



University of Pennsylvania
ScholarlyCommons

Technical Reports (CIS)

Department of Computer & Information Science

December 1984

VP²: The Role of User Modeling in Correcting Errors in Second Language Learning

Ethel Schuster
University of Pennsylvania

Follow this and additional works at: https://repository.upenn.edu/cis_reports

Recommended Citation

Ethel Schuster, "VP²: The Role of User Modeling in Correcting Errors in Second Language Learning", .
December 1984.

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-84-66.

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/cis_reports/728
For more information, please contact repository@pobox.upenn.edu.

VP²: The Role of User Modeling in Correcting Errors in Second Language Learning

Abstract

This paper describes a system, *VP²*, that has been implemented to tutor non-native speakers in English. The system applies Artificial Intelligence techniques developed in Natural Language research. In particular, it differs from standard approaches by employing a model of its users to customize instruction based on knowledge of the student's native language. The system focuses on the acquisition of English verb-particle and verb-prepositional phrase constructions. It diagnoses errors that students make due to interference of their native language. *VP²* recognizes syntactic variation in English sentences, allowing freer translation. *VP²* is a modular system: its model of a user's native language can easily be replaced by a model of another language. Its correction strategy is based upon comparison of the native language model with a model of English. The problems and solutions presented in this paper are related to the more general question of how modeling previous knowledge facilitates instruction in a new skill.

Comments

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-84-66.

**VP²: THE ROLE OF USER
MODELLING IN CORRECTING ERRORS
IN SECOND LANGUAGE LEARNING**

**Ethel Schuster
MS-CIS-84-66**

**Department Of Computer and Information Science
Moore School
University of Pennsylvania
Philadelphia, PA 19104**

December 1984

Acknowledgements: This research was supported in part by DARPA grants N00014-85-K-0018 and N00014-85-K-0807, NSF grants DCR-86-07156, DCR8501482, MCS8219196-CER, MCS-82-07294, 1 RO1-HL-29985-01, U.S. Army grants DAA6-29-84-K-0061, DAAB07-84-K-F077, U.S. Air Force grant 82-NM-299, AI Center grants NSF-MCS-83-05221, U.S. Army Research office grant ARO-DAA29-84-9-0027, Lord Corporation, RCA and Digital Equipment Corporation.

UNIVERSITY OF PENNSYLVANIA
THE MOORE SCHOOL OF ELECTRICAL ENGINEERING
SCHOOL OF ENGINEERING AND APPLIED SCIENCE

**VP²: THE ROLE OF USER MODELLING IN
CORRECTING ERRORS IN SECOND LANGUAGE LEARNING**

Ethel Schuster

Philadelphia, Pennsylvania

December 1984

A thesis presented to the Faculty of Engineering and Applied Science
of the University of Pennsylvania in partial fulfillment of the
requirements for the degree of Master of Science in Engineering for
graduate work in Computer and Information Science.

Timothy W. Finin - Supervisor

Bonnie L. Webber - Supervisor

Graduate Group Chairperson

ABSTRACT

This paper describes a system, *VP²*, that has been implemented to tutor non-native speakers in English. The system applies Artificial Intelligence techniques developed in Natural Language research. In particular, it differs from standard approaches by employing a model of its users to customize instruction based on knowledge of the student's native language. The system focuses on the acquisition of English verb-particle and verb-prepositional phrase constructions. It diagnoses errors that students make due to interference of their native language. *VP²* recognizes syntactic variation in English sentences, allowing freer translation. *VP²* is a modular system: its model of a user's native language can easily be replaced by a model of another language. Its correction strategy is based upon comparison of the native language model with a model of English. The problems and solutions presented in this paper are related to the more general question of how modelling previous knowledge facilitates instruction in a new skill.

A Tole y a Rina

ACKNOWLEDGEMENTS

I would like to thank my advisors Tim Finin and Bonnie Webber for their help and support throughout this work as well as for their helpful comments in earlier drafts.

