarly to *can*, does not seem to be a factor in the culture and academic writing discussion, as its epistemic usage is, with almost 100% across all corpora, the dominating type. *Might* shows the same phenomenon. The more striking differences here can be found in the discussion of the overall distribution.

For *shall*, we actually find that CLEC has an interestingly high percentage of
deontic usages with 36.1%. Both German corpora, albeit in very few overall samples, do not have any occurrences of the deontic version and only use shall to indicate futurity. Such a high frequency of deontic shall is quite curious, as it is rather infrequent in native Englishes, yet here we observe it in Chinese student writing. It seems to be more and more likely, that this Chinese culture of politeness does not necessarily manifest itself in the usage of English modals, with the relatively high frequency of deontic modals that was observed thus far.

With should we can make such an observation yet again, as the deontic usage is predominant in CLEC where it is at 99%. In the German corpora, the deontic variant is relatively high as well, but still 10-15% lower. It would seem that epistemic usage of should as a hedging device is relatively popular in the Magister theses of ChemCorp with 11.66% which is certainly the highest value for the epistemic usage of should in all three corpora.

Will, as discussed before, is almost uniquely used as an indicator of futurity, and it is the same across all corpora. For would on the other hand, there are quite a few differences across the corpora. The epistemic usage in CLEC is only at 58.25% while it is at 78% in ICLE. Again, the idea that the Chinese culture of politeness would manifest itself in the modals seems to not apply here either. In the case of CLEC, we find that would is also used to indicate dynamic meaning at least 22% of the time, as well as to indicate “futurity in the past” 18.75% of the time; both categories are more frequent in CLEC than in ICLE. In ChemCorp, we find further evidence that the usage of the modals as hedging devices is more emphasized in specific academic writing compared to the popular style, as would is used as an epistemic modal 93.03% of the time.

Must also shows once again, that the politeness question is difficult to answer with modals, since it is used as a deontic modal 92% of the time in CLEC. In ICLE and ChemCorp, the deontic usage is much less frequent at around 70% and epistemic must, with around 30% appears relatively frequently, compared to the Chinese data. After talking about the individual distribution of all the types of modality, we find an even more ambiguous picture than before.
6.3 Summary of the Debate on the Chinese Culture of Politeness

Looking back at the discussion chapter, the comparison of the CLEC and ICLE corpora prove that there are some major differences in the usage of modal auxiliaries of Chinese and German students in essays. However, there are also some similarities. *Can, may, might* and *will* all occur under similar circumstances in all three corpora. *Can* is almost always used dynamically, *may* and *might* are almost always epistemic, and *will* is usually an indicator for futurity. What is different among these modals is, however, the overall frequency with which they occur. *Can, may* and *will* all occur more frequently in CLEC, while *might* occurs more frequently in ICLE. Coincidentally, *may* is the only case of an epistemic modal that occurs more frequently in the Chinese corpus. *Might* rarely occurs in CLEC, while it has quite a few incidences in ICLE.

Interestingly, the epistemic usage of the central modals with a higher degree of semantic flexibility, such as *could, should, would* and *must*, seem to be much more predominant in the German student essays of ICLE. The assumption that Chinese students would use a style that uses more epistemic modals, which consequently would mean more hedging devices, has not become apparent throughout the analysis in the last two parts of this chapter.

What the CLEC corpus does have, is a relatively high amount of deontic modals. *Should, must* and *shall* all occur much more frequently as deontic modals in CLEC than they do in ICLE. It seems that the Chinese use these deontic modals to put emphasis on existing obligations and rules that need to be followed or carried out, so the usage of deontic modals is applied in many of the essays.

It is also quite apparent, that the Chinese students prefer to use dynamic modals and modals as future markers over epistemic variants quite regularly. Modals of these types are much less 'modal', as there is not necessarily a connection to possibility, necessity or obligation, but rather to certain dispositions and factual states or situations. More importantly, they do not serve as hedging devices, but represent language of absolutes and certainty. In this way, it is not only the frequent usage of
deontic modals, but also the frequency of dynamic modals, that seems to take away from the assumption that the Chinese culture of politeness is represented in their usage of modals as hedges and politeness markers.

