A Methodology for Developing an Error Taxonomy for a Computer Assisted Language Learning Tool for Second Language Learners[1]

Linda Z. Suri & Kathleen F. McCoy

Technical Report No. 93-16[†]
January, 1993

Linda Z. Suri & Kathleen F. McCoy
Department of Computer and Information Sciences
103 Smith Hall
University of Delaware
Newark, DE 19716 U.S.A.
suri@cis.udel.edu, mccoy@cis.udel.edu

Abstract:

This paper discusses linguistic issues that one must address in order to design an effective CALL system for second language learners. We focus on how one should develop an error taxonomy and indicate how that taxonomy can affect the design of the entire system.

This work was done in the context of designing a CALL tool to help native signers of American Sign Language (ASL) learn written English. Here we report the methodology used in developing the error taxonomy for the system. Our analysis of writing samples from ASL natives indicates that language transfer (LT) (when considered appropriately) can account for many of the errors we found. Recognizing this possibility has substantially influenced the error taxonomy we have developed. The resulting taxonomy captures the true source of many errors and will allow the eventual system to take advantage of ASL knowledge when teaching English. We provide a characterization of language transfer that provides a broad perspective on the ways that two languages may differ and indicates how these differences could potentially influence acquisition and production of the second language.

In addition, we argue that an effective CALL system must not only examine sentences in isolation, but it must identify and correct *discourse-level errors*. We point out that the source of several error classes we identified rests in discourse-level processing. We show how these errors can also be explained in terms of LT, and in doing so we extend the notion of language transfer to the transfer of features which affect the cohesiveness of a discourse. We discuss a methodology for correcting several classes of discourse-level errors. While we explain our methodology and justify our claim in terms of our study of deaf writing, our findings are applicable to the design of CALL systems for other second language learners.

[1]This research was supported in part by NSF Grant #IRI-9010112. Support was also provided by the Nemours Foundation.

[†]An earlier version of this technical report had typographical errors in Table 1.

1 Introduction

In order to develop an effective computer assisted language learning (CALL) system to correct second language production, one must use a design methodology that is linguistically well-grounded in terms of the underlying sources of errors. This is necessary for two reasons: 1) to be able to anticipate and appropriately classify the kinds of errors the user is likely to make; and 2) to provide more appropriate corrective information to the user. In this paper, we describe a methodology for developing such an error taxonomy. The CALL systems we are interested in take written text, analyze it for errors, engage the user in a corrective tutorial dialogue, and generate corrected text. In this paper, we concentrate on the development of an error taxonomy and its implications for the overall system design. Thus, we are concerned here with the linguistic issues one must address in order to develop a taxonomy of errors for learners of a common first language background learning a common second language.

A survey of literature on CALL systems revealed an abundance of CALL systems that prompt the user for input and compare the user's input to a representation of the expected input [ACRS85], [Sar87], [CJ84], [SK84], [Und84], [KK90], [WVJ78], [Sch86]. Such systems may ask the student to define a word or phrase, fill in blanks with characters or words, translate a sentence, nominalize a verb, change the number and person on a verb, or require the student to produce a sentence in response to a question.

While each of these systems can be said to be successful to some extent, for the most part they offer little flexibility and are often at a loss for identifying the true source of a user's errors. The systems typically *explicitly* represent all (instances of) errors that the system could recognize, and provide canned responses for these errors. Thus, the error identification and correction processing is achieved by some level of pattern-matching.

An exception to this approach is the work of [WVJ78] which used a grammar of English augmented with a set of *mal-rules*. A mal-rule allows the system to accept input with instances of a particular *class* of errors and record the nature of the error. We are interested in CALL systems of this nature - ones which are able to capture whole classes of errors rather than only those that have been hard-coded. However, the focus of this paper is on the development of the set of mal-rules. That is, we are interested in how one should determine what classes of errors the mal-rules must capture.[2] We discuss the importance of developing an appropriate error taxonomy in order to implement appropriate mal-rules.

It is important that a CALL system is designed to identify and correct errors that the target user population is likely to produce. In order to build our taxonomy of expected errors, we analyzed writing samples from our target user population. It is also important that the taxonomy capture the regularity of errors. We claim that language transfer (LT), if considered appropriately, can uncover some regularity of the errors, and thus consideration of LT is important in the design of CALL systems for second language learners.

By language transfer, we mean the influence of knowledge of one language on the production of a second language. This definition is much broader than many standard views of language transfer. In this paper, we provide a characterization of LT (or a characterization of how two languages may differ from one another) which captures our broader definition. Our characterization calls for looking at two languages on a feature-by-feature basis and considering how various features could be handled differently.

There are three ways that LT can play a crucial role in the development of an effective CALL system.

[2] We also discuss errors that can not be detected/corrected by syntactic mal-rules of the type used in [WVJ78]. We discuss other mechanisms which could handle classes of such errors.

- By considering the features of the user's native language and the second language, the developer is more likely to *appropriately* classify an error. As a result, some errors which may otherwise seem to occur randomly, or to occur in some situations but not in others without apparent explanation, can be seen as forming a class of errors with an underlying explanation. Thus, appropriate classification is likely to uncover the source of the error.

- By recognizing differences between the native or first language (L1) and the second language (L2) as the possible source of an error, the system will be able to provide the user with a better correction. While this paper is not directly concerned with the details of generating the correction provided to the user of a CALL system, the point to note is that the corrective information will be dependent on the appropriateness of the error classification. For example, for the CALL system that we are developing, the native language (ASL) of the users only marks time at the beginning of an event or discourse segment, while the target language (English) marks tense on every finite verb. Thus, certain missing tense errors may require a correction indicating *when* tense should be marked in English, rather than *how* to mark past tense on a verb in English (which might seem to be the cause of the error if LT were not considered). The point being emphasized is that the better the error classification is, the better the correction provided by the system will be.

- By taking LT into account, one is likely to predict more classes of errors. Since an analysis of writing samples from the target user population will necessarily be limited, some kinds of errors that are likely to be made by that population may not occur in the analyzed samples, or may not occur frequently enough for the designer to recognize the errors as constituting a single class of errors.

The influence of LT has been acknowledged in other work on CALL systems [WVJ78], [Art85], [Sch86]. Schuster and Finin ([Sch86] and [Art85]) designed a system which considered the user's native language as a potential source of error in the production of one particular structure in the second language (L2): verb phrases (VP's) with VP-particles. It should be noted that their system did not correct second language text written by the user, but rather drilled the user with a set of exercises for which a set of possible incorrect answers had been hard-coded. The incorrect answers were motivated by a comparison of the grammars of the first and second language. The generated responses pointed out that the potential source of an error involving VP-particles was a mismatch between the VP-particles of the first and second language.

We distinguish our work from Schuster and Finin in that we consider LT on a much broader basis. Rather than concentrating on a single error class, we look for LT as underlying language production in several different ways. We use our broader definition of LT to allow us to understand more sophisticated sources of errors which might not be apparent if LT were not considered. In addition, as will be discussed below, the notion of finding LT by comparing two grammars (as used by Schuster and Finin) is too narrow to capture the effects of LT in all circumstances. We provide a new characterization of LT which makes clear what the influence of LT is over a much broader spectrum of phenomena.

By taking LT into account, the CALL system design we are proposing takes advantage of the user's knowledge of his or her native language. The particular system we are developing is a CALL system to help (American) deaf writers produce well-formed written English text.[3] Our design proposal may prove to be particularly interesting since the native language of some deaf writers

[3]This tool would be very useful to the deaf population. While data on writing skills is not well-documented, we note that the reading comprehension level of deaf students is considerably lower than that of their hearing counterparts, "... with about half of the population of deaf 18-year-olds reading at or below a fourth grade level and only about

is American Sign Language (ASL), which differs from English in both its syntax and its discourse strategies and thus may have an interesting influence on their written English.

ASL is a visual-gestural language whose grammar is distinct and independent of the grammar of English or any other spoken language [Sto60], [BP78], [BC80], [HS83], [KB79], [BPB83]. In addition to sign order rules, ASL syntax includes systematic modulations to signs as well as non-manual behavior (e.g., squinting, raising of eyebrows, body shifts, and shaking, nodding or tilting the head) for morphological and grammatical purposes [BC80], [Lid80], [Pad81], [KB79], [KG83], [Ing78], [Bak80]. The modality of ASL encourages simultaneous communication of information which is not possible with the completely sequential nature of written English.

In trying to find systematic explanations for the errors in the writing samples we collected from deaf people, we found that LT could be seen as the underlying source of many errors, if it were considered appropriately. In part because of the differences between ASL and English (in terms of their grammars and modality), we were motivated to develop an abstract characterization of LT involving a comparison of the languages on a feature-by-feature basis. This LT characterization should be helpful for other CALL systems because it forces one to look for errors not on the basis of a broad characterization of the two languages, but on the basis of individual features considered one at a time. This approach motivates one to look at errors from the perspective of trying to identify causes in a more conceptually based manner (for example, by considering how the languages mark time or convey the idea of *being*), rather than looking at errors in terms of what syntactic rules of the target language have been violated. This can lead to the more appropriate classification of errors.

A second novel aspect of this work with respect to LT is an extension of the notion of LT to include the transfer of discourse-level features. The writing sample analysis that underlies our findings uncovered several errors which could only be detected and appropriately explained if discourse processing was taken into account. These errors involved, for instance, the appropriate use of cohesion devices (e.g., pronouns). Correcting errors involving such devices involves processing a sentence in the context of the preceding sentences, and these errors affect the processing of subsequent sentences.

Our work is novel in that we provide evidence that several of the discourse-level errors found in our transcript analysis can be explained in terms of LT given our characterization (extended to cover discourse-level features). Thus, we claim that several errors can be explained by the writer carrying over discourse and cohesion strategies from ASL to written English. This kind of discourse-level LT has not, to our knowledge, been previously documented. Previous LT literature documented discourse-level LT involving the overuse of certain lexical items, syntactic structures or pragmatic strategies, and L2 production that seems overly repetitive to L2 native speakers (e.g., Koch [Odl89], Bartelt [Odl89], [Rut83]).

Considering LT in the way we describe, we have found evidence that LT occurs between ASL and written English. Our claim that LT occurs between ASL and English and the evidence we provide for this claim are significant because LT between a visual-gestural language such as ASL and a spoken/written language has not (to our knowledge) previously been documented[4]. In fact, there has been some research ([LM88]) which one might construe as indicating that LT does not occur between ASL and English, but we will show that this would be an unwarranted conclusion. This finding of LT is particularly noteworthy due to the tremendous differences between the two languages (both in grammar and modality). We provide evidence for our belief that LT occurs

10% reading above the eighth grade level..."[Str88] Therefore, this population is likely to benefit a great deal from additional second language interaction.