I have no words to thank Julia Hirschberg, Martha Pollack, Haim Levkowitz and Kathleen McCoy for putting up with all my questions, for providing me with insightful comments and for their help in the preparation of this manuscript.

I would also like to thank the people in the Natural Language Seminar for their comments and the people in TESOL at the University of Pennsylvania for providing me with some of the literature. To them and others, many thanks.

ACKNOWLEDGEMENTS

I would like to thank my advisors Tim Finin and Bonnie Webber for their help and support throughout this work as well as for their helpful comments in earlier drafts.

I have no words to thank Julia Hirschberg, Martha Pollack, Haim Levkowitz and Kathleen McCoy for putting up with all my questions, for providing me with insightful comments and for their help in the preparation of this manuscript.

I would also like to thank the people in the Natural Language Seminar for their comments and the people in TESOL at the University of Pennsylvania for providing me with some of the literature. To them and others, many thanks.

This work was supported by a grant from the National Science Foundation (IST81-12439).

Table of Contents

1. Introduction	1
2. Computer-Assisted Language Instruction	6
2.1. Review of the Literature	6
2.2. A Tutor for English Speaking Students Learning German	6
2.2.1. Error Handling	7
2.3. Generative Systems	8
2.4. Systems that Handle Ungrammatical Input	10
2.4.1. Pulman's System	10
3. User Models	12
4. The Process of Acquisition	18
4.1. Linguistic Theories of Second Language Acquisition	18
4.1.1. Contrastive Analysis	18
4.1.2. Creative Constructive Theory	21
4.1.3. Recent Theories of Second Language Acquisition	23
4.2. Evidence of Using Analogy to Prior Knowledge	24
4.3. Our Approach to Second Language Learning	26
5. Overview of <i>VP</i>²	28
5.1. Identifying the Problem	28
5.2. Scope of the Phenomenon	30
5.3. Goals of the Approach	32
5.4. User Model in <i>VP</i> ²	35
5.4.1. Knowledge Facilitates Performance	35
5.4.2. Knowledge Hinders the Performance	36
5.5. Form of Instruction	38
5.6. User of <i>VP</i> ²	42
5.7. Implementation	43
5.7.1. Components of <i>VP</i> ²	45
5.8. Procedure Used to Handle Responses	47
5.9. Error Handling	50
5.9.1. Missing Particle	50
5.9.2. Additional Preposition	51
5.9.3. Incorrect Preposition	52
5.9.4. Spelling Errors or Unrelated Answers	52

6. Conclusions	53
6.1. Further Research	54
APPENDIX A. Grammars	57
APPENDIX B. Sample Session	63

CHAPTER I

Introduction

Many sophisticated, intelligent tutoring systems have been developed for Computer Assisted Instruction (CAI) [Weischedel et al. 78], [Bates et al. 81], [Bates and Wilson 80], [Collins et al. 75], [Brown et al. 75], [Woods et al. 72], [Carbonell 70], that use Artificial Intelligence (AI) techniques. These systems have been developed with the hope that AI techniques will play an important role in education. The fundamental motivation for building intelligent computer tutoring is the immense success that human tutors have on an individual basis as compared to group or classroom instruction. In general, it seems that private human tutors are more effective in their teaching than instructors in the classroom. [Anderson et al. 84] The expectation then, when building intelligent tutors, is to find ways of emulating individual human tutors with computers.

Expert systems appear to be one of the most promising approaches to developing these tutoring systems. The key idea is to treat the human tutor as an expert whose knowledge can be extracted and build systems that include this knowledge. For example, [Stevens et al. 79] have developed a tutor, modelled as an expert system, that teaches about physical phenomena. The system attempts to correct misconceptions that the students may have by probing the student's

knowledge and comparing the student's responses with a pre-compiled list of possible misconceptions. In general, systems such as this include in their design such a list of misconceptions. The system works by mapping the students' errors onto this list and responding with some pre-determined text associated with each listed misconception. Nevertheless, human tutors can not always have a complete list of possible errors that the student may make. Human tutors use different methods in their tutoring style and usually rely heavily on the use of natural language, their knowledge of the domain, and common methods of reasoning to deduce and correct students' errors.

