Title
A Cross-Sectional Investigation of the Development of Modality in English Language Learners’ Writing: A Corpus-Driven Study

Permalink
https://escholarship.org/uc/item/19z4h5h0

Journal
Issues in Applied Linguistics, 20(0)

ISSN
1050-4273

Authors
Elturki, Eman
Salsbury, Tom

Publication Date
2016

Supplemental Material
https://escholarship.org/uc/item/19z4h5h0#supplemental

Peer reviewed
A Cross-Sectional Investigation of the Development of Modality in English Language Learners’ Written Narratives: A Corpus-Driven Study

Eman Elturki
Washington State University

Tom Salsbury
Washington State University

The authors investigate development of English modality in written narratives of Arab second language (L2) English learners across six levels of English proficiency. Two hundred texts were randomly selected from each of the six levels resulting in a total of 1,200 texts. Following the concept-oriented approach (CoA) to second language acquisition (SLA), modal expressions were analyzed for frequency, type and combinations of modal auxiliaries and verbs. Results indicate that Arab learners express the concept of modality before they have target-like linguistic means to do so. Salient forms (high frequency and low variation) are most recurrent at lower levels of proficiency. The authors relate this finding to the learners’ Arabic L1 modal system. At high levels of proficiency, the data show a sudden and dramatic shift to new modal forms to distinguish modal meanings. Finally, expressions of modality become more productive and variant as learners progress in their language proficiency.

This paper looks at how the concept of modality emerges during the acquisition of English as a second/foreign language (L2) by Arab English language learners (ELLs). The analytical framework for this paper is the concept-oriented approach (CoA) used in research on second language acquisition (SLA). In this approach, researchers study the linguistic means that ELLs use to express a certain semantic concept (Bardovi-Harlig, 2007). The CoA has provided valuable insights into understanding the acquisition of temporality (Bardovi-Harlig, 1995, 1999, 2000; Dietrich, Klein & Noyau, 1995). It has also been used to study the acquisition of English modality by ELLs (Salsbury, 2000; Salsbury & Bardovi-Harlig, 2000, 2001). Modality is used to communicate many different meanings, and it is expressed through various linguistic devices. It can be grammatically conveyed through modal auxiliary verbs or lexically expressed with lexical verbs, adjectives, adverbs, nouns, or idioms. The complex and subtle nature of modal expressions in English poses challenges to ELLs.

Our paper takes a corpus-driven approach (Biber, 2010) to studying the acquisition of L2 English modality. A language learner corpus can be an invaluable source for investigating the development of such linguistic concepts in an ELL’s target language. It provides evidence of the development of grammatical concepts.
and usages across different levels of language proficiency and offers quantitative information on, for example, frequency. Corpus-based and corpus-driven research provide a “quantificational basis” (Housen, 2002, p. 78) that informs the field of SLA.

The purpose of this study is to investigate the development of English modality in Arab ELLs’ written narratives across different levels. The majority of corpus research studying learner corpora uses a homogenous corpus consisting of texts by learners who share the same L1 (Paquot & Granger, 2012). Most existing studies investigated English learners with L1s such as Chinese, German, Spanish, and Russian. There is a scarcity in the literature for learner corpora studies examining the development of linguistic concepts by Arab ELLs despite the fact that this is an increasing population studying English as an L2. According to the Institute of International Education (2013), government-sponsored students from Saudi Arabia alone constitute the fourth-largest group of international students in the United States after China, India and South Korea. Accordingly, a corpus of Arab ELLs’ narratives is analyzed. The study is organized as follows. First, an overview of the concept of modality in English and then in Arabic is presented. Then, the importance of modality to SLA is discussed followed by a brief description of the CoA, which is adopted for the method of analysis. Finally, the analysis is presented followed by a discussion of the results.

Modality in English

Modality conveys different semantic notions such as necessity, probability, possibility, ability, intention, obligation, permission, and hypotheticality. These notions are referred to as “modal meanings” (Aarts, 2011, p. 275). Narrog (2012) has distinguished between two main approaches for defining modality. The first views modality in terms of subjectivity or speaker attitudes (Lyons, 1977, 1983, 1994), and the second in terms of actuality (Papafragou, 2000), factuality (Palmer, 2001), or reality (Portner, 2009). However, it seems that there is a consensus among researchers in most of the recent literature on defining modality in terms of the former approach (Aarts, 2011; Bache, 2000; Collins, 2009; Declerck, 2011; Depraetere & Reed, 2006; Narrog, 2012). For instance, Declerck (2011) sees modality as “the phenomena that a situation is located in a nonfactual world” (p. 27). Narrog (2012) provided a similar definition as “a linguistic category referring to the factual status of a proposition” (p. 6). Depraetere and Reed (2006) offered a more specific view of modality as “a speaker’s judgment that a proposition is possibly or necessarily true or that the actualization of a situation is necessary or possible” (p. 269). In sum, despite the fact that linguistic theories of modality seem unconnected, they all have in common the notion that modality is used to express non-factual or non-actualized situations (Aarts, 2011; Collins, 2009; Narrog, 2012). Non-factuality refers to possible propositions or situations and also refers to obligations (Aarts, 2011). This research focuses on the use of modality in British English rather than
American English because the learner corpus utilized in this study has been developed by the British University in Dubai.