[4]Other researchers (e.g., [PQ73], [QSW74], [QWM76], [RQP76], [QPS77], [KK78] [QP84]) studied errors in deaf writing but did not attribute the errors to LT.

between these two languages by considering the language features of ASL and English, and by explaining, with examples, how these features coupled with LT can explain some of the errors we have found in the written English of ASL natives.

The focus of this paper is a methodology for developing an error taxonomy for a CALL system which takes LT into account in a sophisticated manner. We characterize LT in a novel way, and extend LT to account for errors at the discourse (as well as the sentence) level. While we are describing a general methodology for designing a taxonomy for a CALL system for second language learners, we developed this model in the context of building a CALL system for ASL natives learning written English. We motivate and explain this methodology through explaining the design and development of that particular taxonomy and system. Thus, in this paper, we first discuss how we developed a taxonomy of errors we have found in the written English of deaf writers.

In this discussion, we will briefly summarize previous LT research (section 2.1) and provide a characterization of LT which was motivated by and captures the kinds of errors we have found (section 2.2). This characterization is consistent with previous LT research. We then will discuss examples of sentence-level error classes in our taxonomy in order to illustrate how our characterization of LT explains many of the errors we have found (section 3).

Next, we will clarify what claims we are making about the role of LT in the written English of the deaf (section 4). We will also explain the difficulties in trying to predict the frequency of errors in a population of writers, explaining why it is often difficult to classify a particular error in a given sample, and what factors complicate making generalizations about groups of writers. In the next section (section 5), we will motivate the need to identify and correct discourse-level errors. We suggest that LT should be considered as a source of such errors since LT played such a prominent role at the sentence-level. We will then introduce three features (of subsequent or anaphoric reference) that function differently in ASL and in English. We will show how these differences can explain three classes of errors (inappropriate or incorrect pronominalization, and illegally omitted NP's) that we have found in the written English of ASL natives. Our explanations of these errors extend the notion of LT to errors which can only be identified by processing a sentence in the context in which it occurs, rather than as an isolated sentence, and which affect the processing of subsequent sentences. We believe that the definition of LT must be extended to capture this kind of transfer in order to identify and correct errors which affect the cohesiveness of the discourse.

Then (in section 6) we discuss the design of the particular CALL system that we are developing, and the role that our characterization of LT played in this design. In particular, we discuss how our characterization of LT helped us to more appropriately classify and explain errors at both the sentence- and discourse-levels, and how we plan to use this information to identify and correct these errors in our CALL system.

In conclusion, we discuss implication of our findings and our methodology on the design of other CALL systems.

2 Characterizing and Explaining the Errors

In order to build an error taxonomy for our CALL system, we collected writing samples from a number of different schools and organizations for the deaf. We concentrated on eliciting samples from Deaf people who are (native) ASL signers.[5] This was done in order to increase the probability of finding errors specific to the deaf population since the writing tool is being designed for use by deaf writers. This also enabled us to more easily identify errors that could be explained by

[5]Identifying native signers of ASL is a difficult task. (See [Sur91].)

knowledge of ASL and would thus provide us insight into whether LT occurs and how ASL and English function (e.g., How is discourse structured in these languages? How are pronouns resolved in these languages?).

For the purpose of developing our taxonomy, we have analyzed forty-eight Freshman and Sophomore writing evaluation samples from Gallaudet University, a liberal arts university for the deaf, seventeen writing evaluation samples from the National Technical Institute for the Deaf (NTID), twelve first draft papers from students in the high school program at the Margaret S. Sterck School, a deaf school in Delaware, and five letters and essays written by ASL natives and collected through the Bicultural Center in Washington, DC. The size of this collection of samples is approximately 25,000 words. We have performed a partial analysis, concentrating on discourse-level errors, on (approximately 100) other writing samples.

The analyzed writing samples led us to an initial taxonomy classifying many of the errors we found. While the taxonomy was useful, we desired a general explanation for the errors in order to be sure we appropriately classified the errors, to predict more error classes, and to provide a useful tutorial correction. We contend that LT can account for a large number of the errors we found, and this finding has helped us classify errors which were difficult to explain by other means. In the next several sections, we will explain why we believe LT occurs between ASL and English, and discuss implications of this belief.

2.1 LT Research

The term language transfer has been used to refer to the influence of knowledge of one's native language (L1) on the production and/or comprehension of a second language (L2). This definition, which we adopt, subsumes other narrower definitions of the term. While the existence of LT has been a rather controversial subject over the years (see [McL87], [GS83], [Sur91]), much recent research has provided convincing evidence of LT resulting in the transfer of L1 lexical items, syntax rules and pragmatic production rules to L2 (e.g., [Sch82] and [SR79] as described in [McL87]; [Kle77], [Hak76], and [Gas79] as described in [Gas84]; [GS83]; and [McL87]).

Given that transfer has been documented between spoken languages, one might ask whether or not LT could occur between ASL (a visual-gestural language) and written English. On the surface, transfer may seem surprising since the modalities of the two languages and thereby the components of ASL grammar and written English grammar are so different. ASL grammar components include sign order, morphological modulations of signs, and non-manual behavior (e.g., squinting, raising of eyebrows, body shifts, and shaking, nodding or tilting the head) which occurs simultaneously with the manual signs. Written English grammar components include word order, morphological modulations of words, and punctuation, but nothing that clearly corresponds to the simultaneous non-manual behavior. At a glance, the fact that ASL and written English occur in different modalities may also seem problematic for the claim that LT occurs between ASL and English. However, research shows that much of ASL processing occurs on the side of the brain primarily used for processing spoken and written languages (the left-side of the brain), as opposed to the (right) side of the brain primarily used for visual and spatial functions [Sac90]. Thus, we expect transfer is likely to occur between ASL and written English, but it is not clear how the transfer will manifest itself (particularly the non-manual component of ASL grammar).

2.2 A Characterization of LT

Because of the differences in grammar and modality between ASL and English, we attempted to abstractly characterize how languages could differ in a way which is independent of the grammar components and modalities of the languages. The characterization calls for looking at languages on a feature-by-feature basis, and comparing whether, when and how each feature is marked in the two languages. We now identify ways in which two languages may differ for individual features which could lead to errors, or the persistence of errors, in L2 production. For each case, we give an example of how LT might affect the written English of ASL natives.

- Two languages may differ in *when* they mark a particular feature. As a result the marking of that feature in the L2 may seem *redundant* (or, alternatively, *overly concise*) in the L2 learner's native language. For example, in ASL it is usual to establish tense at the beginning of a discourse, and then not to mark it again until the time frame changes. Of course, in English tense is marked (on the verb) in every finite clause. So, marking tense in every finite clause in English may seem redundant to the ASL native. Transfer of such a feature (i.e., when to mark tense) might explain omission errors (in this case, omission of tense markings) in the L2.

- Two languages may differ in *how* they mark a feature. For example, in ASL, Yes/No questions are distinguished from declarative statements with non-manual markers (facial expression and body shifts) that are conveyed simultaneously with the propositional content being conveyed. Note that there is no sign order change relative to the same propositional information being conveyed declaratively. This method of forming Yes/No questions is *radically different* from the word order changes which typically mark Yes/No questions in written English. Thus LT might explain errors in Yes/No question formation by the ASL signer learning written English.

- Two languages may also differ in regard to *whether* they *have* a particular feature. As a result, the marking of that feature in the L2 may seem *redundant* or *unnecessary* to the L2 learner whose language does not require that feature, and the writer may have difficulty learning to remember to mark that feature. (Alternatively, if the L2 learner's native language has a feature that the L2 does not, the L2 may seem *overly concise, less rule-governed* or *ambiguous* to the learner.)

  – Often, when one language has a feature that the other language does not, that feature involves the use of syntactic structures or the addition of lexical items or morphological changes that do not carry any informational content. For example, English has the feature of requiring a subject-verb agreement marking (an "s") on the verb for the third person singular subject in the present tense. This morphological marking conveys no extra information, but is a requirement of the English language. The situation with regard to subject-verb agreement is more complex in ASL; when subject-verb agreement is marked it involves a radically different marking than the agreement markings of English (see section 3.2) and the marking is (generally) not empty of informational content. This may explain omissions of the "s" on third-person singular present tense verb in the the written English of ASL signers.

- Finally, in general, there need not be a one-to-one mapping between the lexical items of one language and those of another. In the case where the L1 uses one lexical item for two given concepts, and the L2 uses two separate lexical items for those two concepts, a second language learner is likely to sometimes make an incorrect choice among the L2 lexical items. This would be another form of language transfer.

In fact, English often has two or more words or phrases which correspond to a single ASL sign (or sign sequence), and vice versa. For example, ASL uses the same sign (i.e., lexical item) for "other" and "another". Thus, LT might explain why ASL learners of written English might take some time to learn which word ("other" or "another") to use in English.

3 Examples of Sentence-Level Errors Attributable to LT

In order to give the reader a better understanding of our characterization of LT, we will discuss a few error classes from our taxonomy of errors occurring in the written English of ASL natives. (Our error taxonomy will be discussed below. A table summarizing the taxonomy can be found on page 10.) For each explanatory error class, we provide an example of the error class and then discuss how such an error could be captured by our LT characterization. We provide at least one explanatory error class for each part of the characterization.

3.1 NP and VP Conjunctions

- Conjunction Deletion:
  - "He taught _ directed, for almost 30 years ..."[6]

While researchers have identified several kinds of conjunctive markings in ASL from body shifts to particular lexical items ([Pad81], [BS88]), there are many places where an explicit conjunction would be required in English, but not in ASL. For instance, conjoined verbs do not require an explicit separate lexical item; instead the verbs are signed one after the other [Fan83]. Therefore, it is not surprising that an ASL signer would omit 'and' between (the final and next-to-final) conjoined verbs in written English. This omission could be the result of the conjunctive marking (by the use of a conjunctive lexical item between verbs) in English appearing *redundant* or *radically different* to the ASL native.

3.2 Subject-Verb Agreement

- Incorrect Subject-Verb Agreement:
  - "*My brother like* to go..."

In English, subject-verb agreement is required on all verbs, although the agreement does not carry any informational content.

In ASL, not all verbs mark subject agreement for person and number. For certain verbs (some directional and classifier verbs) subject agreement is indicated by a change in handshape, a change in movement, or (rarely) the use of an overt noun phrase (NP) where it would not normally be needed [Pad81]. There is a large class of verbs in ASL which do *not* vary in form for person and number of the subject ([Pad81]). In addition, some directional verbs vary in form according to the person and number of the *object* [BC80], [Fan83].

[6] **Note:** In this paper, "_" is used to mark places where we think the writer has omitted one or more words from the corresponding correct English sentence.