Other approaches to building tutoring systems have concentrated on identifying principles of effective tutoring techniques displayed by human tutors and develop systems that include these methods in their design. Some of these computer tutors actually involve expert systems as subparts in their design. For example, [Brown et al. 82] have implemented a tutor for troubleshooting circuits. This tutor has an expert circuit-analysis system that can reason about the domain and provide answers to the student. In another system, [Sleeman 82] uses a rule-based approach to infer the errors that the students have when solving algebraic problems. He discusses different approaches to determining the incorrect rules that the students used. These approaches include (1) methods for inferring "mal-rules" based on the student's answer, and (2) the definition of a number of incorrect rules. The system attempts to explain the student's answer by applying one of those incorrect rules.

Related to this work on recognizing students' incorrect rules is work on cooperative man-machine interaction. [Joshi 82] has suggested that for effective man-machine interaction, it is necessary for the system to be able to recognize and correct possible user misconceptions. That is, not only must a cooperative computer system give a truthful and informative response, they must also be able to "square away" the beliefs of the user if there exists some discrepancy between what the system believes and what the system believes the user believes. [McCoy 83] has suggested that when a user reveals a misconception in his/her interaction with an expert system, it is the system's job to characterize and reason about that misconception in order to figure out its sources and respond adequately to the user. Even though this work is not concerned with tutoring but rather focuses on effectively responding to the users of an expert system by recognizing their misconceptions, many of the same principles apply.

There have also been AI-based tutorial systems developed for Language Instruction (LI), e.g. [Bates and Wilson 80], [Weischedel et al. 78], allowing students to communicate with the tutor via typewritten natural language. [Weischedel et al. 78] describes an intelligent tutor for foreign language instruction which can point out mistakes and hypothesize their cause, again based on a list of incorrect forms that is included in its grammar.

The system presented in this work, *VP²*, differs from previous tutoring systems in that it includes a user model and does not include a library of incorrect forms. Rather, the system is provided with correct forms in both languages, from which

it deduces the reasons for errors and tailors its response accordingly. That is, it approaches the problem of the influence of previous knowledge on learning material that may be similar to already known material.

The development of *VP²* makes an interesting claim about language understanding and knowledge representation—namely that speakers' grammars can serve as user models. *VP²* must address the following question: How can correspondences between the grammars of two languages provide an account of grammatical errors made by native speakers of one language attempting to learn the other.

VP² focusses on the acquisition by non-English speakers of English verbal constructions formed from a verb plus particle or verb plus prepositional phrase. A well-known claim is that people often rely heavily on their previous knowledge when learning a new skill [Winston 80], [Rumelhart and Norman 81]. This previous knowledge can sometimes hinder their learning [Halasz and Moran 82]. In other words, people reason by analogy from a previous skill and these analogies are sometimes incorrect. Many errors caused by such reasoning can be predicted if prior knowledge is taken into account.

Cooperative CAI systems should include a model of relevant aspects of users' prior knowledge to predict and prevent errors or detect and correct them more easily. The ultimate goal of *VP²* is not only to identify particular mistakes and point out their possible causes to the student but also to explain the differences and similarities in the verbal constructs of the two languages in focus.

VP² allows some flexibility in the English translations it will accept as correct. That is, the order of the words in the sentence that the student types does not have to be exactly the same as the order of the given Spanish sentence. For example, when the student is asked to translate the sentence:

Pense en ti cuando vi ese libro.

VP² is able to accept as correct sentences:

I thought of you when I saw that book.

as well as

When I saw that book, I thought of you.

Furthermore, *VP²* is a modular system: its model of a user's native language can easily be replaced by a model of another language.¹ Its correction strategy is based upon comparison of the native language model with a model of English.