Expressions of Modality

Modality in English is expressed through diverse grammatical and lexical means. It is primarily conveyed through modal auxiliary verbs. It can also be expressed through “a set of periphrastic expressions” (Collins, 2009, p. 11), which includes marginal modals, modal lexical verbs, modal idioms, modal nouns, modal adjectives, modal adverbs, and hedges. Periphrastic modals might also be referred to as “quasi-modals” or “semi-modals.” This study adopts the term “periphrastic modals.” In order to be able to examine the range of modal devices that Arab ELLs use to express the concept of modality, an inventory with the expressions used to convey modality was developed. The following are the types of modal expressions and their descriptions that are examined in this study.

**Modal auxiliary verbs.** English has nine central modal auxiliary verbs: can, could, may, might, should, shall, will, would, and must. Drawing on different researchers’ classification of the modal meanings of these modal auxiliaries (Aarts, 2011; Bache, 2000; Biber, Johansson, Leech, Conrad, and Finegan, 1999; Coates, 1983; Collins, 2009), Table 1 summarizes the meaning that each modal auxiliary might carry.

**Table 1**

<table>
<thead>
<tr>
<th>Modal auxiliary verb</th>
<th>Modal meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>will/would</td>
<td>prediction, volition, intention, predisposition, obligation</td>
</tr>
<tr>
<td>can</td>
<td>possibility, ability, permission, offer, request, order</td>
</tr>
<tr>
<td>could</td>
<td>possibility, ability, permission, conditionality, hypotheticality</td>
</tr>
<tr>
<td>should</td>
<td>necessity, obligation, supposition, conditionality, hypotheticality</td>
</tr>
<tr>
<td>shall</td>
<td>obligation, intention, volition, prediction</td>
</tr>
<tr>
<td>may</td>
<td>possibility, permission, wish, malediction</td>
</tr>
<tr>
<td>might</td>
<td>possibility</td>
</tr>
<tr>
<td>must</td>
<td>obligation, necessity</td>
</tr>
</tbody>
</table>

All these modal auxiliary verbs are unmarked for tense and followed by a bare infinitive form of the verb.

**Marginal modals.** This kind of modal includes verbs such as need, dare, and ought to. The first two can be used as lexical verbs and are marked for tense (e.g., “I don’t dare to speak with him,” “She needs to call him.”). They can also be used as modal verbs and are unmarked for tense and followed by a bare infinitive verb (e.g., “She need follow the instructions,” “I dare not ask questions.”). Ought to is another marginal modal. What makes ought to not a central modal auxiliary is the fact that it is followed by the infinitive to, whereas central modal auxiliaries are
followed by a bare infinitive (Aarts, 2011). Need and ought to may express necessity or obligation, and dare can be utilized to deliver a warning or a volitional force.

**Modal lexical verbs.** This type of modal includes have/has to which expresses obligation and necessity, and be going to, used for predictions and volitions. There are various lexical verbs that can convey modality such as suggest, recommend, wish, intend, propose, advise, urge, permit, and require.

**Modal idioms.** This type of modal includes have/has got to, had better, would rather, and be to. They are considered idiomatic because they are made of a combination of words. Have/has got to can express necessity and obligation. Had better and would rather are followed by a bare infinitive. The latter conveys the meaning of necessity while the former expresses preference. Be to is marked for tense (is/am/are/was/were to) and followed by a bare infinitive. It holds the meaning of obligation, necessity, and future plans (e.g., “This is to be taken seriously.”).

**Modal nouns.** There are some nouns that might express modality such as necessity, request, intention, wish, demand, possibility, and probability.

**Modal adjectives.** Examples of modal adjectives are possible, supposed, sure, able, likely, bound, necessary, and probable.

**Modal adverbs.** There are different adverbs in English that may express modality such as possibly, probably, perhaps, surely, arguably, supposedly, necessarily, hopefully, and maybe.

**Hedges.** Hedges are defined as “expressions that qualify a statement with regard to its truth” (Aarts, 2011, p. 311). Examples of hedges are I think, I guess, I believe, sort of, and kind of.

**The importance of modality in SLA.** Like English temporality, modality — both form and function — is a fundamental component of English language learning because “[r]arely would one find a speech situation in which modals are not used” (Saeed, 2009, p. 75). ELLs need this essential concept to express various notions in speaking and writing such as hopes, desires, requests, opinions, speculations, permissions, abilities, possibilities, and dreams as well as to formulate conditions. Additionally, Hyland and Milton (1997) claimed that a large literature has shown the pragmatic importance of modality in writing as “a discoursal resource for negotiating knowledge claims and conveying a stance towards one’s propositions and readers” (p. 184).

Nevertheless, modality is considered one of the most difficult English structures for ELLs to master (Celce-Murcia & Larsen-Freeman, 1999). First, the forms of English modal verbs differ from ordinary verb forms and vary in their own structure. Modal auxiliary verbs such as will and could are followed by the base form, they have no third person singular form, their negative form is created by adding not, and their interrogative form is formed by the auxiliary verb itself. On the contrary, the modal have to has a third person singular form (has to), and it requires an auxiliary verb (do or does) to form negations and questions. A hypothetical modal structure would require a different formula (e.g., “You could have taken a taxi.”). Additionally, the various modal expressions with their different semantic
properties add to the challenge. For instance, Hyland and Milton (1997) stated that in addition to expressing doubt and certainty via will, may, and would, “at least 350 other lexical devices are also used for this purpose” (p. 185), including verbs, nouns, adjectives, and adverbs. They also added that the polypragmatic nature of modal expressions, i.e., that it “can simultaneously convey a range of different meanings,” (p. 185) adds to the complexity of mastering modality for ELLs as they sometimes fail to associate the appropriate modal that a certain situation requires.