That subject-verb agreement is a syntactic constraint in English, coupled with the differences in *when* and *how* agreement is marked in the two languages, might explain errors in marking subject-verb agreement in the written English of ASL natives. For an English verb that corresponds in meaning to an ASL verb that has subject agreement, we might find subject-verb agreement errors in the written English of ASL natives since the manner of agreement is radically different between the two languages. For an English verb that corresponds in meaning to an ASL verb that does not have subject agreement, such as the verb "like", we might find subject-verb agreement errors since the ASL verb never requires agreement but the English verb does. In addition, if the writer recognizes that all English verbs require subject agreement even though not all ASL verbs do, we may still see subject-verb agreement errors since the manner of agreement is radically different between the two languages.

3.3 Tense

- Dropping tense on verbs (within or across sentences):
  - "We *went* to see Senator Biden's office ... Then we *go to see* the Vietnam memorial ...."
    (start of new paragraph)
    Then we *ate* ... We *go* long walk... Then we *go* to see ..."[7]

In our data we found missing and incorrect tense markings. These errors might be explained by the differences between when and how ASL and English mark tense.

Some ways that English marks tense are through the use of modals and auxiliary verbs, and through morphological changes to the verbal elements. ASL does not use auxiliaries and it does not modulate verb signs for tense.[8] In ASL, tense is generally established once at the beginning of a discourse segment (e.g., by using a time sign), and that time frame is understood to persist until the time frame or the discourse segment changes.

The dropping of verb tenses in the written English of the deaf might be explained by this; marking tense by modulating every verb in English might seem *redundant* to someone fluent in ASL. Similarly, the problems we found with the formation of English verb tenses might also be explained by the fact that ASL marks time in a *radically different* manner from the way that English does (i.e., by using a specific time indicator instead of adding morphemes and auxiliary verbs).

Notice that a lack of tense marking on a verb in a given writing sample could be the result of at least two different problems: (1) the writer did not think he or she *needed* to mark tense (perhaps because he or she carried over rules about when to mark tense from ASL to English), or (2) the writer did not know *how* to mark the appropriate tense for a particular verb (which would not be surprising since tense is marked in a radically different manner in ASL).

In the first excerpt shown, there is an indication that problem (1) is likely to be the source of the errors of missing tense markers, since although the verb "go" is correctly marked for tense when used at the start of a time/discourse segment, it is not correctly marked for tense elsewhere in the excerpt. Further evidence for this is that the tense is marked on the first verb of each paragraph,

[7]For clarity, we have italicized the words of a sentence which are related to the error classification being discussed. We have sometimes corrected portions of a sentence, placing the corrections in square-brackets, [ ], when the lack of correction of those portions would have made it difficult to understand the error in the sentence that is being classified. Also, to prevent the identification of writers, we have changed the names of institutions and people (to "XYZ college," "XYZ school," or "XYZ") in quoted excerpts from writing samples. We followed these conventions when quoting excerpts throughout this paper.

[8]The positioning of a verb with respect to the ASL time line may reflect tense. However, in most cases, verbs are signed neutrally with respect to this time line.

but then dropped within the paragraph; this seems to suggest that the writer is equating paragraph with discourse segment and is using the standard ASL tense/time marking rules.

3.4 BE, HAVE

- missing BE:
  - "Once the situation changes they _ different people."
- lack of distinction between "be" and "have" (as main verbs):
  - "... some birth controls *are* side-effect." (Correction: *"... have side-effects..."*)
  - "I wish to go to Hawaii because it *is* beautiful *and nice weather.*" (Inappropriate "ellipsis" implies "... and *is* nice weather.")

ASL does not have a "be" sign. In ASL, the idea of *being* is conveyed by *radically different* means: it is (usually) conveyed by the use of a topic-comment structure. In a topic-comment structure, a topic is set up, and then properties are attributed to the topic. The topic and comment are distinguished non-manually in ASL. In English, the idea of *being* is conveyed by the use of the verb "to be" (or sometimes by another verb, e.g., "He has become quite arrogant," or "The room grew silent"). This difference between ASL and English could explain the omission of "to be" as a main verb in the in the written English of the deaf.

While ASL does have a "have" sign, it is often omitted if it can be inferred from the context. The omission of the main verb "have" in the second excerpt could be explained by differences between ASL and English in terms of when the idea of *having* needs to be provided by an explicit lexical item.

3.5 Lexical Items

- Mixing up English words or phrases which share a single ASL sign:
  - "Somehow, *I am interesting in ASL* and I want to learn it."

We noted that when a second language learner's native language uses one lexical item for two given concepts, and the L2 uses two separate lexical items for those two concepts, the second language learner may make an incorrect choice among the L2 lexical items. In this case, a single sign in ASL corresponds to both "interesting" and "interested," and so this error could be attributed to LT.

4 Discussion of the Prominence of LT in the written English of Deaf Writers/ASL natives.

Our characterization of LT has helped us come up with an error taxonomy for the written English of ASL natives based on an analysis of writing samples. The error taxonomy is shown in Table 1. Included in the table is the number of sentences which contained at least one error which could be explained by each classification. These numbers are based on a subset of the writing samples we have analyzed. The subset consists of 21 writing samples (370 sentences, 3490 words) that have already been entered into a computer database. The intention of the numbers is to give an indication of the overall frequency of errors, and to give the reader some idea of the relative frequencies of the different error classes.

- **NP and VP Conjunctions**: 18
  - Omitted conjunction: 10
  - Inappropriate conjunction: 7
  - Extra conjunction: 1
- **Inappopriate sentential conjunction:** 2
- **Prepositions**: 79
  - Omitted preposition: 27
  - Inappropriate preposition: 35
  - Extra Preposition: 17
- **Determiners**: 92
  - Omitted determiner: 49
  - Inappropriate determiner or determiner formation: 17
  - Extra Determiner: 26
- **Incorrect Subject-Verb Agreement**: 14
- **Tense and Aspect**: 95
  - Dropped Tense: 6
  - Extra Auxiliary: 7
  - Missing Auxiliary: 2
  - Incorrect Modal: 3
  - Missing Modal: 2
  - Extra Modal: 1
  - Other tense/aspect problems: 74
- **BE, HAVE (non-Auxiliary)**: 28
  - Omitted BE: 18
  - Lack of BE/HAVE distinction: 10
- **Other Omitted Main Verbs**: 15
- **Incorrect Main Verbs**: 13
  (Poor lexical choice)
- **Relative Clauses**: 26
  - Relative pronoun deletion: 5
  - Resumptive pronoun: 1
  - Incorrect WH-relative pronoun: 4
  - Other: 16
- **Mixing up English words or phrases which share a single ASL sign**: 24
- **Adjective and Adverb Problems**: 27
  - Incorrect Adjective Choice: 3
  - Incorrect Adjective Formation: 13
  - Mixing up Adjectives and Adverbs: 2
  - Incorrect Adjective Order: 1
  - Missing Adjective: 2
  - Other Adverb Problems: 6
- **Incorrect Number on Noun**: 36
- **Problems with Noun Formation** : 6
- **Problems with Referent Formation** : 5
- **Pronouns**: 25
  - Incorrect pronoun choice (including pleonastic): 12
  - Inappropriate pronoun use (where full definite descriptions are required): 11
  - Lack of pronoun use (overuse of definite descriptions): 2
- **Pleonastic Pronoun Problems**: 20
  - Object Deletion: 4
  - Subject Deletion: 8
  - Incorrect Pleonastic Pronoun: 5
  - Other Pleonastic Pronoun Problems: 3
- **Redundancy Problems**: 7
- **Other Problems that may be related to Focus/Discourse Structuring**: 76
  - Noun Phrase Omission (subject: 16; object: 15): 31
  - Problems carrying over general/specific description strategies: 6
  - Structuring Problems with "because": 9
  - Other (may be related to topic-comment structures, or verbs of ASL): 25
  - Other: 5
- **Other Illegal Ellipses**: 10
- **Run-on Sentences**: 8
- **Idioms, word choice, hard to define problems**: 36

Table 1: Error Taxonomy

The numbers in this table are based on samples which have been entered into an on-line database, although we have analyzed many more samples. The numbers represent the number of sentences which have *at least one* error that *could be* classified as the corresponding error. Note that there are multiple error classifications of single errors in the

In fact, the majority of the error classes in our taxonomy could be explained by LT if one compares English and ASL on a feature-by-feature basis in the manner that we illustrated in section 3. It is important to note that we are not claiming that LT explains every error found in the written English of ASL natives. Nor are we claiming that because LT could explain errors of a particular class, that LT is the source of error for every instance of that class. There are other factors at work as well. For instance, an error may arise for the same developmental reasons that someone learning English as a native language might produce the error.[9] There are also production errors, including orthographical errors, that occur for reasons such as fatigue, carelessness, distraction, etc.

Obviously, the amount of exposure to a language influences the rate of acquisition of a language. For deaf ASL natives, exposure to English is limited to poor (or no) audio input, speechreading (or "lipreading"), and written English. In this particular case, the limited audio input may affect the rate of acquisition of certain morphological markers, such as verb endings for subject-verb agreement or tense. In fact, exposure to speechreading is not likely to help improve matters very much since "In casual speech as opposed to painstaking speech, approximately 40% of the phonemes are visible" (p. 243, [Swi89]), and "many of the sounds that are visible on the lips look identical ... and many vowels are confusable as well" (p. 243, [Swi89]). As a result, the speechreader must make use of linguistic and extralinguistic context, facial expression, gestures, knowledge of the language and pragmatics, world knowledge, and residual hearing to understand a spoken message by speechreading. Yet, if one has not acquired much of the spoken language, one can not make much use of linguistic context or of knowledge of the language in order to speechread and further acquire the language. Swisher [Swi89] gives reasons why learning a language from text is difficult. First, a child's first language is typically acquired in the course of conversations about concrete things and/or things salient to the present situation, and findings by Snow show that "semantically contingent speech" (p. 245, [Swi89]) is associated with more accelerated language acquisition. Yet, most written text is devoid of such rich contextual support. Also, "... written language does not have the information provided by intonation and stress patterns that may help the child to perform syntactic analysis, by signaling where constituent boundaries occur" (p. 245, [Swi89]).

Again, we believe that there are explanations based on LT for the majority of our the error classes in our taxonomy for the written English of the deaf, but we do not claim that LT can explain every instance of an error class that could be explained in terms of LT. In addition, we believe that even if differences between an L1 and L2 could not be the source of a particular error class, because writers of all L1 backgrounds produce that kind of error, these differences might still lead to longer persistence of that kind of error. In this way, L1 knowledge could still be said to influence L2 production of these errors.