A review of Computer Assisted Language Instruction (CALI) literature is described in Chapter 2. Chapter 3 provides a description of user models in computer systems and the advantage they provide for a system like *VP²*. Chapter 4 presents some aspects of the linguistic theories of Second Language Acquisition (SLA). Chapter 5 describes a version of *VP²* that has been implemented for native speakers of Spanish. It provides an overview of the system and discusses the way *VP²* recognizes and corrects students' errors. Finally, Chapter 6 presents some conclusions and further issues raised by this study.

¹Note that one would not replace the target language L_2 , since the use of certain verbs with prepositions/particles is a feature of a particular L_2 , namely English. The system would have to be changed somewhat to tutor a different grammatical phenomenon.

CHAPTER II

Computer-Assisted Language Instruction

2.1. Review of the Literature

Much work has been done on traditional CAI systems for language instruction. Some people who learn a second language as adults have great difficulty with the grammatical structure of the language they are trying to learn [Pica 83] [Taylor 75]. Computer-assisted language instruction (CALI) systems have been developed for tutoring language structure, but have been quite limited. These systems consist in carefully devised sequences of questions in which the answers are directly matched against a library of correct answers. They do not give the students much freedom in their use of the language, and are specifically geared toward drill and practice exercises. This chapter describes some of the different approaches that have been used.

2.2. A Tutor for English Speaking Students Learning German

[Weischedel et al. 78] have developed a tutorial system designed for English-speaking college students taking their first course in German.

The students are presented with a text in German and a set of questions which they are expected to answer. The tutor accepts both good (grammatical) and a

number of bad (ungrammatical) forms of German and it can recognize why a bad form is ungrammatical. The goal of the system is to diagnose and pinpoint the student's syntactic and semantic errors. The student is free to use the language naturally without being constrained by a specific grammatical form--that is, s/he is not restricted by a specific word order.

2.2.1. Error Handling

When learning German, many native English-speaking students make errors due to interference from their first language. For example, they use English word order in the past participles of German perfect tenses. In German, the past participles must come at the end of a clause. In order to handle such anticipable incorrect syntactic forms, the tutor has these incorrect forms included in its grammar.

German is a highly inflected language, with complex constraints on subject-verb agreement, inflected noun phrase endings, and word order of adverbial elements. These are another major source of syntactic errors made by the students because they are much more complex than their English counterparts. In order to accept incorrect subject-verb agreement and still diagnose the student's mistake, the system uses predicates that check whether the form is correct without rejecting it. If a predicate evaluates to false, the system generates an error message. To diagnose errors in noun phrase endings (e.g. in article and adjective endings), the system checks to see if the endings would be correct, assuming that the student

forgot the gender of the noun and tells the student if that is the case. The tutor also checks whether the endings would be correct if the student had merely forgotten what syntactic case should have been used. In this last case, it can inform the student about the correct case for the noun phrase.

The tutor can handle spelling errors too. To unanticipated errors it responds that it cannot understand and proceeds to the next question.

The German tutor can also handle certain kinds of semantic errors. These include errors with respect to reference, irrelevant answers, and errors due to presuppositions. These errors are handled using general principles. For example, an interpretation in which all references can be made is always preferred over one where not all references can be completed.

This German tutor demonstrates that AI approaches to computer-assisted language instruction offer benefits that can complement the traditional CALI systems. These AI approaches consist in allowing the student to more freely use the language, pinpoint grammatical errors, and find semantic and comprehension errors as well.

2.3. Generative Systems

[Bates et al. 81] and [Bates and Wilson 80] describe a system called ILIAD which tutors in both the production and comprehension of written English. The system generates exercises for the student using a dictionary, a grammar of

English and a set of sentence characteristics, many of which are chosen by the student. The exercises that the system generates focus on a wide variety of syntactic structures. This generative approach appears to be a flexible and powerful one in the design of tutorial systems. It can generate many examples without having them "canned", it has a vocabulary that can be increased easily, and it generates its information from rules that are part of its grammar.

The system can be used by people learning English as a Second Language or people deaf from birth. ILIAD differs from other CALI systems in that:

- The examples and exercises that are posed to the student are generated by the computer from a dictionary and a grammar of English as opposed to being pre-stored.
- The vocabulary level, type, content and complexity of the exercises are controlled by the student. That is, the student can choose the level of the exercises, tailoring them to his/her individual needs.