This is not to say, that the Chinese students use epistemically marked modals less frequently across the board. We have some instances, where the Chinese essays expose a higher frequency of epistemic modals. This becomes visible with *may*, a uniquely epistemic modal, which occurs with a higher frequency in CLEC than it does in ICLE. *Could* falls into this category as well. Despite its lower overall frequency in CLEC compared to ICLE, we find that the Chinese students prefer to use *could* as an epistemic modal more often than their German counterparts in ICLE.

Problematically, these seem to be the only two examples, where the assumptions from earlier receive any support, so we have to conclude that this study is not able to support the theory that the culture of politeness is supplemented by a higher usage of epistemic modals in Chinese student writing, since it seems to be the case that Chinese students use deontic and dynamic modals more frequently than German students.

It is, however, possible that this last phenomenon has its roots in the Chinese culture or language in some way. It has become quite clear, that deontic and dynamic modals are preferred by Chinese students in their essays, while Germans do not have this preference. Reasons for this could potentially lie with the Chinese culture in relation with their education systems and conventions. China has a stronger focus on rules and formalities, which do not receive as much attention in western countries such as Germany. With a focus on rules, it would only make sense that deontic modals occur more frequently.

### 6.4 Summary of the Debate on Specific vs. Popular Academic Writing

We do find some relatively clear results for the comparison of specific and popular academic writing. In the comparison of ICLE and ChemCorp, it becomes apparent that, due to their common country of origin, they share a lot of similarities in terms of the usage of modals. *Can, may, shall and will* occur with similar
frequencies overall, although it needs to be noted that shall is almost non-existent in both corpora. ICLE has higher overall frequencies of the remaining central modals, except for might which occurs less frequently in ICLE than in ChemCorp.

The most important observation on this discussion is, again, related to the epistemic usage of the modals. The results for the comparison of ICLE and ChemCorp show, that the epistemic usage is more popular in ChemCorp across the board. This should stem from the conventions that exist within academic writing – hedging and weakening of statements is one of the basic skill sets in this style, and since epistemic modals are a very useful hedging device, they are employed quite frequently as well. This does not necessarily hold true for student essays, which allow for more stylistic freedom – ’catchiness’ and controversy are more important in this genre, so fewer epistemic modals are employed and more deontic and dynamic modals are used instead.

6.5 Limitations of this Study

This study has compared corpora of Chinese and German English usage in relatively similar environments. While the corpora share a wide range of features, there are a few problems that occur from the contents of the texts. The CLEC and the ICLE corpora contain student essays from Chinese and German university students. While both corpora have texts of the same genre, the subject of these essays are not the same. The Chinese corpora are limited on a few selective topics, while the German ICLE corpus contains essays from what appears to be a much broader range of topics. Such a multitude of topics influences the vocabulary range of the given texts. Thus, it is not out of the question, that certain differences in the usage of modals can be attributed to the differences of the essay topics.

A second issue is presented with the sheer amount of tokens that CLEC contains. There were certain modals that frequently occurred with 1000 tokens and up to almost 9000. Analysing such a tremendous amount of tokens would simply exceed the boundaries of this study, and thus a barrier of at most 400 tokens for each modal was implemented. The downside of this is, however, a slightly less accurate
representation of the distribution of modals, which were presented in Table 5.2. With an analysis of the remaining tokens, a slightly better accuracy may be achieved.

Thirdly, this study has completely disregarded the effects, which the Chinese modal system might have on a Chinese speaker's English usage. The goal of this study was to carry out a comparative analysis based on statistics, so the implementation of a comparison between the Chinese and the English modal systems would have exceeded the scope of this paper. However, it is very certain, that a Chinese speaker's native language would also influence English as their L2. This paper has only focussed on certain cultural aspects and did not take such effects of language systems and second language acquisition into account.

The fourth point is the problem which is presented with the analysis of only one small group of modals. Outside of the central modals, there are various other means of expressing epistemic modality, some of which are presented in Table 2.1 and also manifest themselves as adjectives, adverbs, nouns, certain grammatical constructions and more. These are not included in the analysis of this study, but they might present further relevant data for this debate.

Finally, it has to be noted that no native variety was included in this study. Essentially, this study compared the Englishes of speakers, who are not “norm providing”. (Kachru: 1997) If there are non-standard phenomena, they can often even be regarded as simple error. It may be interesting to see, whether the inclusion of a native variety as a third comparative component could change the results of this study.

6.6 Recommendations for Further Research

After having analysed the central modals and their usage, it may be useful to look at other categories of modals from Quirk's classification (1985: 137). With such a study, the results of this paper could be confirmed or potentially extended into these other categories. A similar comparative analysis could be carried out.