**Modality in Arabic**

Modality in Arabic is a semantic category rather than a semantic-grammatical one as in English. It is semantically delivered because Arabic does not have a modal system equivalent to English. Therefore, it is claimed that the lack of such a system imposes difficulties on Arab ELLs in producing English modality (Saeed, 2009). Nevertheless, Arabic makes use of other lexical means to achieve modality including verbs, adverbs, adjectives, particles, prepositional phrases, and nouns. Alharbi (2002) listed some possible concepts for expressing modality in Arabic which are summarized as follows:

**Table 2**

<table>
<thead>
<tr>
<th>Arabic Modal Expressions</th>
<th>Meanings in English</th>
</tr>
</thead>
<tbody>
<tr>
<td>yureed/araad</td>
<td>want/wanted/would like</td>
</tr>
<tr>
<td>yastatee’/istataa’/yqdir/qadir</td>
<td>can/could/be able to</td>
</tr>
<tr>
<td>yimkin/mumkin/yuhtamal/muhtamal/qad</td>
<td>may/might/could/be probable</td>
</tr>
<tr>
<td>jaayiz/yajuuz/yastaeel/mustaheel</td>
<td>can/could be/possible/impossible</td>
</tr>
<tr>
<td>sawfa/sa- (prefix)</td>
<td>will/shall/be going to</td>
</tr>
<tr>
<td>yanbaghi/yajib/labud</td>
<td>should/must/ought to/be obliged to/have to</td>
</tr>
<tr>
<td>yalzam/laazim</td>
<td>be obliged to/should/have to</td>
</tr>
</tbody>
</table>

Additionally, Holes (2004) clarified that “mood and modality are intimately bound” (p. 423) in Arabic because modality can be expressed through inflectional marking of verbs (i.e., morphological markers). For instance, the prefix sa- might be attached to a verb to express intentions and futurity. Moreover, when the particle ‘qad’ precedes a verb, “it generally implies the modality of doubt, uncertainty, [and] possibility” (Kinberg, N., Kinberg, L., & Versteegh, 2001, p. 123). Holes (2004) also illustrated that modality in Arabic covers different categories such as possibility, obligation, and ability and can further be subdivided into deontic and epistemic modality.
The Concept-Oriented Approach to SLA

This study uses the CoA (Bardovi-Harlig, 2007; von Stutterheim & Klein, 1987) to investigate the concept of modality in SLA. It is also referred to as the semantic-oriented approach or the meaning-oriented approach. Researchers using concept-oriented analyses “are interested in the linguistic devices that speakers use to express a particular concept” (Bardovi-Harlig, 2007, p. 62). This approach proposes that, unlike in the acquisition of a first language, adult learners come to an additional language with fully developed semantic concepts. Their task is to develop the linguistic means of expression in a new language “to get the message across” (von Stutterheim & Carroll, 2012, p. 110). Within this approach, certain functions that learners want to deliver are mapped to the form they need to express. Accordingly, Gass and Selinker (2008) define the CoA as “an approach that maps language functions that a learner wants to express to the form that she or he needs to express it” (p. 515). As learners advance in the target language over time, various linguistic forms emerge. They also become more frequent and more productive over time (Author 2 & Bardovi-Harlig, 2000, 2001).

The CoA has been used to investigate the acquisition of temporality (Bardovi-Harlig, 1992, 1994, 1999; 2000, Klein, 1995) in different target languages such as English, Dutch, German, French, and Italian. These studies have been mainly longitudinal, ranging from 3 months to 3 years and involving 2 to 16 participants, usually through oral and written personal narratives, film retell tasks, and/or guided conversations. In addition to past temporality, the CoA has been used to study the emergence of lexical futures (Bardovi-Harlig, 2005) and of modality and unreal conditionals (Author 2, 2000).

Much of the past work within the CoA has been longitudinal in nature with a small number of participants, usually from different L1s. The small sample size raises the question of generalizability. The present research investigates the development of English modality by ELLs, specifically Arab ELLs. This is achieved cross-sectionally, as opposed to longitudinally, through a corpus comprised of 1200 texts. The cross-sectional research design on a single language group (Arabic) promises to be a valuable source for understanding the development of the complex concept of modality. Within the CoA, it is assumed that adult Arab ELLs have access to semantic concepts from their L1 linguistic and cognitive experience (Bardovi-Harlig, 2007).

The prediction in this paper is that given the lack of an equivalent English modal system in Arabic, Arab ELLs initially seek salient forms, which “are those that are most noticeable; learners tend to notice forms that are salient” (VanPatten, Williams, & Rott, 2004, p. 17). In regards to the expression of present and past modal meanings, learners should first rely on the most frequent modal auxiliaries: will followed by can in conversational English and can followed by will in academic prose (Biber et al., 1999). As learners advance in their language proficiency, they should begin to distinguish between forms to deliver specific modal meanings. This research explores how Arab learners express semantic concepts in L2 English. The
exploration of just one language group highlights unique patterns across different levels. Although adult Arab learners have the concept of modality from their L1, at low levels of proficiency they do not have sufficient linguistic means to express this concept in English.

**Methodology**

The purpose of the present cross-sectional study is to explore the development and emergence of English modality in Arab ELLs’ written discourse across six levels of English proficiency. This was achieved through a language learner corpus comprised of Arab ELLs’ writing. It specifically examined the types and frequencies of the modal devices used across the six levels to express the concept of modality.