Although the ILIAD system is geared to receiving the correct answer, it can handle simple student errors. The student can request help or hints to lead him to the correct answer (as well as requesting the answer itself). Capitalization and punctuation errors are ignored by the system. If the student types an incorrect answer and the system recognizes it, the system allows the student additional chances. If after these attempts, the student fails to answer correctly, the system

provides the answer. The system does not attempt to tell the student the possible reason for the error or how to correct it him/herself, except by providing the correct answer before proceeding to the next exercise.

2.4. Systems that Handle Ungrammatical Input

[Pulman 84] describes a natural language system that can handle ungrammatical input. His system includes some actual and potential applications to Second Language Learning, by providing practice and instruction in "tricky" parts of English grammar.

2.4.1. Pulman's System

Pulman's system uses Generalized Phrase Structure Grammar (GPSG) [Gazdar 82], with a parser extended to detect errors. The system attempts to parse a sentence, and when it fails, it keeps a "picture" (or record) of what it has been able to parse, along with a parse tree. From there, it proceeds to generate an error message. The error messages include: (1) saying what constituent(s) the system is trying to complete, (2) saying what kind of item it was expecting at the time of failure, (3) offering information about what would be necessary to complete the sentence correctly, by providing examples.

The advantage of this mechanism is that it is simple and defined for the whole class of context-free grammars. The system provides help by explicitly locating the position of the incorrect constituent in the sentence.

All these systems provide mechanisms for dealing with language teaching. However, these systems have not addressed the problem of the effect of knowledge of one language on the learning of a second language. They thus fail to provide the student with some possible explanation for his/her error. These issues are the focus of this work.

CHAPTER III

User Models

Much research in AI has been devoted to the development of consultation programs, expert systems and tutorial systems. Part of this research has involved enabling such systems to answer questions, recognize misconceptions and provide explanations.

Some systems have user models in order to better handle their interaction with their users. These include ones that: 1) provide diagnosis, using the user model to tailor explanations to level of user expertise [Wallis and Shortliffe 81], and 2) answer questions about the domain, generating their responses based on the user's interests [Rich 79].

Good answers and explanations serve several functions in expert and tutorial systems as follows:

- If errors occur, the system can provide a method to examine the program's reasoning.
- They assure the user that the program's reasoning is logical.
- They can persuade the user that "unexpected" advice may be helpful.
- They can teach the user in areas in which s/he may not be very knowledgeable.

All these functions impose requirements on the system. That is, the system should be able to show its user its reasoning. It should also allow the user to examine its reasoning and underlying knowledge at different levels of detail. In general, the system should be able to respond to the user, regardless of his/her level of expertise. Therefore, the system "must have the capability to tailor its explanations to the varying needs and characteristics of its users."²

For a system to customize its answers and explanations to its users, it must be able to model the users' knowledge and their motivations for using the system. This **user model** can represent the minimal information that the user knows. One approach is to rate the user's level of expertise. This approach can then be extended to distinguish subareas of a user's expertise in different portions of the system's knowledge. The system should also be able to respond to questions that the user may ask that the system may have thought the user already knew. In other words, the system must be able to update dynamically its user model as its interaction with the user proceeds. Some systems build their user models with the aid of stereotypes and use those models to guide them in their task [Rich 79]. These systems have small sets of models of typical users at various levels of expertise, which are invoked either by requesting the user to characterize him/herself or by deducing, from a series of questions posed to the user, the closest model that appears to fit him or her.