4.1 Why LT in Deaf Writing might be Overlooked

We have noted that no one else (to our knowledge) has documented LT between ASL and English (or between any other visual-gestural language and a spoken language). There have been studies done which could erroneously lead one to believe that ASL has little impact on written English. For example, Langston and Maxwell performed a study in which "... 30 judges (professionals in audiology, speech pathology, deaf education, and language teaching) who were asked to distinguish holistically between the compositions of deaf signers, deaf nonsigners, and [English as a Second Language] students were unable to make the judgments correctly (Langston & Maxwell, 1988)." To conclude that ASL does not influence English based on such a study would be unwise. In fact,

[9] Perhaps these reasons are based in Universal Grammar, although we do not want to address such questions in this paper.

for the reasons enumerated below, we are not at all surprised that the judges could not reliably identify deaf signers from the other writers.

1. Recall that we look for evidence of LT on a feature-by-feature basis. Thus, if two languages share a feature in common which they do not share with the language being learned, we would expect that learners from the two different first languages would make similar errors with respect to that feature. For example, we would expect second language learners who know Chinese to have problems in English similar to ASL natives with regard to determiners and dropping verb tenses since neither ASL nor Chinese has articles and they both use time markers rather than verbal tense markings to mark time.

2. The kinds of errors produced by a native of a particular language in the production of a second language vary greatly according to his or her current stage of L2 acquisition. This means that while one might expect (on the basis of LT, for example) a particular error from a person with a particular L1 background learning a particular L2, the error might not occur because the learner has already acquired the L2 to a point where the error no longer occurs. Similarly, we might expect to not see a particular error although we've inadvertently misjudged that the writer has acquired the L2 enough to no longer make that error.[10]

3. In addition, since we think that L1 knowledge is not the only factor that influences L2 acquisition[11], writers from different language learning backgrounds are likely to share some errors in common, independent of the features shared by their languages.

Given the problems specified in 1 and 2, we would think it to be extraordinarily difficult to judge the language background of a writer based on a single writing sample. Instead, one might be able to hypothesize certain features of the learner's language (e.g., that it has no tense markings, no modals, no determiners). Because of the reasons enumerated above, and because one must account for production errors and test-taking strategies, it would seem difficult if not impossible to reach a conclusion about a writer's language background based on one sample.

Another interesting thing to note about the [LM88] study has to do with the backgrounds of the people who were judging the origin of the samples. The judges in the study were teachers, audiologists, and speech pathologists, who had "experience" with *either* deaf *or* foreign students, but not both. If someone is not familiar with all three[12] of these language backgrounds (deaf non-signers, deaf signers, and foreign students), he or she could not be expected to recognize errors specific to one population but not the other two. In fact, as we've already argued, there are several reasons why it would be difficult for someone familiar with several foreign language backgrounds, with the English of ASL signers, and with the English of non-signers to classify writing samples in the way required by the [LM88] study. Given how difficult distinguishing an L1 would be with all of this background knowledge, it should not be surprising that someone who did not have this knowledge would not be able to reliably classify writers according to their language backgrounds based on a single writing sample.

[10] This would not be surprising for two reasons: 1) the order of acquisition of features of an L2 for a given L1 is generally not well known, and is likely to be hard to determine for the reasons given in section 4.2 for why it is hard to make claims about the frequency of errors in the written English of ASL natives. 2) It is difficult to assess the stage of L2 acquisition of a writer for the reasons mentioned in 1, and this is particularly difficult if one has limited L2 data from the writer.

[11] In particular, we do not rule out developmental errors similar to those found for an L1, since factors in L1 acquisition may be factors in L2 acquisition.

[12] We use the number three here somewhat reluctantly, since obviously the number of language backgrounds is not three, but (at least) two plus the number of foreign languages known among the foreign students.

In closing on this matter, let us note that just as two languages need to be compared on a feature-by-feature basis to determine how LT between them might manifest itself, two languages need to be compared on a feature-by-feature basis with each other to see in what ways they might similarly and differently affect L2 production in a third language. Thus, to judge whether a writer is of one language background or another based on second language production is a difficult task. Furthermore, to judge the language background of a writer on the basis of a single writing sample is complicated further by the difficulties one faces when classifying errors (see section 4.2).

4.2 Frequency of Errors and Our Claims about LT

The purpose of Table 1 was to give the reader an idea of the kinds of errors that often occur in the written English of an ASL native, and some idea of the frequency with which the errors occur. However, we should note that the numbers should be considered with caution. First of all, for a given writing sample, it is often difficult to decide how to classify the errors in the sample for several reasons, some of which we will briefly note below. Secondly, generalizing error tendencies across writers is also difficult, as we will also discuss below.

4.2.1 Counting Errors in a Single Sample

- We often can not be certain of the writer's intent, particularly since other errors in the same sample may mislead us.

- One error might look like another. For example, if the writer doesn't know how to pluralize a particular noun, the resulting error may look the same as if the writer does not know that they need to (or how to) mark subject verb agreement. (For example, consider "The dress look pretty.")

- If the surface manifestation of an error is clear, the underlying source may still be unclear. For example, a verb that isn't properly marked for past tense could be the result of the writer not knowing how to mark tense or not knowing that they need to mark tense on that verb in that particular context. The limited language data that we have (usually only one or two writing samples per student) does not necessarily contain enough extra language data for us to be sure which error was made. (For the example given above, there may not be any other instances of the same verb (or a verb of the same kind) in a context where past tense would be required.)

- Errors also interact. If the writer omits an auxiliary, you often get an incorrect tense marking as well (e.g., "They _ become..." instead of "They had become..."). Should one count that as one instance of each kind of error or only as an instance of one of them?

- Another question is whether one should count every instance of "the same error". For example, if the writer fails to form the past tense of "drop", but correctly forms the past tense for all other verbs, do you want to count *every* instance of a missing tense on "drop" as an error?

- Another possible influence on the error frequencies in our samples is an artifact of the source of many of our writing samples: essays to be read and corrected by a teacher or essays for the purpose of evaluating the author's written English. Some writers may employ a test-taking strategy of using one guess about the correct form of a verb, adjective, noun, clause, sentence, etc. in one place, and using another guess for the same thing elsewhere in the essay to minimize the number of errors.

4.2.2 Generalizing across Writers

It is difficult to generalize about what kinds of errors and with what frequencies kinds of errors are made by deaf writers as a group. Depending on several factors, different deaf writers will make different errors (with different frequencies).

We believe that for two given error types and two given students, one error may account for a much higher percentage of the first student's errors than the second error, and the second error may account for a much higher percentage of the second student's errors. This may occur for many reasons, including differences in first languages (or in level of acquisition of a common first language), second language instruction, stage of second language acquisition, or hearing differences (e.g., perhaps certain morphological endings will be easier to acquire for someone with better hearing). Yet, to classify writers according to such differences in order to generalize about writers at a given stage of second language acquisition, with a common first language, common level of hearing loss, etc., is difficult since measuring many of these things (stage of second language acquisition, knowledge of a first language, etc.) is difficult.[13] As a result, to try to quantify error frequency for a population of writers is very difficult. So, to claim that deaf writers (even a particular group of deaf writers, e.g., students in grade x at school y) make a particular error which accounts for n% of the errors in their writing seems inappropriate.

4.2.3 Discussion of of Counting and Generalization Problems: Other Researchers will Face the Same Problems

The problems of how to count errors described above (with the exception of problems arising from test-taking strategies, perhaps) would be faced by any researcher trying to count errors in a set of writing samples (without interacting with the student). Unfortunately, the method of counting errors is rarely or never reported at this level of detail in literature on LT or the written English of the deaf. (We know of no research on LT or the written English of the deaf discussing the problems described above and how they addressed these problems.) So, while we urge the reader to use caution when considering the numbers in our table, we also urge the reader to use caution when reading other literature on the written English of the deaf or literature on errors in second language production.

4.3 Conclusions about LT in the written English of ASL natives

Despite the difficulties in estimating error frequencies for deaf writers, we believe that many errors in the written English of native signers of ASL fit the characterization of LT given above. Thus, we should use this characterization to verify we have appropriately classified errors, to hypothesize new error classes, and to decide what corrective information to give the user.

5 Carrying through with the Methodology: Uncovering Discourse-Level Errors that Arise from LT

The errors discussed in section 3 are intended to show how languages may be compared on a feature-by-feature basis in order to uncover cases where language transfer might explain a class of errors. The majority of the error classes we have found could be explained as resulting from language transfer if one compares features of ASL and features of English according to the type

[13] In fact, determining one's level of first or second language acquisition is difficult because of the problems in counting error frequency discussed above.

of analysis suggested by our characterization of LT and illustrated above. In many cases where LT is the (possible) source of an error, we believe that explaining the similarities and differences between the L1 and the L2 will be useful to the student.[14] This is one way that using knowledge about similarities among and differences between the learner's language and the target language would be useful in a CALL system.

5.1 The Need to Identify Discourse-Level Errors

The kinds of errors that we have discussed so far, what we term *sentence-level errors*, are not different in kind from those found in standard error taxonomies for systems correcting written English. The errors described so far are the kinds of errors that a CALL system would be able to identify by considering the (target language) rules for forming an individual *sentence* as an independent unit. These types of errors could be identified and corrected by a system that parsed text one sentence at a time, and searched for errors using a standard English grammar augmented with appropriate syntactic error rules in conjunction with a set of rules for finding semantic anomalies (such as those created by a lack of have/be distinction).

However, if a user of a second language is to gain proficiency in the language, not only must he or she have appropriate command of individual syntactic structures, but he or she must learn to use them appropriately *in context.* Wilbur [Wil77] has extensively studied the writing of deaf people and suggests that many errors she found could result from the writer not understanding when and how to use particular structures (such as relative clauses and pronouns). She suggests that this problem is the consequence of deaf education concentrating on the sentence level. This educational approach results in the students not being exposed to and appropriately tutored in discourse-level decision making. In the long run, a student may create texts that seem awkward even though the individual sentences are well-formed. Discussions with educators of the deaf and ASL linguistic researchers confirm that a skilled deaf writer may develop his/her writing skills to a point where he or she produces text which is not very cohesive, even though the individual sentences are grammatically correct and the content of the sentences is appropriate.

We believe that attention to discourse-level phenomena (e.g., text structuring and cohesion devices) should be included in any CALL system in order for the system to be truly beneficial to the user. Our conviction is supported by the number of errors in our samples that could only be appropriately explained by reference to discourse-level processing. We term such errors *discourse-level errors.* These are errors that either manifest themselves at the sentence level (resulting in a sentence which is ill-formed syntactically or semantically), or they may only be apparent in a longer stretch of text. In either case, these errors seem best explained in terms of the writer not having learned how to appropriately use the cohesive devices of English appropriately [HH76]. The errors which fall into this class range in severity from making the text seem "choppy" to making the text incomprehensible (due to the reader's inability to resolve pronouns, for example).