**Research Questions**

The study was guided by the following research questions:

1. Which modal devices are used in each level to express the concept of modality?
2. What characterizes the expressions of modality at different levels of L2 English proficiency among Arab ELLs?
3. What stages of development can be discerned in terms of the acquisition of English modality by Arab ELLs?

**BALC Corpus**

The data of this research is obtained from the BUiD Arab Learner Corpus (BALC), a corpus of Arab ELLs built by the British University in Dubai (BUiD) in the United Arab Emirates (UAE). The corpus contains 1,865 texts, which make up 287,227 word tokens and 20,275 word types. These texts consist of essays composed by last year high school students for their Common Educational Proficiency Assessment (CEPA), which is a required pre-university national test produced and assessed by the UAE Ministry of Higher Education and Scientific Research, as well as essays written by first-year university students.

In order to achieve the aim of this cross-sectional investigation, the analyses are based on the CEPA tests because they had been grouped into six levels of language proficiency. Based on BUiD’s description of the corpus, the majority of the CEPA texts were attained from the UAE’s Ministry of Higher Education and Scientific Research. The CEPA test consists of English and math exams and is used for placement and admission purposes. The CEPA-English has three components; Grammar and Vocabulary, Reading, and Writing (information obtained from the CEPA’s webpage). The CEPA-writing test takers are provided with one writing prompt on a specific topic and offered 30 minutes to complete the writing task. The types of the CEPA essays included in this corpus can be classified as mainly narrative (e.g., writing about the worst vacation ever or a perfect holiday) and descriptive
(e.g., describing a favorite possession, person, or movie) in nature. The essays are assessed according to specific criteria looking at content, grammar, vocabulary, and spelling, and then a holistic score from 1 to 6 is assigned to the essay. A score of 1 indicates a low level of English proficiency and 6 a high proficiency level. Consequently, essays with the same score are grouped together forming 6 levels of language proficiency.

The BALC corpus is quite possibly the first corpus comprised entirely of Arab ELL writing. It has been developed for conducting research on this population. It has been successfully utilized by Randall and Groom (2009) to explore the acquisition of L2 English spelling by Arab students and eventually to develop materials for teaching common spelling patterns.

Data Analysis Procedures

Two hundred texts were randomly selected from each of the six CEPA’s folders resulting into a total of 1,200 texts and 175,145 words. Table 3 provides a summary of the total number of words in each level.

Table 3

<table>
<thead>
<tr>
<th>CEPA</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
<th>Level 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total words</td>
<td>7,762</td>
<td>17,659</td>
<td>24,626</td>
<td>36,700</td>
<td>43,166</td>
<td>45,232</td>
</tr>
</tbody>
</table>

In preparation for the analysis of the learner corpus, all the texts were spell-checked in order to facilitate the search for modal devices with concordance software. Replacing errors by reconstructing, than classifying them by tags, is considered a fast and reliable approach for annotating a language learner corpus because the chance of misclassifying errors is eliminated (Fitzpatrick & Seegmiller, 2004). Additionally, some texts that contained codes for spelling from previous research (Randall & Groom, 2009) were removed as well. Unintelligible utterances were replaced by a specific code. In order to confirm that the identified modal expressions carry modal meaning, another coder in addition to the researcher looked independently at 10% of modal expressions from each level and from a variety of different modal expression types. Inter-rater reliability resulted in 99% agreement. This procedure was undertaken as a check on the coding scheme for modal expression.

The concordance software program WordSmith Tools version 5 (Scott, 2008) was used to search and sort modal expressions in the data. We created inventories of modal expressions and calculated frequency scores using normalized frequencies across the six levels. Normalized frequency was calculated by dividing the number of occurrences of a modal expression by the total number of words in a particular level. The result was multiplied by an arbitrary value (2,000 for this study) so that whole numbers and not decimals or very small numbers were used for the analysis. The value used usually depends on the corpus size (Römer & Wulff, 2010). To illustrate the calculation in Level 6, the total number of modal auxiliary
verbs (544) was divided by 45,232, which is the total number of words in the level. Then the resulting value (0.012026) was multiplied by 2,000, yielding 24.1. Gries (2010) stressed the importance of this procedure because corpus frequencies can be compared or used to draw conclusions on the phenomenon under investigation only when the frequencies have been normalized.

Findings and Discussion

We begin by answering our first two research questions around the types of modal auxiliary verbs and periphrastic modals utilized in each level and their frequency of occurrences. Then our third research question regarding stages of development is addressed.

Types of Modals and Frequency of Occurrences

We created inventories of modal auxiliary verbs and periphrastic modals used across the six levels and their total frequencies of occurrence. Then, the token frequencies across the six levels were normalized because the texts were not equal in length. The normalized frequencies for modal auxiliaries are presented graphically and clustered into three groups (Levels 1 and 2, Levels 3 and 4, Levels 5 and 6). Normalizing the frequencies and presenting the data in clusters highlights trends in the data.

Figure 1. Normalized frequencies of modal auxiliary verbs across the six levels.
First, modal auxiliary verbs were used across the six levels as shown in Figure 1. Shall, may, might, and must occur minimally, and should occurred equally across the levels. These frequency patterns are not surprising due to the narrative nature of the writing task (essays about past experiences, favorite vacations, or favorite possessions or hobbies). Therefore, the analyses will focus more on can, will, would, and could. Can and will were the most frequent in all the levels. Can is disproportionately most frequent in Levels 1 and 2. Can and will are also the most frequent modal auxiliaries in the highest two levels, but only marginally more frequent than would and could. Learners relied more on can and will to express the concept of modality in lower levels.

