

King Alfred: A Translation Environment for Learners of Anglo-Saxon English

Lisa N. Michaud

Computer Science Department
St. Anselm College
Manchester, NH 03102
lmichaud@anselm.edu

Abstract

King Alfred is the name of both an innovative textbook and a computational environment deployed in parallel in an undergraduate course on Anglo-Saxon literature. This paper details the ways in which it brings dynamically-generated resources to the aid of the language student. We store the feature-rich grammar of Anglo-Saxon in a bi-level glossary, provide an annotation context for use during the translation task, and are currently working toward the implementation of automatic evaluation of student-generated translations.

1 Introduction

Criticisms of the application of computational tools toward language learning have often highlighted the reality that the mainstays of modern language teaching—including dialogue and a focus on communicative goals over syntactic perfectionism—parallel the shortcomings of computational environment. While efforts continue to extend the state of the art toward making the computer a conversational partner, they nevertheless often fall short of providing the language learner with learning assistance in the task of communicative competence that can make a real difference within or without the classroom.

The modern learner of ancient or “dead” languages, however, has fundamentally different needs; learners are rarely asked to produce utterances in the language being learned (L2). Instead of communication or conversation, the focus is on translation from source texts into the learner’s native language (L1). This translation task typically involves annotation of the source text as syntactic data in the L2 are

decoded, and often requires the presence of many auxiliary resources such as grammar texts and glossaries.

Like many learners of ancient languages, the student of Anglo-Saxon English must acquire detailed knowledge of syntactic and morphological features that are far more complex than those of Modern English. Spoken between circa A.D. 500 and 1066, Anglo-Saxon or “Old” English comprises a lexicon and a grammar both significantly removed from that of what we speak today. We therefore view the task of learning Anglo-Saxon to be that of acquiring a foreign language even to speakers of Modern English.

In the Anglo-Saxon Literature course at Wheaton College¹, students tackle this challenging language with the help of *King Alfred’s Grammar* (Drout, 2005). This text challenges the learner with a stepped sequence of utterances, both original and drawn from ancient texts, whose syntactic complexity complements the lessons on the language. This text has recently been enhanced with an electronic counterpart that provides the student with a novel environment to aid in the translation task. Services provided by the system include:

- A method to annotate the source text with grammatical features as they are decoded.
- Collocation of resources for looking up or querying grammatical- and meaning-related data.
- Tracking the student’s successes and challenges in order to direct reflection and further study.

¹Norton, Massachusetts



Figure 1: The main workspace for translation in King Alfred.

This paper overviews the current status of the King Alfred tutorial system and enumerates some of our current objectives.

2 System Overview

King Alfred is a web-accessible tutorial environment that interfaces with a central database server containing a curriculum sequence of translation exercises (Drout, 1999). It is currently implemented as a Java applet using the Connector/J class interface to obtain curricular, glossary, and user data from a server running MySQL v5.0.45.

When a student begins a new exercise, the original Anglo-Saxon sentence appears above a text-entry window in which the student can type his or her translation as seen in Figure 1. Below this window, a *scratch pad* interface provides the student with an opportunity to annotate each word with grammatical features, or to query the system for those data if needed. This simultaneously replaces traditional annotation (scribbling small notes in between lines of the source text) and the need to refer to auxiliary resources such as texts describing lexical items and morphological patterns. More on how we address the latter will be described in the next section.

When the student is finished with the translation, she clicks on a “Submit” button and progresses to a second screen in which her translation is displayed alongside a stored instructor’s translation from the database. Based on the correctness of scratch pad annotations aggregated over several translation exercises, the system gives feedback in the form of a simple message, such as *King Alfred is pleased with your work on strong nouns and personal pronouns*, or *King Alfred suggests that you should review weak verbs*. The objective of this feedback is to give the students assistance in their own self-directed study. Additional, more detailed informa-

tion about the student’s recorded behavior is viewable through an open user model interface if the student desires.

3 Resources for the Translation Task

As part of the scratch pad interface, the student can annotate a lexical unit with the value of any of a wide range of grammatical features dependent upon the part of speech. After the student has indicated the part of speech, the scratch pad presents an interface for this further annotation as seen in Figure 2, which shows the possible features to annotate for the verb *feoll*.

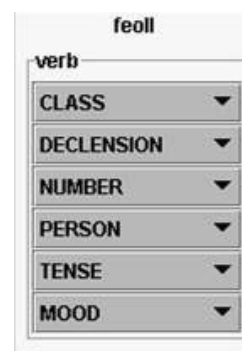


Figure 2: A scratch pad menu for the verb *feoll*.