Table 2 contains a subset of the classes in our error taxonomy (derived from sample analysis) which fall into the category of discourse-level errors. Although, some of these errors appear to manifest themselves at the sentence level (e.g., dropped verb tense, referent formation, noun phrase omissions), we consider them to be discourse-level errors because the appropriate use of the associated devices requires an appreciation of the previous context. Thus, while some of the errors apparent at the sentence level may not require additional machinery to be *recognized*, we claim that their correction (and the appropriate tutoring related to them) would require discourse tracking mechanisms.

[14] See section 6.1.2 for further discussion of using information about the L1 in the instruction of an L2.

- **Inappropriate determiner or determiner formation:** 17
- **Dropped Verb Tense:** 6
- **Relative Clauses**: 27
- **Redundancy Problems**: 7
- **Pleonastic Pronoun Problems**: 20
  - Object Deletion: 4
  - Subject Deletion: 8
  - Incorrect Pleonastic Pronoun: 5
  - Other Pleonastic Pronoun Problems: 3
- **Run-on Sentences**: 8
- **Problems with Referent Formation** : 5
- **Pronouns**: 25
  - Incorrect pronoun choice (including pleonastic): 12
  - Inappropriate pronoun use (where full definite descriptions are required): 11
  - Lack of pronoun use (overuse of definite descriptions): 2
- **Other Problems that may be related to Focus/Discourse Structuring**: 76
  - Noun Phrase Omission (subject: 16; object: 15): 31
  - Problems carrying over general/specific description strategies: 6
  - Structuring Problems with "because": 9
  - Other (may be related to topic-comment structures, or verbs of ASL): 25
  - Other: 5
- **Other Illegal Ellipses**: 10

Table 2: Errors that may be related to LT of discourse-level strategies

Other errors in the table would be apparent only if the discourse level was considered (e.g., inappropriate determiner formation, description strategy problems). These classes of errors would require discourse machinery for their detection as well as for correction and tutoring. For instance, the use of a definite (as opposed to an indefinite) determiner is dependent on whether or not the referent of the noun phrase has already been introduced in the discourse. Therefore, the error cannot be detected unless the system has a way of knowing what has been introduced.

5.2 Does LT play a Role in Discourse-Level Errors?

Given the quantity of discourse-level errors that we have found, one must question how they should be addressed by the language instructional system. In particular, the question of whether or not (many of) these discourse errors are explainable in terms of LT in a way similar to the sentence-level errors discussed above arises. In addressing this question, we extended our methodology for uncovering language transfer to comparing languages on a feature-by-feature basis at the discourse level (as we did in section 2.2) where the features we consider are discourse features which affect the discourse level. That is, we proposed that a language learner may not only be transferring lexical items and syntactic rules, but may also transfer *discourse strategies* and cohesion marking devices from one language to another. In the following sections, we give evidence that this is the case by examining three discourse features (used for subsequent or anaphoric[15] reference) that are handled differently in ASL and written English. We submit that several classes of errors that we have found can be explained in terms of the user carrying over information about when these features can be used from ASL to English.

In particular, in the following sections we argue that some of the error classes involving pronouns (incorrect pronoun choice and inappropriate pronoun use) and many illegal NP omissions can be explained in terms of LT when the discourse feature of subsequent reference is examined more closely.

5.3 Subsequent Reference: An Introduction to Three Language Features

The devices available to a user of a language to refer to items already introduced into the discourse are quite varied, and the rules for their appropriate use can be quite complicated. Notice that *the form* of subsequent reference may serve several purposes in the discourse in addition to picking out a referent (i.e., an object, event, etc.). For example, the variety of subsequent reference tools available in English serve as cohesive devices (e.g., to help indicate a continuance or a change in topic) in addition to allowing the user to pick out the appropriate referent. We argue that the different roles that subsequent reference devices might play in two languages is a potential source of error for second language learners.

We consider a subsequent reference device to be a feature of a language. Languages commonly have three subsequent reference features: full NP's, pronominal NP's, and zero pronouns. Of course, both ASL and English have the feature of full noun phrase reference. We will not be explicitly discussing this feature since by looking at the features of pronominal reference and zero pronominal reference for ASL and English, we will implicitly be discussing when full noun phrase reference is used for each of the languages.

We will first discuss the feature of pronominal reference in English; that is *how* pronominal reference is achieved and *when* pronominal reference can/should be used in English. Then we will

[15]By anaphoric reference, we mean any reference to an item that has been introduced in previous discourse. Notice that this use of the term is broader than reference to reflexive pronouns, as found in much of the Government and Binding literature.

discuss *how* pronominal reference is achieved and *when* pronominal reference can/should be used in ASL. The differences[16] between the languages in terms of when pronominal reference can/should be used in the languages allows us to explain some pronominalization errors we have found in terms of language transfer.

Then, we will discuss a kind of of zero pronominal reference that is used in ASL but not in English. Our claim is that transfer of knowledge about when zero NP's can be used in ASL may explain some of the illegal NP omissions in the written English of ASL natives. We give examples of such omissions.

5.4 Pronoun Use in English

While it may or may not be apparent to a native speaker of English, the rules for when and when not to pronominalize a noun phrase in English are very complicated and dependent on several factors. These factors include the semantics of the items in the discourse, the semantics of the verb that an NP is an argument of, pronominalization and focus history of entities in previous discourse, intended focus of entities in the current sentence, pragmatics, world knowledge, knowledge shared by the discourse participants, etc. To illustrate the effects of some of these factors, consider the following examples.

1a John gave Sarah an apple.

1b She cut it up.

2a John gave Peter an apple.

2b He cut it up.

In 1b, the semantics of "she" and that Sarah is the only female in the preceding discourse, allow Sarah to be pronominalized with no ambiguity.

However, in 2b, "He" could refer to either "John" or "Peter", and thus pronominalization in 2b may seem ambiguous to the reader and is thus less appropriate. In fact, because things are normally presented in temporal sequence in English, and because (typically) once a person has given an object to somebody else we don't expect the first person to do anything more with the object, it is likely that a reader would prefer Peter as the referent of "He". Nonetheless, because of the possible ambiguity of "He", either the use of a fuller description or a different sentence structure (e.g., "John cut up an apple and gave it to Peter" or "John gave an apple to Peter who cut it up") would be clearer.

We can also use 2a and 2b to illustrate how pragmatics and shared knowledge influence pronominalization: The "He" in 2b would seem to unambiguously refer to John to most readers, if both the writer and the reader knew that "John" referred to an adult, and "Peter" referred to an 18-month old child.

Tense may also influence the ability to appropriately resolve pronouns in English. Consider:

3a John gave Peter an apple.

3b He had cut it up.

[16] As discussed in section 5.5, this mismatch is partly based on differences in how pronominal reference is achieved in the languages, since pronominal reference in ASL is inherently unique.

While 3a and 3b (in neutral contexts in terms of knowledge about John and Peter) are somewhat ambiguous just as 2a and 2b are, the reader probably prefers the reading that John had cut the apple up, not Peter. The use of the past perfect tense to shift the focus back to a time prior to when John gave Peter the apple makes it more likely that John cut the apple up.

Another influence on whether or not a pronoun is appropriate is whether or not a clear focus indicator has been given ([Sid83], [Sur92b]). For instance, consider:

4a There was a new girl in Mary's class yesterday.

4b She acted very friendly.

In this case, the form of the first sentence (a *there*-insertion) is an indicator that the new girl is very prominent and allows one to use a pronoun to refer to her even though Mary would be a reasonable referent for "She" in 4b in a non-marked context. In addition, the use of a pronoun in 4b indicates that the discourse is probably going to continue being about the new girl, and this indication is important for understanding subsequent sentences. Notice it would be odd for the discourse to continue from this point with a discussion of Mary with no further reference to the new girl.

While we have only given a few extremely simple examples here, the point is that when to use a pronoun and how a pronoun is resolved in English is fairly complicated. There is a small number of pronouns and these carry some semantic information about their referent (e.g., gender and number). Pronoun use in English is dependent on the focus or what has been made prominent in a discourse. The use of a pronoun seems to be an indicator of continued prominence of the referent. However, there are some situations where a pronoun cannot be effectively used even though focus rules might favor a pronoun. These are cases of potential ambiguity resulting from the presence of other items in reasonably high focus in the discourse with the same gender and number features.

5.5 Pronoun Use in ASL

Let us now turn to pronominal reference in ASL. In ASL, a *locus* or location may be associated with an object, concept, event, etc., and then pronominal reference to that referent can be achieved by indicating the locus. A locus may be indicated by pointing or gazing at the locus, or by using that locus as the starting or end point of an inflecting verb.

For a referent that is present, the locus is the location of the referent itself. For referents which are not present, each referent may be associated with a locus in the signing space in front of the signer's body. "This is accomplished by producing the sign for the referent at some arbitrary locus in space, or making the sign and then pointing to the locus with the index finger, or by eyegaze in the direction of the locus while making the sign." (p. 25-6, [LM91]) Abstract loci associations persist in discourse until a new framework is established, and the number of such loci, while theoretically unlimited, typically does not exceed 5.

5.6 Some Differences between Pronominal Reference in ASL and English

In some sense, the use of a locus (for pronominal reference) in ASL is analogous to using pronouns in English because both are forms of (radically) reduced reference. However, the similarity stops there. Perhaps the most important difference is that in ASL pronominal reference is the normal form of reference to an item that has already been mentioned. Lexical repetition is used in a very limited fashion in ASL and when it is used it seems to serve the function of reaffirming the present topic (p. 235, [WP83]). In English, several factors influence whether or not something is referred to pronominally. For instance, the use (and comprehension of) pronominalization is highly dependent on whether or not (and to what extent) that item has been in focus in the preceding discourse. In fact, in English, the use of a pronoun (as opposed to lexical repetition) frequently serves the function of reaffirming the present topic.

5.7 Pronoun Errors that can be Attributed to LT

Many errors in the written English of deaf writers can be explained as resulting from language transfer because of the the differences between the use of pronominal reference in ASL and English. Below we discuss three excerpts with pronominalization errors. (Note that these three excerpts[17]came from three different writers.)

1. "Fraternities and Sororities here at XYZ DO provide [a] social life. Some examples: parties; gettogethers; workplaces; IM; and sports. At parties that *they* host..."

   Possible correction: "At parties that *the fraternities and sororities* host..."