In a similar vein, it may be interesting to compare other corpora. As suggested before, we might find different results if we consider the usage of modals among
Chinese students compared against the usage of modals among native speakers of English.

Another interesting approach may be a corpus-based analysis on the usage of modal auxiliaries by Chinese speakers with the Chinese modal system as a reference point. An analysis across Chinese and English may shed some light on certain preferences among Chinese speakers when using modals in English.

7. Conclusion

The phenomena, that influence Englishes in certain societies and cultural contexts are quite varied. With the example that presents itself with Chinese student writing in English, we find yet another case where we have a lot of variation and some seemingly unusual developments. At the beginning of this paper stood the assumption, that the English usage of Chinese speakers may be influenced by their culture and politeness and that such an influence would be visible within the category of the central modals. The data for this analysis was taken from three corpora. The two corpora that had then been compared were CLEC and ICLE, which consist of student essays from Chinese and German university students. As a reference corpus for the German corpus, ChemCorp was used, a corpus that is composed off Magister theses, which are of a considerably higher level, almost a native level, in terms of the English being used.

The comparative analysis has shown a certain level of ambiguity that was not expected. We discovered that the usage of modal auxiliaries is not necessarily a representation for the Chinese culture of politeness, but potentially uncovers certain other cultural influences. It has become apparent, that these influences manifest themselves in an increased usage of deontic and dynamic modals compared to German students. This could imply, that Chinese students and the Chinese English education put an emphasis on rules and obligations. It could also imply, that Chinese students seem to express stronger confidence in the factuality of their statements when using English in an academic context. The German student essays usually had a higher frequency of epistemic modals which implies a certain level of caution with
their statements – epistemic modals as hedging devices are quite frequent in the German data.

The question of popular vs. specific academic writing has also been discussed additionally to comparative analysis of Chinese and German ESL students. This discussion, while it may not be immediately apparent, goes along with the idea of using a reference corpus composed of Magister theses. It has to be noted that the language in such papers is quite different from student essays and since this comparative study was conducted with the help of two corpora composed of essays, the differences between the essay and the Magister thesis in regard to the usage of modal had to come up in the process. Unsurprisingly, it turned out that these differences manifest themselves in a higher frequency of epistemic modals in specific academic writing. What has also been visible is the fact that German student writing shares a lot of similarities in the usage of modals apart from the higher usage of epistemic modals in specific academic writing. Thus, ChemCorp proved to be a quite useful reference corpus, seeing how the aforementioned, similarities were actually visible in the two German corpora.

This study has tried to show, how English varieties are affected by the native environment in which they come into being. The examples from the outer circle countries (Kachru: 1997), China and Germany, have illustrated that certain features, potentially cultural conventions or characteristics of the respective first languages, shape the way of a group's English usage. In this case, it was the usage of modals among university students. It has also been apparent, that the way in which such characteristics shape the usage of certain features in non-native Englishes is affected by many factors at once and cannot be attributed to culture alone. Certainly, further research into the topic is necessary, but this study has to be viewed as a starting point, rather than a final result. Once again, it has become apparent, that non-native English varieties are quite interesting autonomous systems that present a variety of curious phenomena.
References


Appendix

Appendix 1: Table 5.1a Overall distribution of individual modals in %

<table>
<thead>
<tr>
<th></th>
<th>#words</th>
<th>Can</th>
<th>Could</th>
<th>May</th>
<th>Might</th>
<th>Shall</th>
<th>Should</th>
<th>Will</th>
<th>Would</th>
<th>Must</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLEC</td>
<td>980k</td>
<td>0.91</td>
<td>0.08</td>
<td>0.11</td>
<td>0.02</td>
<td>0.02</td>
<td>0.36</td>
<td>0.51</td>
<td>0.13</td>
<td>0.20</td>
</tr>
<tr>
<td>ICLE</td>
<td>240k</td>
<td>0.49</td>
<td>0.15</td>
<td>0.08</td>
<td>0.08</td>
<td>0.01</td>
<td>0.18</td>
<td>0.24</td>
<td>0.28</td>
<td>0.09</td>
</tr>
<tr>
<td>ChemC</td>
<td>100k</td>
<td>0.54</td>
<td>0.09</td>
<td>0.08</td>
<td>0.11</td>
<td>0.01</td>
<td>0.06</td>
<td>0.27</td>
<td>0.15</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Appendix 2: Figure 5.1a Overall distribution of individual modals in %