This first finding can be attributed to the saliency of can and will. These forms exhibit high frequency and low variability (Biber et al., 2009), which are two requisites for high saliency (Boyd & Goldberg, 2009). These salient forms are usable by beginning ELLs to express present or past modal meanings. In addition to saliency, the fact that Arabic expresses modality via lexical devices adds to the attractiveness of these forms that appear at early exposure to be lexical (can for ability, will for future). Thus, L1 semantics and saliency contribute to Arab ELLs in lower levels employing these forms to express this modal meaning.

As Arab ELLs advance in their English proficiency, the emergence of new forms (could and would) does not follow gradual or linear patterns, as illustrated in Figure 1. The normed frequency of would and could in lower and intermediate levels is similar. However, there is a sudden jump in frequency patterns in upper levels whereby Arab ELLs suddenly distinguish between different modal meanings using appropriate modal devices. This pattern of usage, like so many phenomena in language learning, is suggestive of chaotic, complex, and non-linear development (e.g. Chaos/Complexity theory. See Larsen-Freeman, 1997, 2007).
Second, a dramatic increase in periphrastic expressions of modality (all other expressions excluding modal auxiliaries) as proficiency level increased is observed. This is illustrated in Figure 2. Level 1 learners used just 13 types and Level 2 learners only 14 (be going to, want to, think, if, love to, like to, hope, believe, have to, suggest, need to, sure, allow). In contrast, Level 6 learners used over 40 different types of modal expressions. Normalized frequency scores presented in Figure 2 illustrate that in addition to a greater variety of modalized forms, the advanced learners also used more frequent expressions of modality in comparison to the lower level learners. In lower levels, the learners attempted to express modal meanings using the linguistic means at their disposal such as lexical verbs (BE going to, want to, think, like to) in addition to a limited number of modal auxiliaries. As they progress in their language proficiency, a variety of new periphrastic modal expressions emerges in addition to a greater variety of modal auxiliaries to convey different notions of this semantic concept.

Learners across the different levels sought multiple ways to convey different modal notions through the use of periphrastic modals. The expressions of periphrastic modals become more frequent and diverse as learners advance in their language proficiency (e.g. 13 periphrastic modal types in Level 1 compared to over 40 periphrastic modal types in Level 6). Learners in lower levels rely more on modal auxiliaries than periphrastic modals, as illustrated in Figure 3. At higher proficiency levels, learners’ vocabularies have developed. So too has the means to express modality with a range of linguistic devices such as modal adverbs, modal adjectives, modal nouns, and hedges. Thus, whereas the written discourse of learners in Levels 1 and 2 is dominated by the primary modal auxiliaries, can and will (Author 2, 2000), The written discourse of learners in Levels 5 and 6 is dominated by periphrastic modals (see Figure 3).

Figure 3. Total occurrences of periphrastic versus auxiliary modals across the six levels
Characteristics of the Modal Expressions

Expressing past meanings via primary auxiliaries. Can and will are the most frequent modal auxiliaries in lower and intermediate levels. These modal auxiliaries assume several functions in the learner data, including past time reference. Perkins (1983) distinguishes primary modal auxiliaries, such as will and can, from secondary modal auxiliaries such as would and could. Although Perkins did not use the terms primary and secondary in regards to acquisitional sequences, lower level learners overwhelmingly use the primary modals will and can. In addition to the use of primary modals that indicate a clear function-like a narrative technique to engage the reader or set the scene (e.g. I can tell you that..., as you can see..., I will never forget...), the secondary modals would and could appear in the data of the higher level learners. The prompts of the writing task established an obligatory past time context (write about a favorite or worse vacation or an experience with a favorite person, possession, or movie). In the early stages of the acquisition of English modality, Arab ELLs have not yet acquired the means to express past modal time. Therefore, they rely heavily on the primary modals to express past meanings as in the following examples from Levels 1, 2, and 3 of the BALC corpus:

1. He will go in the Ibn Batota and swimming. (CEPA 1 200600318)
2. When my father told me I will travel, I will be a very happy because I enjoyed in the last traveled. (CEPA 2 200606139)
3. It was so wonderful holiday because sleep hotel and sleep in desert Dubai, and I can see family. I can see uncle and grandfather and grandmother. I can play game. And I can eat only food. I can go to sea. I can go cinema. (CEPA 3 200612553)

Examples 1 to 3 demonstrate the use of will and can to express past modal meanings such as intentionality, possibility, habit, and ability. In higher levels, on the other hand, there is a clear shift from arbitrary modal usages, such as the use of will to express a past meaning, to more focused usages like the utilization of secondary modals to express past and hypothetical meanings. In the following examples, the learners clearly differentiate between expressing a current predisposition using will as in Example 4 and describing past volitions and predictions with would, illustrated in Examples 5 and 6 and with obligations in Example 7.

4. So don’t ask me why I love this place and I won’t stop loving it until the last moment of my life, so I think Barny was singing about us “I love you, you love me, we are happy family”. (CEPA 5 200603190)
5. We were very happy because we would go to UK. (CEPA 4 200611068)
6. We were told that the plane had something wrong and it would be late. We waited for three hours in the airport. (CEPA 5 200600215)
7. We were in the airport waiting in the departure lounge when they announced that we would have to wait about forty-five minutes... (CEPA 6 200619772)

Learners in the upper levels have control over modality to express a variety of meanings. For instance, Examples 8, 9, and 10 show how could + not is used to describe particular situations that were not possible.