In this excerpt, the pronoun "they" in the final sentence could not refer to anything more recent than "Fraternities and Sororities", based on semantics (people or organizations host parties; parties, gettogethers, workplaces, IM or sports do not host parties). Nevertheless, the pronoun resolution of "they" is difficult for a native speaker of English because the focus of the discourse has shifted away from fraternities and sororities to various kinds of social events. Notice that in ASL this shift would not preclude the use of a "pronoun" (locus) to refer to fraternities and sororities in this instance. Thus, a possible explanation of this error is that the writer does not understand the relationship between focus and pronoun use in English since that relationship is not present (or is very different) in ASL. This possibility is further supported by the fact that the corresponding discourse in ASL would use pronominal reference to refer to *fraternities and sororities.*

2. "Abolishing fraternities and sororities at XYZ would be a very unwise thing to do. If so, no fun; no unity; and no social life would result. Being a Greek is what every student dreams of when they enter college. If we abolish *them*, that dream will never come true. Not only should Greek organizations stay open due to traditional reasons, *it* offers a wonderful bond between [...] brothers and sisters. For an example, my close friend, Amy, is a member of a sorority here at XYZ. I have seen the closeness she and her fellow sisters [have] and I envy that. Joining an organization gives you wonderful personal growth. *They* can help you improve in your areas of weaknesses."

In this excerpt, we have italicized three pronouns which have been used incorrectly according to the intuitions of a native speaker of English. We will consider each of them in turn. The first error involves the *them* in "If we abolish *them*..." The most likely correction for this is "If we abolish *fraternities and sororities*," which would be consistent with what the essay is about. This error is analogous to the error described in the previous excerpt. In particular, the writer doesn't recognize that *fraternities and sororities* are no longer prominent enough to be referred to pronominally at this point in the discourse.

The second error involves the *it* in "... *it* offers a wonderful bond..." This error can be explained in two different ways, depending on what the writer is trying to refer to. The first possibility is that the writer has made a pronoun choice ("it" for *Greek organizations*) which is incorrect since

[17] In some cases, the excerpts we are showing have corrections and/or deletions in them which are indicated in square-brackets. Our corrections do not affect pronoun resolution. Also, we have italicized the words of a sentence which are related to the error classification being discussed.

"it" and *Greek organizations* do not agree in number. The second is that the writer is trying to refer to *being a Greek* with the pronoun "it". This would be an incorrect usage since the notion of *being a Greek* isn't prominent enough to be referred to pronominally (in English) at this point in the discourse.

For the third error ("*they* can help you"), there are several things the writer could have intended to refer to, but each of these is problematic. For example, the most recent NP that could co-refer with "they" in terms of number agreement is *she and her fellow sisters*. However, the sentence would not make sense pragmatically if that were the intended referent. On the other hand, the writer could have intended to refer to *the fraternities and sororities*, *(Greek) organizations*, or *the members of fraternities and sororities*). However, these items are fairly far back in the discourse and no longer prominent enough for pronominal reference. Alternatively, the writer could have intended to refer to *joining an organization* or *being a Greek*. However, then there is a number feature mismatch between the intended referent and the pronoun.

3. "My father hired me to run for *my dad*."

   Possible Correction: "My father hired me to run for *him*."

The original sentence is ungrammatical because English requires an NP to be pronominalized if its referent occurs previously in the same sentence (in most circumstances).[18] Here, the writer may be avoiding pronoun use because he or she has recognized that English does not allow pronominal reference in situations where ASL does, but he or she isn't sure when pronominal is allowed (or required) in English.

5.7.1 Conclusions about Pronominalization Errors

In the above examples, we have seen three possible causes for errors involving pronoun use. The first is that the writer does not appreciate that in English the form (i.e., number and gender) of pronoun may not be enough to fully specify its referent, although in ASL a referent is uniquely identified by the locus of the pronominal reference. Such a misconception on the part of the writer may result in the use of a pronoun which is ambiguous because it has two or more possible referents. (In fact, a text might seem awkward even if the reader could ultimately figure out what the intended referent of the pronoun was, if the use of pronominalization doesn't match the reader's expectations about what would most likely be referred to pronominally at that point in the discourse.)

The second possible cause of errors is that the writer does not recognize a mismatch between the number and/or gender of the intended referent and the number and/or gender of the pronoun. This could result in an incorrect choice of a pronoun. The third is the that writer does not appreciate the role of focus in pronominalization, and thus he or she has tried to refer pronominally to an NP that is no longer prominent enough to be referred to pronominally.

These sources of error can be explained by differences in the nature of pronominalization (since pronominal reference uniquely identifies a referent in ASL[19] but not in English), and by differences between the rules of when to use pronominalization in ASL and English. Errors involving pronoun use are likely to persist until the writer learns the relationship between focus and pronominal

[18] More accurately, the original sentence is ungrammatical because a full NP can not be bound. The correction is acceptable since a non-reflexive pronoun must be bound, but must be free in its governing category.

[19] While, informally, our informants indicate that referents assigned loci in space generally remain available for the duration of the conversation, we know of no formal studies documenting the length of their persistence, or reuse of a locus for a different referent during a different discourse segment. Baker and Cokely note that there is "very little known about the rules used by ASL signers when setting up multiple referents in space and how these referents are moved into different space to reflect different 'scenes' in a narrative." (p. 227, [BC80])

use in English, in addition to understanding how/what semantic features are captured by English pronouns.

5.8 Empty Categories in ASL and English

Another form of subsequent reference used in ASL involves the use of *Empty Categories* (EC's) or zero NP's. An empty category is a syntactically real, but phonologically null NP. Here we concentrate on the EC's that occur with *plain verbs* in ASL [LM91]. The plain verbs of ASL have no morphology that would help the "hearer" recover an implicit referent. However, these verbs do allow null arguments (and, thus, empty categories) in some contexts. Lillo-Martin [LM91] argues that the deletions with respect to these ASL verbs are similar to deletions in languages (like Chinese) which are termed *discourse-oriented* languages. In these languages, the deleted NP can be recovered from context. Her analysis of deletions/EC's with respect to plain verbs in ASL is based on Huang's analysis of EC's in Chinese. Specifically, Lillo-Martin argues that ASL allows *Topic NP Deletion*, i.e., for the topic of a sentence to be "deleted under identity with a topic of a preceding sentence" [Hua84].

Although English does have some forms of empty categories, it does not have the feature of allowing Topic NP Deletion, and so it does not have the corresponding zero NP's.

5.9 NP Omission Errors that can be Attributed to LT

ASL has the feature of allowing the rule of Topic NP Deletion, but English does not have this feature. As a result, ASL has zero NP's that arise from the deletion of an NP that co-refers with the topic of a previous sentence (or a discourse topic), but English does not. Therefore, one might expect to find missing noun phrases in the written English of ASL natives, as the result of transfer of the Topic NP Deletion rule from ASL to English.

In fact, we find such omissions. Consider the following examples of English text (written by deaf writers) with NP omissions:

1. "I think that XYZ College should require all deaf students to take speech and speechreading courses. Therefore, they can improve their oral skills for their future use. I am going to tell you [...] why the deaf student should take _."

2. "There are many things I like about XYZ. They offer supporting services like interpreters and notetakers for mainstream classes which I had experiences [with in] my public schools. Now XYZ/PDQ offers [the] same thing that my school offered but [... with] better supporting services. That is [what] I like about XYZ.

   But [the] one thing [that] worries me [the] most about XYZ/PDQ is financial problems. I hope I could find some ways to solve _."

3. "First, in summer I live at home with my parents. I can budget money easily. I did not spend [a] lot of money at home because at home we have [a] lot of good foods, I ate [a] lot of foods. While living at college I spend lot of money because _ go out to eat almost everyday. At home, sometimes my parents gave me some money right away when I need _. While in college, I could not ask my parents for money right away because I live in XYZ and my parents live in QRS. It is too far."

Each of the above examples is missing (one or more) noun phrases and, as a result, the texts are ungrammatical in English. Informally, we can see that each of the deletions in the above excerpts can be explained as resulting from transfer of the rule of Topic NP Deletion. For example, in the last sentence of the first excerpt, it seems the writer has deleted a referent to *speech and speech reading courses* since the courses are the topic of the first sentence (and perhaps the second sentence, if "they" is taken to refer to *speech and speech reading courses* and not to *deaf students*). In the second example, it seems the writer has deleted a referent to *financial problems* since financial problems are the topic of the previous sentence.

We note that for each of the above examples, discussions with ASL informants indicated that the corresponding ASL discourse would be acceptable/grammatical if the omitted NP were not signed, pronominally referenced, or indicated by verb agreement and that the omitted items would be understood from the context. This supports our proposal that these errors can be explained by transfer of the Topic NP Deletion rule from ASL to English.

5.10 LT at the Discourse Level: Summary

Note that while researchers have made predictions about how L1 might affect L2 at the discourse level (e.g., [Rut83], [Rut89]), the limited documentation of discourse transfer (e.g., Koch [Odl89], Bartelt [Odl89], Jordens [Jor83]) has mostly dealt with how the discourse strategies of L1 manifest themselves at the *sentence* level of L2. Some examples of such transfer that have been reported are the overuse of topic-comment structures by Mandarin speakers acquiring English ([Rut83]), and the overuse of repetition by Navajo and Apache students learning English (Bartelt [Odl89]) and repetition of words and phrases in English by Arabic speakers learning English (Koch [Odl89]).

We have extended the notion of LT to discourse-level transfer that involves L2 production errors that manifest themselves at the discourse level, i.e., errors that have implications on how to process subsequent sentences and/or require context to be detected.[20] In the case of ASL and English, we have identified errors in the written English of ASL natives involving the use of an inappropriate pronoun arising from differences between the languages in terms of the interaction of focus with *when* pronominal reference is used (and *how* pronominal reference is achieved); we also identified illegal NP omissions resulting from the presence of the feature of Topic NP Deletion in ASL but not in English.

6 The Design of Our CALL System and the Role of LT in the Design

We have given a characterization of LT, and we've discussed how this characterization can be used to explain many of the sentence-level and discourse-level errors in the written English of ASL natives. We now explain how we plan for our system to identify and correct errors in the written English of ASL natives, and how our LT characterization and methodology affects the design of our system. This discussion should illustrate the usefulness of comparing an L1 and L2 on a feature-by-feature basis and why our methodology would be useful for the design of other CALL systems.

6.1 How our System will Identify and Correct Sentence-Level Errors

Our system will rely on a grammar of English which has been augmented with a set of syntactic and semantic *meta-rules* ([WVJ78], [WS83], or *mal-rules*, in [Sle82] terminology) which extend the language accepted by the parser and semantic interpreter to include the types of errors we expect.

[20]One might argue that using a pronoun where a fuller NP is required or using an inappropriate pronoun could be viewed as a sentence level manifestation of an error resulting from errors in use of discourse strategies. What we want to stress is that such errors have discourse-level implications in terms of how to resolve pronouns in *subsequent* sentences.

That is, in addition to rules which can parse grammatical English sentence, we will include meta-rules which allow the system to parse ungrammatical sentences. The use of a particular meta-rule causes the system to record the nature of the errors that the meta-rule allows the system to parse. (Of course, the system will first try to parse the user's text using only the grammatical rules of English, and only try to apply the meta-rules if that fails.) By the use of meta-rules, our CALL system will be able to identify sentence-level errors in the user's text.