The scratch pad provides the student with the opportunity to record data (either correctly, in which case the choice is accepted, or incorrectly, where the student is notified of having made a mistake) or to query the system for the answer. While student users are strongly encouraged to make educated guesses based on the morphology of the word, thrashing blindly is discouraged; if the information is key to the translation, and the student does not have any idea, asking the system to *Tell me!* is preferable to continually guessing wrong and it allows the student to get “unstuck” and continue with the transla-

tion. None of the interaction with the scratch pad is mandatory; the translator can proceed without ever using it. It merely exists to simultaneously allow for recording data as it is decoded, or to query for data when it is needed.

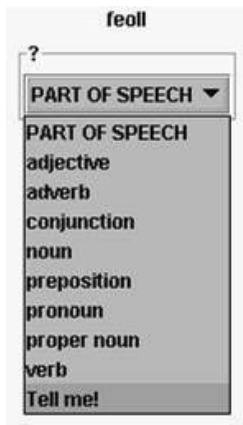


Figure 3: Querying King Alfred for help.

3.1 Lexical Lookup

Like most Anglo-Saxon texts, *King Alfred* also contains a glossary which comprises all of the Anglo-Saxon words in the exercise corpus. These glossaries typically contain terms in “bare” or “root” form, stripped of their inflection. A novice learner has to decode the root of the word she is viewing (no easy task if the inflection is irregular, or if she is unaware, for example, which of seven declensions a verb belongs to) in order to determine the word to search for in the glossary, a common stumbling block (Colazzo and Costantino, 1998). The information presented under such a root-form entry is also incomplete; the learner can obtain the meaning of the term, but may be hampered in the translation task by not knowing for certain how this particular instance is inflected (e.g., that this is the third person singular present indicative form), or which of the possible meanings is being used in this particular sentence.

Alternatively, a text can present terms in their surface form, exactly as they appear in the exercise corpus. This approach, while more accessible to the learner, has several drawbacks, including the fact that glossary information (such as the meaning of the word and the categories to which it belongs) is common to all the different inflected versions, and

it would be redundant to include that information separately for each surface form. Also, in such an entry the user may not be able to discover the root form, which may make it more difficult to recognize other terms that share the same root. To avoid these issues, a glossary may contain both, with every surface form annotated with the information about its inflection and then the root entry shown so that the reader may look up the rest of the information.

We believe we can do better than this. In order to incorporate the advantages of both forms of glossary data, we have implemented two separate but interlinked glossaries, where each of the surface realizations is connected to the root entry from which it is derived. Because electronic media enable the dynamic assembly of information, the learner is not obligated to do two separate searches for the information; displaying a glossary entry shows both the specific, contextual information of the surface form and the general, categorical data of the root form in one presentation. This hybrid glossary view is shown in Figure 4.

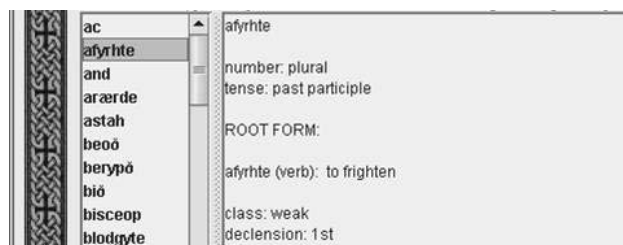


Figure 4: A partial screen shot of the King Alfred glossary browser.

3.2 Surface and Root Forms

To build this dual-level glossary, we have leveraged the *Entity-Relationship Model* as an architecture on which to structure King Alfred’s curriculum of sentences and the accompanying glossary. Figure 5 shows a partial Entity-Relationship diagram for the relevant portion of the curriculum database, in which:

- *Sentences* are entities on which are stored various attributes, including a holistic *translation* of the entire sentence provided by the instructor.
- The relationship *has word* connects Sentences

to *Words*, the collection of which forms the surface level of our glossary. The instances of this relationship include the ordinality of the word within the sentence; the actual sentence is, therefore, not found as a single string in the database, but is constructed dynamically at need by obtaining the words in sequence from the glossary. Each instance of the relationship also includes the translation of the word *in the specific context of this sentence*.²

- The entity set *Words* contains the actual orthography of the word as it appears (*text*) and through an additional relationship set (not shown) is connected to all of the grammatical features specific to a surface realization (e.g. for a noun, *person*=third, *number*=singular, *case*=nominative).
- The relationship *has root* links entries from the surface level of the glossary to their corresponding entry at the root level.
- The *Roots* glossary has the orthography of the root form (*text*), possible definitions of this word, and through another relationship set not in the figure, data on other syntactic categories general to any realization of this word.