Appendix 3: Table 5.2 Relative distribution of types of modality in %

46
<table>
<thead>
<tr>
<th></th>
<th>Can</th>
<th>Could</th>
<th>May</th>
<th>Might</th>
<th>Shall</th>
<th>Should</th>
<th>Will</th>
<th>Would</th>
<th>Must</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CLEC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overa.</td>
<td>8865</td>
<td>811</td>
<td>1078</td>
<td>151</td>
<td>155</td>
<td>3558</td>
<td>4991</td>
<td>1281</td>
<td>1993</td>
</tr>
<tr>
<td>Epist.</td>
<td>.25</td>
<td>25</td>
<td>97.75</td>
<td>96.6</td>
<td>0</td>
<td>.75</td>
<td>0</td>
<td>58.25</td>
<td>7.75</td>
</tr>
<tr>
<td>Deont.</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>36.1</td>
<td>99</td>
<td>0</td>
<td>.75</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>Dyn.</td>
<td>98.25</td>
<td>68.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.15</td>
<td>22</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Uncl.</td>
<td>.75</td>
<td>4.5</td>
<td>.25</td>
<td>3.4</td>
<td>1.9</td>
<td>.25</td>
<td>0</td>
<td>.25</td>
<td>.25</td>
</tr>
<tr>
<td>Fut.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>61</td>
<td>0</td>
<td>98.5</td>
<td>18.75</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>ICLE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overa.</td>
<td>1173</td>
<td>346</td>
<td>190</td>
<td>201</td>
<td>18</td>
<td>439</td>
<td>576</td>
<td>675</td>
<td>214</td>
</tr>
<tr>
<td>Epist.</td>
<td>3</td>
<td>8.38</td>
<td>96.85</td>
<td>98</td>
<td>0</td>
<td>7.25</td>
<td>0</td>
<td>78</td>
<td>26.63</td>
</tr>
<tr>
<td>Deont.</td>
<td>.75</td>
<td>0.28</td>
<td>1.05</td>
<td>0</td>
<td>5.55</td>
<td>91</td>
<td>0</td>
<td>.25</td>
<td>71.49</td>
</tr>
<tr>
<td>Dyn.</td>
<td>96.25</td>
<td>91.32</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6.5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Uncl.</td>
<td>0</td>
<td>0</td>
<td>2.10</td>
<td>2</td>
<td>0</td>
<td>0.75</td>
<td>.25</td>
<td>0</td>
<td>1.86</td>
</tr>
<tr>
<td>Fut.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>95.45</td>
<td>1</td>
<td>99.75</td>
<td>15.25</td>
<td>0</td>
</tr>
<tr>
<td><strong>ChemCorp</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overa.</td>
<td>569</td>
<td>89</td>
<td>82</td>
<td>111</td>
<td>15</td>
<td>60</td>
<td>286</td>
<td>158</td>
<td>45</td>
</tr>
<tr>
<td>Epist.</td>
<td>.75</td>
<td>20.22</td>
<td>98.78</td>
<td>100</td>
<td>0</td>
<td>11.66</td>
<td>0</td>
<td>93.03</td>
<td>28.88</td>
</tr>
<tr>
<td>Deont.</td>
<td>.25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>85</td>
<td>0</td>
<td>0.63</td>
<td>66.66</td>
</tr>
<tr>
<td>Dyn.</td>
<td>98.25</td>
<td>70.78</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.69</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Uncl.</td>
<td>1</td>
<td>8.98</td>
<td>2.22</td>
<td>0</td>
<td>0</td>
<td>3.33</td>
<td>0.69</td>
<td>1.26</td>
<td>4.44</td>
</tr>
<tr>
<td>Fut.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>98.61</td>
<td>6.32</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

**Appendix 4: Table 5.1b Overall distribution of the central modals in %**

<table>
<thead>
<tr>
<th></th>
<th>CLEC</th>
<th>ICLE</th>
<th>ChemCorp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can</td>
<td>2.34</td>
<td>1.6</td>
<td>1.35</td>
</tr>
</tbody>
</table>

**Appendix 5: Figure 5.2a CLEC – Relative distribution of types of modality**
Appendix 6: Figure 5.2b ICLE – Relative distribution of types of modality

Appendix 7: Figure 5.2c ChemCorp – Relative distribution of types of modality