Our characterization of LT was important in the development of our taxonomy for errors at the sentence-level. Once we noted that LT could explain many sentence-level errors, we used this information to determine how we should classify our errors more appropriately. For example, as mentioned several times before, by virtue of considering LT we recognized that sometimes a missing tense marking is not necessarily the result of the learner not knowing how to mark tense on that particular verb, but a result of the learner not knowing that tense needs to be marked on the verb in the current context. We also recognized that many awkward sentence constructions (and awkward stretches of text) seemed to be related to an attempt to use a topic-comment strategy. Similarly, we recognized that many awkward sentence constructions (and stretches of text) seemed to be related to an attempt to describe things generally and then give the specifics, without appreciation for how English connects clauses and sentences together.

6.1.1 Appropriate Correction

The appropriate classification of an error ensures that we are able to provide an appropriate correction to the user. For example, if we believe the writer omitted a tense marking on a particular verb because he or she did not realize that tense needed to be marked on that verb, we should give the user corrective information about how although time need not be marked in every sentence in ASL, time must be marked on every verb in (a finite clause in) English. If we only gave the user information about how to mark tense on that verb (which he or she might already know, although we may not be able to determine that from the rest of the writing sample), the user is not going to benefit from our correction.

Similarly, if we find missing pluralization markings on nouns, we need to tell the user that every NP in English must be marked appropriately for number, unlike in ASL where once you mark something as plural, subsequent referents may be in the singular form [Pad81]. (Also, although requirements about marking plurality in ASL are still not well understood, it seems to be the case that when one uses a quantifier like $\overline{\text{SEVERAL}}$, the lexical item may not need to otherwise be marked for plurality.[21] Thus, correction of missing plural markings is likely to be complicated and dependent on several factors.)

Another example of how corrective information should take information about ASL into account is for the correction of errors that might result because ASL has one lexical item for multiple concepts expressed with multiple lexical items in English. For example, when correcting "I am interesting in ASL", we should note that ASL uses one sign, $\overline{\text{INTEREST}}$, for "interest", and "(be) interesting", "(be) interested". This would cue the learner to look for how these different English constructions are used.

[21] We use capital letters and overbars to mark a "translation" or a gloss of an ASL sign. Here, we use $\overline{\text{SEVERAL}}$ to convey a sign which could be translated to the English word "several." (To properly convey all linguistic information of ASL, we would need to use a more sophisticated notation, but this notation is adequate for the present purposes.)

6.1.2 What if LT isn't the Source of an Error?

From the previous discussion, we can see that if we did not use our characterization of LT to classify errors and to decide what corrective information to provide users, the instruction we gave the user would be less appropriate.

However, one must ask how our characterization is useful for those instances of errors or classes of errors that are not explained by the influence of knowledge of the L1 on the production of the L2. In fact, even if a particular class of errors does not have its basis in the writer's knowledge of the L1, a contrastive comparison of an L1 and L2 might be useful in L2 instruction for a particular class of errors. The idea of using L1 information in L2 instruction has been suggested by others as well (e.g., [Swi89]). We believe, *more specifically*, that pointing out differences between the L1 and the L2 in terms of when, how, or whether a feature is marked, or pointing out the fact that the lexical mapping between the L1 and L2 is not a one-to-one mapping may help the student learn the L2. Thus, our characterization can help the CALL system designer better identify what corrective information to provide the user.

6.2 Using LT to Predict, Identify and Correct Discourse-Level Errors: the Need for Tracking Focus

Our characterization of LT and knowledge about ASL and English was important for our understanding of several discourse-level errors in the written English of ASL natives. We were able to understand the source of inappropriate pronoun use where a fuller NP is needed by considering *when* pronominal reference is used in ASL (e.g., recall the inappropriate use of "they" to refer to *fraternities and sororities* in the first excerpt of section 5.7). By examining Topic NP Deletion we were able explain omitted NP's.

We now turn to the question of what would be required for a CALL system to be able to (detect and) correct these errors effectively. The first subclass of errors that we looked at can broadly be termed inappropriate pronoun use. There seemed to be two major underlying causes for these errors. The first is that the user didn't appreciate the role that focus of attention plays in using pronouns in English. This would explain the errors that involved using a pronoun even though the focus of attention had shifted away from the referred-to item. The second underlying cause is that the user did not appreciate the semantics associated with the English pronoun. This last factor would explain the use of an incorrect pronoun, such as the use of "it" to refer to *fraternities and sororities* as discussed in the second excerpt from section 5.7.

The second subclass of errors we discussed involved the carry-over of Topic NP Deletion from ASL to English. Recall that this rule allows an NP to be deleted if it is the topic of the current sentence and it is coreferential with a previous topic.

Thus, in handling both classes of errors, we ([Sur92a], [Sur92b]) propose that a crucial piece of information that must be tracked is the topic or, in computational linguistic terms, the *focus* of a discourse. There have been several algorithms proposed which attempt to track focus, or the thing that the sentence (and subsequent sentences) is intended to be most centrally about ([Sid79], [GJW83],[Sid83], [BFP87], [DB90], [Sur92b], [Sur92a]). There are several factors that influence what the algorithms take to be in focus; however, surface structure (syntactic form), the use of pronouns and other anaphoric devices and what their referents are, and what has been focused in the past are particularly important. Thus, a current emphasis in our work is developing a focus tracking algorithm for use in our CALL system. The algorithm we are developing most closely follows [Sid79], [Sid83] and is described in [Sur92b], [Sur92a].

Notice how the focus tracking algorithm can be used. For pronouns, a focus tracking algorithm would allow us to flag pronouns that are used without appreciation of the relationship between focus and pronominalization in English. This would be achieved by our algorithm detecting that a pronoun could not be appropriately resolved. The choice of an incorrect pronoun would be flagged in a similar way.

For the illegal NP deletions, the focus tracking algorithm could be used to recover the intended referent of the deleted NP. This recovery is crucial for the correction and tutoring phase of the system. The basic idea in the recovery process is to treat the deleted NP as highly indicative of focus (since the Topic NP Deletion rule allows topics to be deleted under co-reference with a previous topic), and to attempt to find the referent of the deleted NP using items of past focus.

Through their role in the correction of these and other classes of discourse-level errors, our LT characterization and a focus tracking algorithm can add functionality to CALL tools for second language learners.

7 Conclusions

In this paper, we proposed a methodology for designing a taxonomy for a CALL system for correcting the written text of second language learners. This methodology was based on a characterization we provided for language transfer. The characterization of LT was motivated by our findings of LT between ASL and English at the sentence level, and captures a very broad definition of LT: that of the influence of knowledge of one language on the comprehension and/or production of another. Our LT characterization proposes that language transfer might occur when two languages differ with respect to whether they have a particular feature, how they mark a particular feature, when they mark a particular feature, or as a result of the mapping between the lexical items not being one-to-one. Other researchers have documented LT between languages, but we have identified phenomena that requires one to extend the notion of LT beyond that previously considered. In particular, we showed that differences in language structuring and cohesion strategies must also be examined as potential sources of discourse-level errors.

We have argued that during the design of a CALL system for second language learners it is useful to compare the native and target languages on a feature-by-feature basis in the manner suggested by our LT characterization. There are three important ways that our characterization of LT can be useful in the design of a CALL system.

- One is likely to more appropriately classify errors if one takes our characterization of LT into consideration. The correct classification of errors is of primary importance in a CALL system because the better the classification is, the better the correction the system provides will be.
- Our characterization of LT enables the designer to identify error classes which may be overlooked otherwise as an unconnected set of errors.
- Our characterization of LT suggests information to be included in the corrective response given to the user.

We also argued that discourse-level features cannot be ignored by a CALL system that is to be effective. If the discourse-level is ignored, several kinds of errors will likely persist. While some of these errors (e.g., discourse structuring errors) may be deemed relatively unimportant because they are at such a high level, others (e.g., inappropriate pronoun use and omitted NP's) must be addressed if the text is to be considered grammatical.

By extending the notion of LT to the discourse-level, the methodology used at the sentence-level can be used to uncover classes of discourse-level errors and suggest what kind of corrective information to provide the user in connection with these errors.

We have noted one piece of discourse information that will be helpful in detecting and correcting many of the discourse errors we have found: the local focus of the discourse. By adding the ability to track focus and use focus information appropriately, a CALL system will have the ability to deal with several of the discourse-level errors we have identified as being prominent. Other discourse tracking machinery may also be necessary in handling other discourse-level errors.

To illustrate the usefulness of our methodology and LT characterization, we discussed the design of a CALL system to help ASL natives learn written English. The design of this CALL system both motivated our methodology and was in turn affected by the methodology.

In this paper, we reported a taxonomy of errors that we found in the written English of deaf natives. The purpose of the writing sample analysis was to identify patterns of errors that are commonly made by (and should thus be expected from) the deaf population. Our initial analysis led us to hypothesize that language transfer could explain errors in the written English of ASL natives. We developed a language-independent characterization of LT, and the use of this LT characterization helped us to more appropriately classify the errors that we found.

We also discussed three discourse features, pronominal reference, zero pronominal reference, and full noun phrase reference, which differ in ASL and English. We provided examples of inappropriate use of pronouns, incorrect pronoun choice and illegally omitted noun phrases in the written English of the deaf, and discussed how these examples could be explained by the differences between ASL and English with respect to those discourse features. Again, we proposed using a focus tracking mechanism for identifying and correcting these errors ([Sur92a], [Sur92b]).

It should be evident from this paper that our characterization of how two languages may differ from each other (and subsequently lead to LT) helped in the design of our CALL system. We have discussed ways in which our LT characterization helped us explain some sentence-level errors, how the characterization helped us more appropriately classify some sentence-level errors, and the impact of the appropriate classification on the corrective information to be provided to the user. We have also explained how the characterization helped us explain some discourse-level errors. We believe our language-independent LT characterization and system design methodology should be useful for the design of other CALL systems for second language learners. Furthermore, we believe that the development of an appropriate error taxonomy is crucial for the development of any effective CALL system.

8 Acknowledgments

We would like to thank John Albertini of the National Technical Institute for the Deaf (NTID), MJ Bienvenu of the Bicultural Center, Helen Craig and Patty Carnes of the Western Pennsylvania School for the Deaf, Bob McDonald of Gallaudet University, Stephen P. Quigley (formerly) of the University of Illinois, Lore Rosenthal of the Pennsylvania School for the Deaf, and George Schellum (formerly) of the Margaret S. Sterck School, for helping us gather writing samples.