Since the root form must be displayed in some form in the glossary, we have adopted the convention that the root of a verb is its infinitive form, the roots of nouns are the singular, nominative forms, and the roots of determiners and adjectives are the singular, masculine, nominative forms.

Other related work does not explicitly represent the surface realization in the lexicon; the system described by (Colazzo and Costantino, 1998), for example, uses a dynamic word stemming algorithm to look up a surface term in a glossary of root forms by stripping off the possible suffixes; however, it is unable to recognize irregular forms or to handle ambiguous stems. GLOSSER (Nerbonne et al., 1998)

²This does not negate the necessity of the holistic translation of the sentence, because Anglo-Saxon is a language with very rich morphology, and therefore is far less reliant upon word order to determine grammatical role than Modern English. In many Anglo-Saxon sentences, particularly when set in verse, the words are “scrambled” compared to how they would appear in a translation.

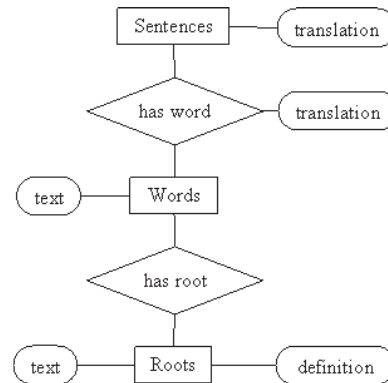


Figure 5: A piece of the Entity-Relationship diagram showing the relationships of Sentences, Words, and Roots.

for Dutch learners of French also automatically analyzes surface terms to link them to their stem entries and to other related inflections, but shares the same problem with handling ambiguity.

Our approach ensures that no term is misidentified by an automatic process which may be confused by ambiguous surface forms, and none of these systems allows the learner access to which of the possible meanings of the term is being used in *this* particular context. The result of King Alfred’s architecture is a pedagogically accurate glossary which has an efficiency of storage and yet dynamically pulls together the data stored at multiple levels to present the learner with all of the morphosyntactic data which she requires.

3.3 Adding to the Glossary

Because there is no pre-existing computational lexicon for Anglo-Saxon we can use and because creating new translation sentences within this database architecture via direct database manipulation is exceedingly time consuming—and inaccessible for the novice user—we have equipped King Alfred with an extensive instructor’s interface which simultaneously allows for the creation of new sentences in the curriculum and the expansion of the glossary to accommodate the new material.³

The instructor first types in an Anglo-Saxon sentence, using special buttons to insert any non-ASCII characters from the Anglo-Saxon alphabet. A holis-

³All changes created by this interface are communicated directly to the stored curriculum in the central server.

tic translation of the entire sentence is entered at this time as well. The interface then begins to process each word of the sentence in turn. At each step, the instructor views the entire sentence with the word currently being processed highlighted:

- Sum mann feoll on ise.

The essential process for each word is as follows:

1. The system searches for the word in the surface glossary to see if it has already occurred in a previous sentence. All matches are displayed (there are multiple options if the same realization can represent more than one inflection) and the instructor may indicate which is a match for this occurrence. If a match is found, the word has been fully processed; otherwise, the interface continues to the next step.
2. The instructor is prompted to create a new surface entry. The first step is to see if the root of this word already exists in the root glossary; in a process similar to the above, the instructor may browse the root glossary and select a match.
 - (a) If the root for this word (*feallan* in our example) already exists, the instructor selects it and then provides only the additional information specific to this realization (e.g. *tense*=past, *person*=3rd, *number*=singular, and *mood*=indicative).
 - (b) Otherwise, the instructor is asked to provide the root form and then is presented with an interface to select features for both the surface and root forms (the above, plus *class*=strong, *declension*=7th, *definition*="to fall").

When this process has been completed for each word, the sentence is finally stored as a sequence of indices into the surface glossary, which now contains entries for all of the terms in this sentence. The instructor's final input is to associate a contextual gloss (specific to this particular sentence) with each word (these are used as "hints" for the students when they are translating and need extra help).

4 Automatically Scoring a Translation

When initially envisioned, King Alfred did not aspire to automatic grading of the student-generated translation because of the large variation in possible translations and the risk of discouraging a student who has a perfectly valid alternative interpretation (Drout, 1999). We now believe, however, that King Alfred's greatest benefit to the student may be in providing accurate, automatic feedback to a translation that takes the variety of possible translation results into account.