² [Wallis and Shortliffe 81]p.1

Much work has been done on developing user models in computer systems. [Genesereth 79] has developed a system for "Automatic Consultation" called the Advisor which provides consultation for MACSYMA³ users. The implementation of the Advisor is based on the observation that "a good consultant must possess not only a substantial knowledge of its material but also a good model of its user's knowledge."⁴ In order to obtain the model of its current user, the Advisor reconstructs the user's plan by converting the user's actions into a dataflow graph and then interrogating the user about his/her goal. The system tries to recognize the plan and identify possible misconceptions by using a "parsing" procedure that searches the partially reconstructed plan for plan fragments (from its "Plan Library") or error fragments (from its "Error Library"). The first such "parsing" of the graph that the Advisor finds is assumed to be the correct plan. If later on the Advisor realizes that this plan does not correspond to the user's, it eliminates this plan and searches for another one.

Once the Advisor has a version of the user's plan, it tries to identify possible misconceptions. This is done by acquiring a suspicion of what may be wrong. Suspicion can be aroused by either recognition of some standard error in the user's plan or by a general model debugging process. Once a suspicion is aroused, the Advisor confirms that this is a misconception by asking the user whether or

³MACSYMA is a large, interactive computer system designed to assist mathematicians, scientists, and engineers in performing symbolic manipulation of mathematical expressions.

⁴ [Genesereth 79],p. 319

not s/he believes it. Once the misconception has been identified, the Advisor corrects it and provides the user with advice. In general, the Advisor tries to help the user achieve his/her goal by providing useful information which include alternative approaches to reaching the goal.

Other works have focussed on student models in tutorial systems. In general, an intelligent tutor should have the ability to adapt its discourse to each individual student. It should be able to adequately present information both to a student who knows a lot about the domain and to a student who knows very little about the domain.

Some tutorial systems that contain user models are [London and Clancey 82], [Sleeman 82], and [Farrell et al. 84]. [London and Clancey 82] describe a student modeler for the GUIDON2 tutor. GUIDON2 is a tutoring program that uses the case method approach to teach medical diagnosis. The system not only accounts for the student's knowledge but also for his/her planning. GUIDON2 has three components: an "expert", a student model, and an instructional manager. The expert represents the knowledge about the medical domain and diagnosis strategies. The student model interprets the student's answers by using the expert's knowledge, evaluating the student's knowledge and producing alternatives. The instructional module applies discourse and teaching strategies to decide whether to interrogate the student or provide him/her with advise. The student model uses two separate approaches to infer and evaluate the student's plans. First, it forms a model-driven *range of predictions* which represent the

plans that the student should be following. If the student's answers match the predictions, the predictions are used to describe the student's answers. If the answers are incompatible with the predictions, then more specific processing is required to explain the data. The student model generates its predictions by simulating the expert. It also makes some adjustments to increase the computational efficiency, the level of detail, and the likelihood of successful recognition and evaluation of student's answers. The student model also has to be able to perform reasonably well even when the student acts in unusual ways and should be able to recover from its own errors. The student model in GUIDON2 does not base its knowledge of what the student actually "knows" but on the system's predictions.

[Farrell et al. 84] describe an intelligent computer tutor for LISP that incorporates methods used by good private tutors. The students learn LISP with the tutor and work through a series of problems. The tutor consists of a problem solver which generates steps towards the solution and an advisor that compares the student's steps (student model) with those of the problem solver. This problem solver continually monitors the students' progress and tries to assess the *knowledge* that the student has in order to produce certain behavior. This knowledge is represented as production rules and goals. The tutor also has a set of bad rules and goals that it can recognize. Hence, the student model is obtained interactively by inferring those rules and goals from the tutor's library that produce the student's behavior. These production rules also serve as a novice

model that follow the student as s/he solves the problem. The LISP tutor presents a system that can infer and update the student model in order to correct errors.

The systems described above include user models in their design. However, very few of these systems describe the knowledge in their user model to be "prior knowledge" (because this prior knowledge is not always available). It is more common for these systems to infer the knowledge based on the interaction, disregarding the prior knowledge that may affect the learning. In the system presented here, the user model consists of (1) a body of correct knowledge of a language -- that is, the student's native language -- and (2) a mapping process that the student is believed to use when lacking knowledge of the target language. We feel that including the information about the student's native language in the user model is a better approach to the kind of tutoring we want to do. This way, the system has a closer approximation to what the student knows .