8. I looked to the place where he pointed it was a lake full of frogs. I couldn’t do anything it was really horrible. (CEPA 4 200607662)
9. Then we went to Disneyland. Wow, I could not believe that it is full of games and so many people there. (CEPA 5 200607633)

10. I was so frustrated that I couldn’t control my temper and what a summer holiday it was. (CEPA 6 200620682)

Examples 11 and 12 provide instances of the expressions utilized to express conditionality. The utterance from Level 3 (Example 11) shows that the learner relied mainly on if to convey a condition; whereas in the writing sample from a Level 6 student (Example 12), various modal expressions are combined to express a condition or wish.

11. If you have a holiday with your parents and your family, don’t keep away from it. (CEPA 3 200612006)

12. If I could, I would buy a house up the mountains facing the lake. That’s my one and only wish. (CEPA 6 200621115)

New forms emerge to take on more specialized meanings. This is the case with both modal auxiliaries as well as with the periphrastic modals. This divergence and specialization in expressions of modality can be explained within the CoA. Beginner adult ELLs are aware of the need to deliver a modal meaning to specific utterances. However, since their range of modal devices in the lower levels is limited, ELLs sought to use the basic modals that they possess in these levels to establish a modality reference. These basic modal forms were also combined with the past form of a verb to mark a past event, as in the following examples:

13. I can saw in the zoo animal and go Sahara. (CEPA 1 200604247)

14. I can went in Al-Ain is very perfect holiday saw and did wonderful. (CEPA 2 200605590)

15. We will had go the schools. (CEPA 2 200612511)

16. When I went to India last summer firstly I felt happy when my parents told me we will went India this summer. (CEPA 3 200611019)

17. My Family was very angry and the children sad. I can’t studied when I back to home. (CEPA 3 200610975)

These examples show that beginner Arab ELLs have attempted to establish past modal references through the integration of tense/aspect morphology. Since they still have not acquired the proper way for formulating past modals, they sought other modes to indicate that a specific incident took place in the past through combining past morphology. We argue that it is likely a development in tense/aspect morphology. Learners at this stage would benefit from focused instruction on how modality is expressed in the past. Such learners seem ready for this instruction (Schmidt, 1995) since they have a stable past morphology. These findings parallel those from research on temporality within the CoA (e.g. Bardovi-Harlig, 1995, 2000; Dietrich, Klein & Noyau, 1995). Because ELLs still have not developed the morphological means to indicate a past event at the beginning stages of acquisition, they seek other means to convey past meanings with lexical items such as the adverb ‘yesterday’ (e.g. I visit my friend yesterday.). In the case of modality in earlier levels of English proficiency, the secondary types of modal auxiliaries
have not yet emerged in the service of expressing past time. Thus, learners seek other means to mark past modal meanings.

Modals as the main verb. Another pattern observed in lower levels was the use of modal auxiliary verbs as main verbs, illustrated in the following examples:

18. He can family and friend and many. (CEPA 1 200612260)
19. Then I will very happy with my family. (CEPA 2 200612138)
20. I can’t happy because the zoo is only for see not for play this is was so bad. (CEPA 2 200606360)
21. In Saturday I will perfect happy holiday. (CEPA 3 20067753)
22. My brother can very happy because don’t now the problem. (CEPA 3 200611022)

The missing main verb is the verb to BE. For instance, in example 18, the student very likely wanted to say that ‘He can be with family…’ and in the rest of the examples, it is clear that what is missing is be. Within the CoA, these examples suggest that some of the beginner Arab ELLs have not yet developed complex combinations such as a modal auxiliary verb plus verb to BE. Prior research on this population (e.g. Muftah, & Eng, 2011) has indicated that Arab ELLs often omit verb to BE when they produce English utterances.

Productivity as an indicator of language proficiency. As mentioned earlier, a modal auxiliary verb is considered productive when it occurs with various verbs. Productivity of modal expressions is regarded an indicator that the learner has developed a target-like modal system (Giacalone Ramat, 1992, 1995; Author 2, 2000; Author 2 & Bardovi-Harlig, 2000, 2001). In levels 1, 2, and 3 will + go, will + be, and will + happy were the only combinations that occurred at least 5 times (34, 7 and 7 times respectively). In levels 4, 5, and 6 there were many frequent combinations of modal auxiliary verbs and periphrastic modals with main verbs (see Table 4). The majority of the modal auxiliary verbs were associated with verbs such as go, enjoy, see, forget, and visit which can be ascribed to the nature of the writing task. Additionally, periphrastic modals appear to be productive with combination types of wish, if, think, and hope. Combinations of would and could with main verbs is most productive in Level 6. As learners advance in their level of English proficiency, further modal expressions become more productive. This is an indicator that the concept of modality has grammaticalized (Giacalone Ramat, 1992, 1995). That is, learners have moved from lexical modes of expressing modality to grammatical means through the use of various forms of modal auxiliary verbs and periphrastic modals. This is a sign that learners in the upper levels have established target-like (English) modal expressions (Giacalone Ramat, 1992, 1995; Author 2, 2000).
Table 4
Modal Combinations and their Total Occurrences in Higher Levels