Recent work on machine translation evaluation has uncovered methodologies for automatic evaluation that we believe we can adapt to our purposes. Techniques that analyze *n*-gram precision such as BLEU score (Papineni et al., 2002) have been developed with the goal of comparing candidate translations against references provided by human experts in order to determine accuracy; although in our application the candidate translator is a student and not a machine, the principle is the same, and we wish to adapt their technique to our context.

Our approach will differ from the *n*-gram precision of BLEU score in several key ways. Most importantly, BLEU score only captures potential *correct* translations but equally penalizes errors without regard to how serious these errors are. This is not acceptable in a pedagogical context; take, for example, the following source sentence⁴:

- (1) *Sum mann feoll on ise.*

The instructor's translation is given as:

- (2) One man fell on the ice.

Possible student translations might include:

- (3) One man fell on **ice**.
- (4) **Some** man fell on the ice.

In the case of translation (3), the determiner before the indirect object is implied by the case of the noun

⁴This example sentence, also used earlier in this paper, reflects words that are very well preserved in Modern English to help the reader see the parallel elements in translation; most sentences in Anglo-Saxon are not nearly so accessible, such as shown in example (5).

ise but not, in the instructor's opinion, required at all. Translation (3) is therefore as valid as the instructor's. Translation (4), on the other hand, reflects the presence of the *faux ami*, or false friend, in the form of *sum*, which looks like Modern English 'some' but should not be translated as such. This is a minor mistake which should be corrected but not seen as a reflection of a serious underlying grammatical misconception.

Adverbs that modify the main verb also have flexible placement:

- (5) *Pa wurdon þa mynstermen miccle afyrhte.*
- (6) **Then** the monks became greatly frightened.
- (7) The monks **then** became greatly frightened.
- (8) The monks became **then** greatly frightened.
- (9) The monks became greatly frightened **then**.

And there are often many acceptable translations of a given word:

- (10) Then the monks became greatly **afraid**.

What we wish to focus our attention on most closely are misinterpretations of the morphological markers on the source word, resulting in a misinflected translation:

- (11) Then the monks **become** greatly frightened.

This is a difference which is most salient in a pedagogical context. Assuming that the student is unlikely to make an error in generating an utterance in her native language, it can be concluded that such an error reflects a misinterpretation of the source morphology.

A summary of the differences between our proposed approach and that of (Papineni et al., 2002) would include:

- The reliance of BLEU on the diversity of multiple reference translations in order to capture some of the acceptable alternatives in both

word choice and word ordering that we have shown above. At this time, we have only one reference translation with which to compare the candidate; however, we have access to other resources which can be applied to the task, as discussed below.

- The reality that automatic MT scoring usually has little to no grammatical data available for either the source or target strings of text. We, however, have part of speech tags for each of the source words encoded as part of the curriculum database; we also have encoded the word or short phrase to which the source word translates, which for any target word occurring in the candidate translation essentially grants it a part of speech tag. This means that we can build in flexibility regarding such elements as adverbs and determiners when the context would allow for optional inclusion (in the case of determiners) or multiple placements (in the case of adverbs).
- Multiple possible translations of the word can come from a source other than multiple translators. We intend to attempt to leverage WordNet (Fellbaum, 1998) in situations where a candidate word does not occur in the reference translation to determine if it has a synonym that does. The idea of recognizing a word that does not match the target but nevertheless has a related meaning has previously been explored in the context of answers to reading comprehension questions by (Bailey, 2007).
- Minor mistranslations such as *sum*/some due to *faux amis* can be captured in the glossary as a kind of "bug rule" capturing typical learner errors.
- Other mistranslations, including using the wrong translation of a source word for the context in which it occurs—a common enough problem whenever a novice learner relies on a glossary for translation assistance—can be caught by matching the multiple possible translations of a root form against an unmatched word in the candidate translation. Some morphological processing may have to be done

to match a stem meaning against the inflected form occurring in the candidate translation.

- The primary focus of the automatic scoring would be the misinflected word which can be aligned with a word from the reference translation but is not inflected in the same way. Again, morphological processing will be required to be able to pair together mismatched surface forms, with the intention of achieving two goals:

1. Marking in the student model that a misinterpretation has occurred.
2. Giving the user targeted feedback on how the source word was mistranslated.