CHAPTER IV

The Process of Acquisition

4.1. Linguistic Theories of Second Language Acquisition

The relationship between first and second language acquisition has interested linguists and language teachers for many year. Of particular interest is how the relationship between the native language (L_1) and the second language (L_2) affects second language acquisition. Two theories have been proposed to explain how second languages are acquired. One is the Contrastive Analysis Theory and the other is the Creative Constructive Theory. These theories have been considered to be mutually exclusive, but neither by itself provide a detailed and complete theory for second language acquisition.

4.1.1. Contrastive Analysis

The contrastive analysis (CA) theory claims that "where features of the source and target languages match, learning will be facilitated; where features of the two languages do not match, language learning will be difficult."⁵ When the features of L_1 don't match those of L_2 , difficulties appear as "interference". Such difficulties provide evidence that the learner uses features of L_1 in acquiring L_2 .

⁵ [Lado 57]p. 2

Second language learning is usually seen as being acquired by analogy with [Dommergues and Lane 76], association to, reinforcement by [James 81], and transfer of the native language, that is, simple reliance on the structure of the native language when not enough of the second language is known.

The fundamental claim of the contrastive analysis theory is that when people are learning a second language, the patterns of the language to be learned are matched with those of the native language. Those that indeed match require 'no learning' while those that do not match produce errors and result in interference. When there are features of the second language that do not match those of the native language, the L_2 learner has to **learn** the features of L_2 that are not present in L_1 .

The CA theory of SLA is supported by various kinds of data that show interference of L_1 in L_2 where those features of L_2 do not match the ones of L_1 . A study by [Taylor 75] has shown evidence of syntactic interference in the acquisition of English as a second language for adult native speakers of Spanish. The study was designed to analyze the errors in their use of English Auxiliaries and Verb Phrases. The results show that the largest percentage of errors could be accounted for by interference of Spanish on English. For example, the translation of

(1) The men are there at 8 o'clock

from

'Los hombres estan alla a las 8 en punto'

instead of

The men will be there at 8 o'clock

demonstrates how learner's errors can be attributed to direct translation of the Spanish verb form. The results of this study indicate that learners of a second language use analogy when learning a new language. Because of their lack of familiarity with the new linguistic system, they rely extensively on their native language for support. This study also found that with increased proficiency in the new language, the learners rely less frequently on their native language grammar and more on their knowledge of the new language.

[Dommergues and Lane 76] conducted a grammaticality judgment study to measure errors in learning the syntax of a new language (English) for native speakers of French. They found that sentences such as (2) and (3)

(2) **My father is teacher in London.**

(3) **Mommy made two fruits cakes.**

are accepted as grammatical by French students learning English.⁶ These sentences are both directly translated from the French, hence the omission of the article in the first one and the use of a plural inflection on the adjective in the second one.

The results of this study confirm that the partial congruence of syntactic patterning in a speaker's second language and his/her first language is a significant source of his/her errors.

⁶ [Dommergues and Lane 76]p. 111

On the other hand, there have been studies that prove that the CA theory is unable to adequately account for some learners' errors in L₂. CA predicts that similar features in the native language and the second language should make such features easy to learn. Nevertheless, this is not always the case. For example, [Sciarone 70] describes how Dutch speaking children make mistakes when learning French auxiliaries. The French past tense auxiliaries, 'avoir' and 'etre' have corresponding equivalents in Dutch: 'hebben' and 'zijn'. He has found that the children tend to confuse the two French verbs in any tense and generate sentences such as:

- (5-1) Je suis un livre
 (6-1) J'ai malade

instead of

- (5-2) J'ai un livre
 (6-2) Je suis malade

4.1.2. Creative Constructive Theory

The Creative Constructive Theory (CCT) views second language learning as a creative constructive process -- [Dulay and Burt 74] on child SLA and [Bailey et al. 74], [Richards 71] on adult SLA -- which is rooted in innate and universal structural properties of the mind. These properties are similar in determining the acquisition of both L₁ and L₂. CCT claims that there is 'almost' no transfer of the L₁ features onto L₂: instead, the general creative process of language learning is transferred. This theory is also guided by a generative transformational model of linguistics which views the notions of innateness and universality as part of its framework for language learning.