<table>
<thead>
<tr>
<th>Combination</th>
<th>Level 4 Total Occurrences</th>
<th>Combination</th>
<th>Level 5 Total Occurrences</th>
<th>Combination</th>
<th>Level 6 Total Occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td>can + spend</td>
<td>7</td>
<td>can + play</td>
<td>5</td>
<td>can + see</td>
<td>19</td>
</tr>
<tr>
<td>can + see</td>
<td>19</td>
<td>can + use</td>
<td>9</td>
<td>can + enjoy</td>
<td>7</td>
</tr>
<tr>
<td>can + help</td>
<td>5</td>
<td>can + see</td>
<td>12</td>
<td>can + be</td>
<td>12</td>
</tr>
<tr>
<td>can + enjoy</td>
<td>7</td>
<td>can + do</td>
<td>12</td>
<td>can + do</td>
<td>5</td>
</tr>
<tr>
<td>can + do</td>
<td>6</td>
<td>can + get</td>
<td>7</td>
<td>can + go</td>
<td>7</td>
</tr>
<tr>
<td>can + take</td>
<td>5</td>
<td>can + be</td>
<td>11</td>
<td>could + be</td>
<td>9</td>
</tr>
<tr>
<td>can’t + forget</td>
<td>6</td>
<td>can + visit</td>
<td>7</td>
<td>could + have</td>
<td>5</td>
</tr>
<tr>
<td>must + go</td>
<td>5</td>
<td>should + be</td>
<td>7</td>
<td>should + go</td>
<td>8</td>
</tr>
<tr>
<td>should + go</td>
<td>6</td>
<td>should + go</td>
<td>5</td>
<td>should + be</td>
<td>7</td>
</tr>
<tr>
<td>will + die</td>
<td>6</td>
<td>will + never</td>
<td>18</td>
<td>will + forget</td>
<td>9</td>
</tr>
<tr>
<td>will + never</td>
<td>14</td>
<td>will help</td>
<td>5</td>
<td>will + always</td>
<td>13</td>
</tr>
<tr>
<td>will + be</td>
<td>23</td>
<td>will + always</td>
<td>5</td>
<td>will + never</td>
<td>8</td>
</tr>
<tr>
<td>will + forget</td>
<td>8</td>
<td>would + be</td>
<td>8</td>
<td>will + come</td>
<td>5</td>
</tr>
<tr>
<td>will + enjoy</td>
<td>6</td>
<td>would + like</td>
<td>5</td>
<td>will + enjoy</td>
<td>5</td>
</tr>
<tr>
<td>will + good</td>
<td>5</td>
<td>hope + be</td>
<td>8</td>
<td>will + be</td>
<td>15</td>
</tr>
<tr>
<td>will + go</td>
<td>17</td>
<td>hope + can</td>
<td>5</td>
<td>will + have</td>
<td>8</td>
</tr>
<tr>
<td>hope + go</td>
<td>10</td>
<td>think + will</td>
<td>6</td>
<td>would + say</td>
<td>5</td>
</tr>
<tr>
<td>think + is</td>
<td>8</td>
<td>think + was</td>
<td>18</td>
<td>would + be</td>
<td>36</td>
</tr>
<tr>
<td>if + go</td>
<td>5</td>
<td>thought + will</td>
<td>5</td>
<td>would + never</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>want + go</td>
<td>8</td>
<td>would + love</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>wish + could</td>
<td>7</td>
<td>would + like</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>wish + can</td>
<td>5</td>
<td>would + go</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>would + have</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>would + get</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>would + always</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>if + could</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>if + were</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hope + will</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hope + go</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>think + is</td>
<td>10</td>
</tr>
</tbody>
</table>

Adult Arab ELLs express the concept of modality in their L2 with limited means before they have fully acquired the linguistic tools to do so. As they progress in their language proficiency, more varied and frequent expressions of modality emerge. This phenomenon is consistent with Salsbury’s (2000) findings:

New forms are added to the expression of a concept, but these newer forms do not replace older forms, at least not initially. As we see these new forms emerge...
and added to the overall number of linguistic forms available to learners in the expression of the concept. (p. 292)

Language learners make use of the linguistic means that they possess at a given time, and later on in the acquisition process, they attempt to enrich their inventory of the concept that they already have (von Stutterheim & Klein, 1987; Bardovi-Harlig, 2007).

**Conclusion**

This study investigated the development of English modality by Arab ELLs across six levels of English proficiency, through analyzing a total of 1,200 texts from the BALC corpus. This was guided by the CoA as a theoretical framework which enables researchers to understand what devices an L2 learner uses to express a certain concept, and how this concept is expressed at a given time or level. The findings have yielded the following phenomena:

- Arab ELLs express the concept of modality in earlier levels before they have the linguistic means to do so, and it becomes more productive and variant as learners progress in their language proficiency.
- In the acquisition stages of the English modality, learners rely heavily on the primary modals of can and will because these are the most salient forms (high frequency and low variation) as well as being regarded by the learners as lexical instead of grammatical (based on the learners’ L1). More forms and types of modal expressions with focused meanings emerge suddenly at high levels of proficiency. At this high level of proficiency learners make clear distinctions between the usages of, for example, can and could. This phenomena illustrates dynamic and chaotic development as opposed to gradual or linear development.
- Productivity in the modal combinations increases dramatically and new forms emerge to take on new meanings in higher levels.