A good part of our knowledge of ASL comes from discussions with ASL signers. We would like to thank our informants, April Nelson of Rosemont College, Don Ruble of Bloomsburg State College, and Jean Quillen and Carmine Salvato of the Pennsylvania School for the Deaf. We also thank Lore Rosenthal of the Pennsylvania School for the Deaf for interpreting for us.

In addition, we want to thank the numerous people from deaf schools and organizations who have discussed this project with us.

We thank Julie Van Dyke for her implementation of the English grammar, Jeff Reynar for his implementation of the error productions, and Karen Hamilton for implementing a database for us to store and access our analyzed writing samples. We thank K. Vijayashanker for many useful comments and helpful discussions about the content and organization of this paper. We also thank Anil M. Shende, John Hughes, Mark Jones and Chris Pennington for their comments.

References

[ACRS85] Khurshid Ahmad, Greville Corbett, Margaret Rogers, and Roland Sussex. *Computers, Language Learning and Language Teaching.* CUP, NY, 1985.

[Art85] Artificial Intelligence and Simulation of Behavior. *VP2: The Role of User Modelling in Correcting Errors in Second Language Acquisition*, 1985.

[Bak80] C. Baker. Sentences in American Sign Language. In C. Baker and R. Battison, editors, *Sign Language and the Deaf Community*, pages 75–86. National Association of the Deaf, Silver Spring, MD, 1980.

[BC80] C. Baker and D. Cokely. *American Sign Language: A Teacher's Resource Text on Grammar and Culture.* TJ Publishers, Silver Spring, MD, 1980.

[BFP87] Susan E. Brennan, Marilyn W. Friedman, and Carl J. Pollard. A centering approach to pronouns. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*, pages 155–162, 1987.

[BP78] C. Baker and C. Padden. Focusing on the non-manual components of American Sign Language. In P. Siple, editor, *Understanding Language through Sign Language Research*, pages 27–58. AP, NY, NY, 1978.

[BPB83] K. Bellman, H. Poizner, and U. Bellugi. Invariant characteristics of some morphological processes in American Sign Language. *Discourse Processes*, 6:199–223, 1983.

[BS88] C. Baker-Shenk. Comparative linguistic analysis for interpreters. In D. Cokely, editor, *Sign Language Interpreter Training Curriculum*, pages 84–108. Fredericton, NB: University of New Brunswick, 1988.

[CJ84] Carol Chapelle and Joan Jamieson. Language lessons on the PLATO IV system. In David H. Wyatt, editor, *Computer-assisted language instruction*, pages 13–20. Pergamon Press, NY, 1984.

[DB90] Deborah Dahl and Catherine N. Ball. Reference resolution in PUNDIT. Technical Report CAIT-SLS-9004, UNISYS, Paoli, PA, March 1990.

[Fan83] Lou Fant. *The American Sign Language Phrase Book.* Contemporary Books, Inc., Chicago, 1983.

[Gas79] Susan Gass. Language transfer and universal grammatical relations. *Language Learning*, 29:327–344, 1979.

[Gas84] S. Gass. A review of interlanguage syntax: Language transfer and language universals. *Language Learning*, 34(2):115–132, 1984.

[GJW83] Barbara J. Grosz, Aravind K. Joshi, and Scott Weinstein. Providing a unified account of definite noun phrases in discourse. In *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, pages 44–50, Cambridge, Mass, June 1983.

[GS83] S. Gass and L. Selinker, editors. *Language Transfer in Language Learning.* Newbury House, Rowley, MA, 1983.

[Hak76] K. Hakuta. A case study of a Japanese child learning English as a second language. *Language Learning*, 26:321–51, 1976.

[HH76] M. A. K. Halliday and Ruqaiya Hasan. *Cohesion in English.* Longman, London, 1976.

[HS83] R. J. Hoffmeister and C. Shettle. Adaptations in communication made by deaf signers to different audience types. *discourse processes*, 6:259–274, 1983.

[Hua84] C.-T. James Huang. On the distribution and reference of empty pronouns. *Linguistic Inquiry*, 15(4):531–574, Fall 1984.

[Ing78] R. M. Ingram. Theme, rheme, topic and comment in the syntax of American Sign Language. *Sign Language Studies*, 20:193–218, Fall 1978.

[Jor83] Jordens. Discourse functions in interlanguage morphology. In S. Gass and L. Selinker, editors, *Language Transfer in Language Learning*, chapter 20. Newbury House, Rowley, MA, 1983.

[KB79] E. S. Klima and U. Bellugi. *The Signs of Language.* Harvard University Press, Cambridge, MA, 1979.

[KG83] J. Kegl and P. Gee. Narrative/story structure, pausing and American Sign Language. *Discourse Processes*, 6:243–258, 1983.

[KK78] Richard R. Kretschmer Jr. and Laura W. Kretschmer. *Language Development and Intervention with the Hearing Impaired.* University Park Press, Baltimore, MD, 1978.

[KK90] M-M. Kenning and M.J. Kenning. *Computers and Language Learning: current theory and practice.* Ellis Horwood Limited, England, 1990.

[Kle77] H. Kleinmann. Avoidance behavior in adult second language acquisition. *Language Learning*, 27:93–108, 1977.

[Lid80] Scott K. Liddell. *American Sign Language Syntax.* Mouton Publishers, 1980.

[LM88] C. Langston and M. Maxwell. Holistic judgment of texts by deaf and ESL students. *Sign Language Studies*, 60:295–312, 1988.

[LM91] Diane C. Lillo-Martin. *Universal Grammar and American Sign Language.* Kluwer Academic Publishers, Boston, 1991.

[McL87] R. McLaughlin. *Theories of Second-Language Acquisition.* Edward Arnold, London, 1987.

[Odl89] T. Odlin. *Language Transfer.* Cambridge University Press, NY, 1989.

[Pad81] C. Padden. Some arguments for syntactic patterning in American Sign Language. *Sign Language Studies*, 32:239–259, Fall 1981.

[PQ73] D. Power and S. Quigley. Deaf children's acquisition of the passive voice. *Journal of Speech and Hearing Research*, 16:5–11, 1973.

[QP84] S. P. Quigley and P. V. Paul. *Language and Deafness.* College-Hill Press, Inc., San Diego, 1984.

[QPS77] S. P. Quigley, D. J. Power, and M. W. Steinkamp. The language structure of deaf children. *The Volta Review*, 79(80):72–84, February-March 1977.

[QSW74] S. P. Quigley, N. L. Smith, and R. B. Wilbur. Comprehension of relativized sentences by deaf students. *Journal of Speech and Hearing Research*, 17:325–341, 1974.

[QWM76] S. Quigley, R. Wilbur, and D. Montanelli. Complement structures in the language of deaf students. *Journal of Speech and Hearing Research*, 19:448–457, 1976.

[RQP76] W. K. Russell, S. P. Quigley, and D.J. Power. *Linguistics and Deaf Children: Transformational Syntax and its application*. The Alexander Graham Bell Association for the Deaf, Inc, Washington, D.C., 1976.

[Rut83] W. Rutherford. Language typology and language learning. In S. Gass and L. Selinker, editors, *Language Transfer in Language Learning*, chapter 21. Newbury House, Rowley, MA, 1983.

[Rut89] William Rutherford. Interlanguage and pragmatic word order. In S. M. Gass and J. Schachter, editors, *Linguistic Perspectives on Second Language Acquisition*, chapter 7, pages 163–182. Cambridge University Press, NY, 1989.

[Sac90] Oliver W. Sacks. *Seeing Voices*. University of California Press, Berkeley and Los Angeles, CA, 1990.

[Sar87] Olivia Saracho. Teaching second-language literacy with computers. In Douglas Hainline, editor, *New Developments in Computer-Assisted Language Learning*, chapter 5, pages 53–68. Nichols Publishing Company, NY, 1987.

[Sch82] J. Schumann. Simplification, transfer and relexification as aspects of pidginization and early second language acquisition. *Language Learning*, 33:337–66, 1982.

[Sch86] Ethel Schuster. The role of native grammars in correcting errors in second language learning. *Computer Intelligence*, 2:93–98, 1986.

[Sid79] Candace L. Sidner. *Towards a Computational Theory of Definite Anaphora Comprehension in English Discourse*. PhD thesis, MIT, June 1979.

[Sid83] Candace L. Sidner. Focusing in the comprehension of definite anaphora. In Robert C. Berwick and Michael Brady, editors, *Computational Models of Discourse*, chapter 5, pages 267–330. MIT Press, Cambridge, MA, 1983.

[SK84] David Sanders and Roger Kenner. Whither CAI? the need for communicative courseware. In David H. Wyatt, editor, *Computer-assisted language instruction*, pages 33–40. Pergamon Press, NY, 1984.

[Sle82] D. Sleeman. Inferring (mal) rules from pupil's protocols. *Proceedings of ECAI-82*, 9:160–164, 1982.

[SR79] J. Schachter and W. E. Rutherford. Discourse function and language transfer. *Working Papers on Bilingualism*, 19:1–12, 1979.

[Sto60] W. C. Stokoe, Jr. Sign Language structure. *Studies in Linguistics occasional papers*, (8), 1960.

[Str88] Michael Strong. *Language Learning and Deafness.* Cambridge University Press, NY, 1988.

[Sur91] Linda Z. Suri. Language transfer: A foundation for correcting the written English of ASL signers. Technical Report TR-91-19, Dept. of CIS, University of Delaware, 1991.

[Sur92a] Linda Z. Suri. Correcting illegal NP omissions using local focus. In *Proceedings of the 30th Annual Meeting of the Association for Computational Linguistics (Student Session)*, pages 273–275, June 1992.

[Sur92b] Linda Z. Suri. Using local focus to correct illegal NP omissions (a Ph.D. proposal). Technical Report TR-93-07, Dept. of CIS, University of Delaware, 1992.

[Swi89] M. Virginia Swisher. The language-learning situation of deaf students. *TESOL Quarterly*, 23(2):239–257, June 1989.

[Und84] John H. Underwood. *Linguistics, Computers, and the Language Teacher.* Newbury House Publishers, Inc., Rowley, MA, 1984.

[Wil77] R. B. Wilbur. An explanation of deaf children's difficulty with certain syntactic structures of English. *The Volta Review*, 79(80):85–92, February-March 1977.

[WP83] R. Wilbur and L. Petitto. Discourse structure in American Sign Language (or, how to know a conversation when you see one). *Discourse Processes*, 6:225–241, 1983.

[WS83] Ralph M. Weischedel and Norman K. Sondheimer. Meta-rules as a basis for processing ill-formed input. *American Journal of Computational Linguistics*, 9(3-4):161–176, 1983.

[WVJ78] Ralph M. Weischedel, Wilfried M. Voge, and Mark James. An artificial intelligence approach to language instruction. *Artificial Intelligence*, 10:225–240, 1978.