With this extension, King Alfred would be empowered to record much richer data on student competency in Anglo-Saxon by noting which structures and features she translates correctly, and which she has struggled with. Such a model of student linguistic mastery can be a powerful aid to provide instructional feedback, as discussed in (Michaud and McCoy, 2000; Michaud and McCoy, 2006; Michaud et al., 2001).

5 Other New Directions

Ongoing work with the glossary browser includes enhancements to include dynamically generated references to other occurrences of words from the same stem or root throughout the translation corpus in order to reflect other inflected forms in their contexts as many dictionaries do.

This, however, is a relatively simplistic attempt to illustrate the pattern of morphological inflection of a root to the learner. A long-term plan is to incorporate into King Alfred a full morphological engine encoding the inflection patterns of Anglo-Saxon English so that the surface glossary is only needed as a collection of the feature values active in a specific context; with the ability to dynamically generate fully inflected forms from the root forms, King Alfred would empower the learner to access lessons on inflection using the specific words occurring in a sentence currently being translated.

We are unaware of any existing efforts to encode Anglo-Saxon morphology in such a fashion, although in other learning contexts the system Word

Manager (Hacken and Tschichold, 2001) displays a lexicon grouping other words applying the same inflection or formation rule in order to aid the learner in acquiring the rule, a similar goal.

6 Conclusion

King Alfred was deployed in the Anglo-Saxon literature course at Wheaton College in the Fall semesters of 2005 and 2007. Preliminary feedback indicates that the students found the hybrid glossary very useful and the collocation of translation resources to be of great benefit to them in completing their homework assignments. Ongoing research addresses the aggregation of student model data and how the system may best aid the students in their independent studies.

We are most excited, however, about how we may leverage the structuring of the curriculum database into our dual-level linguistic ontology toward the task of automatically evaluating translations. We believe strongly that this will not only enhance the student experience but also provide a rich stream of data concerning student mastery of syntactic concepts. The primary objective of student modeling within King Alfred is to provide tailored feedback to aid students in future self-directed study of the linguistic concepts being taught.

7 Acknowledgments

The Anglo-Saxon course at Wheaton College is taught by Associate Professor of English Michael Drout. Student/faculty collaboration on this project has been extensively supported by Wheaton grants from the Davis, Gebbie, and Mars Foundations, and the Emily C. Hood Fund for the Arts and Sciences. We would particularly like to thank previous undergraduate student collaborators David Dudek, Rachel Kappelle, and Joseph Lavoine.

References

- Stacey Bailey. 2007. On automatically evaluating answers to reading comprehension questions. Presented at CALICO-2007, San Marcos, Texas, May 24-26.
- Luigi Colazzo and Marco Costantino. 1998. Multi-user hypertextual didactic glossaries. *International Journal of Artificial Intelligence in Education*, 9:111-127.

- Michael D. C. Drout. 1999. King Alfred: A teacher controlled, web interfaced Old English learning assistant. *Old English Newsletter*, 33(1):29–34, Fall.
- Michael D. C. Drout. 2005. *King Alfred's Grammar*. Version 4.0.
- Christiane Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database*. MIT Press.
- Pius Ten Hacken and Cornelia Tschichold. 2001. Word manager and CALL: structured access to the lexicon as a tool for enriching learners' vocabulary. *ReCALL*, 13(1):121–131.
- Lisa N. Michaud and Kathleen F. McCoy. 2000. Supporting intelligent tutoring in CALL by modeling the user's grammar. In *Proceedings of the Thirteenth International Florida Artificial Intelligence Research Society Conference (FLAIRS-2000)*, pages 50–54, Orlando, Florida, May 22-24. FLAIRS.
- Lisa N. Michaud and Kathleen F. McCoy. 2006. Capturing the evolution of grammatical knowledge in a CALL system for deaf learners of English. *International Journal of Artificial Intelligence in Education*, 16(1):65–97.
- Lisa N. Michaud, Kathleen F. McCoy, and Litza A. Stark. 2001. Modeling the acquisition of English: an intelligent CALL approach. In Mathias Bauer, Piotr J. Gmytrasiewicz, and Julita Vassileva, editors, *Proceedings of the 8th International Conference on User Modeling*, volume 2109 of *Lecture Notes in Artificial Intelligence*, pages 14–23, Sonthofen, Germany, July 13-17. Springer.
- John Nerbonne, Duco Dokter, and Petra Smit. 1998. Morphological processing and Computer-Assisted Language Learning. *Computer-Assisted Language Learning*, 11(5):421–37.
- Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. BLEU: a method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, PA, July 6-12. ACL.