The present study was successful in explaining the emergence of English modality in Arab ELLs’ written discourse. However, there are some limitations to the study. The first limitation relates to the medium of communication. The CoA is like a window to an ELL’s L2. In addition to looking at the L2 learners’ written discourse, it examines their use of the concept under research while learners are engaged in communicative tasks. Through the inclusion of the latter element, the researcher would also be able to see the interaction of the research participant with the interlocutor. Second, most research that follows the CoA is longitudinal in nature. In the present study, it cannot be determined whether Arab ELLs went through the pragmatic stage in the acquisition of modality, as shown in the research findings on the acquisition of temporality, because we only had access to a single text from a given learner at just one point in time. Nonetheless, looking cross-sectionally at Arab ELLs’ written discourse through a relatively large corpus allowed for a comparison between the different levels of language proficiency and yielded results that parallel some of those of the longitudinal designs (e.g. Author 2, 2000).
This research also adds to the body of the literature on SLA of English modality especially of Arab ELLs since there have been no studies that have particularly examined this population.

Pedagogical Implications: Teaching English Modal Forms, Meanings, and Uses

As our findings have shown, Arab ELLs rely heavily on primary modals. Secondary modals emerge when learners reach an advanced level of English proficiency. Not only do Arab ELLs use limited modals to express different meanings, but they also have problems with structures such as the use of past tense forms after modal auxiliary verbs. Accordingly, more focused instruction on the different patterns of modals, including the periphrastic modals, is needed. One suggestion is to provide multiple exemplars from corpus-based materials from which learners may derive their own rules (Thornbury, 1999; McEnery & Xiao, 2010). In the current study, Arab ELLs were capable of conveying the concept of modality in earlier levels but with limited linguistic means. Offering Arab ELLs in beginner levels with a greater number of focused exemplars from authentic texts will create greater salience to modal form and function pairings (Boyd & Goldberg, 2009). Several English grammar textbook authors have used corpus-based techniques in developing pedagogical materials (Azar & Hagen, 2009; Reppen, 2011).

In addition to the complex grammatical structure of English modals, a single modal expression like could may convey multiple meanings, thus complicating their use or function in communicative settings. For example, the misuse of English modals by Arabic English speakers is likely related to the differences between the social norms and cultural values in English-speaking and Arab communities (Abdelrazeq, 2011). To avoid unexpected (or inappropriate) use of English modals by Arab ELLs, teachers need to develop communicative activities to reinforce the various modal meanings. Yule (1999) suggests providing students with “the context or circumstances in which those modal forms are used” (p. 91). Teachers can provide many contextualized examples so that students see how the different modal expressions are used in context and what meanings they convey.

Finally, students should experience the different modals in various communicative settings through, for example, role plays where students work in pairs or groups to practice giving advice and suggestions, talking about personal abilities and preferences, and making requests and offers. Teachers can develop guided activities such as providing students with specific scenarios and modal formulas like offering a ride to a friend or investigating a crime incident where students need to use past modals (modal auxiliary + have + past participle). Students can also work collaboratively on projects like creating invitation cards and appropriately responding to the invitation by accepting or rejecting. This kind of activity will in turn familiarize students with social norms in the target community. They can also design surveys or questionnaires to collect information about their classmates’ personal abilities (or friends and family members). Additionally, teachers can incorporate everyday objects and visuals such as using images of road signs. Students then interpret these
through the use of modals of obligation. Excerpts from magazines, newspapers, and movies can be a rich source to analyze and discuss the use of modals.

**Resources for Teaching English Modals**

Some of the recent grammar textbooks that have been effective for teaching English modality are the Grammar and Beyond series (2011, 2012) and Understanding and Using English Grammar 4th edition (Azar & Hagen, 2009). The latter is appropriate for intermediate and advanced English learners. It offers charts with straightforward explanations to form and meaning of modals followed by various communicative exercises including listening, speaking, and reading activities. Grammar Speaks is an interesting online resource offered by www.azargrammar.com, it provides grammar explanations in an engaging way. The former is also a powerful source to teach modals for beginner, intermediate, and advanced learners as it incorporates various authentic and communicative contexts elicited from the Cambridge International Corpus. It highlights differences in the use of modals in spoken and written language. Some online communicative activities that include modals can be accessed through http://www.cambridge.org/grammarandbeyond. Another excellent resource for teachers is Explaining English Grammar (Yule, 1999). It offers a chapter on teaching English modality with a thorough explanation of form and meaning as well as teaching tips.

Teachers might also make use of the free corpora available online, like the Corpus of Contemporary American English (COCA), and have students search for specific modal expressions so that they can examine how they are used by the native speakers of the language. Another online user-friendly concordance is StringNet http://www.lexchecker.org/, which contains two billion multiword patterns extracted from the British National Corpus. This tool can be particularly useful for upper-intermediate and advanced learners in academic writing to find various examples for the same modal pattern in context. To conclude, offering students with different opportunities to notice and practice modals would make them salient and frequent in the input, and this in turn would facilitate the developmental process of this complex yet vital concept in English.

**References**


BIOGRAPHICAL SKETCHES

Eman Elturki is a doctoral candidate in the Language, Literacy and Technology program at Washington State University. She holds a master’s degree in TESOL from the University of Southern California. She teaches ESL at the Intensive American Language Center of Washington State University. Elturki’s research focuses on exploring second language acquisition using corpus linguistics as a methodology.

Tom Salsbury is Associate Professor in the Department of Teaching and Learning, College of Education, at Washington State University. He has published in the areas of pragmatics and language learning, discourse analysis and content-based instruction. His current work is in second language lexical development.