The studies presented by [Richards 71] provide strong evidence that: 1) while some second language learner errors appear to exhibit native language transfers, many do not, and 2) many second language learners errors are both systematic and similar for learners of diverse backgrounds. [Bailey et al. 74] claim that the errors made by children in L_2 are "developmental" rather than interference. That is, they are similar in kind to the errors made by children learning their native language.

As in the CA theory, there are studies that support the approach presented by CCT. These include work by [Cook 73], [Cooper et al. 79], and [Gass 80]. [Cooper et al. 79] for example, looked at the acquisition of complex structures in English for Arabic and Hebrew speakers. They found that both groups performed similarly to one another and to two French Canadian groups even though the structure of L_1 was very different. Their results indicated that adult learners, despite their native language, interpreted English sentences in the same way that children did when learning English as their first language. All the same, these studies do not seem to provide enough evidence to support the CCT. As shown before, the CA studies have proven that **some** interference does occur. The CCT studies seem to suggest that by some general processing mechanism, the knowledge of a native language is used to actually organize various aspects of the second language. In general, the CCT does not provide a formal model to explain the process that goes on in learning a second language.

4.1.3. Recent Theories of Second Language Acquisition

Recent studies have looked at the theory of markedness and its application to second language acquisition. The theory of markedness is a theory of the phonological structure of human language, in particular, the structure of the lexicon. The notion of "markedness" was developed to characterize certain features in a language: ones in which the normal value of the feature can be called unmarked (U) and the less normal, called marked (M). For example, vowels may be nasalized but they are generally non-nasal, and although there are languages that have only oral vowels, there is no language with only nasal vowels. Hence, nasality is considered to be marked for vowels. Also, there is evidence that the phonological elements that are acquired later and lost earlier are the marked ones [Guitart 76].

The theory of markedness also assumes that the lexicon of a language is organized along certain universal principles that take into account the asymmetrical nature of features. [Rutherford 82] provides a review of the literature in which he notes that the theory of markedness has not yet been able to account for a theory of second language acquisition.

Other studies have focussed on the application of other theories of language universals to second language acquisition. These include the work of [Gass 80] in which a "multifactor approach" to second language acquisition is presented. That is, several factors play an important role in determining the learning patterns of a

second language learner. These include: (1) universal factors, (2) specific facts about the learner's native language, and (3) specific facts about the target language. Nevertheless, these studies have not been generalized to different language phenomena besides the use of restrictive relative clauses.

[Flynn 83] has proposed a basis for a theory of second language acquisition which explains both contrastive and constructive aspects of L_2 . She argues that principles of Universal Grammar (UG)⁷ that have been used to determine the acquisition of a first language also determine the acquisition of a second language. In addition, those principles of UG that involve parameters whose values are set by their use in the native language will be learned differently depending on whether they match in L_1 and L_2 .

4.2. Evidence of Using Analogy to Prior Knowledge

There have been some studies on the beneficial effects of analogy to prior knowledge when learning new information.

For example, [Schustack and Anderson 79] conducted two experiments to explore how memory for new information is affected by awareness of parallels to pre-experimental knowledge. In one experiment, the subjects had to study brief biographies of fictional characters analogous to famous people. Next the subjects

⁷The theory of Universal Grammar [Chomsky 80] specifies that there exist linguistic principles which underlie all natural languages. These principles create the initial state of a language learner's mind (the basis on which knowledge of a language develops) and constrain the language acquisition process.

were given a group of sentences for each fictional character, where each fictional character's name was paired with sentences from both his/her biography, and with sentences from other biographies (as distractors). The subjects were asked to judge whether a sentence had occurred in that character's biography. In the other experiment, the subjects were given the same biographies to study but the names of the characters were paired with names of famous people, e.g. the name corresponding to Golda Meir's biography was Hannah Enkol. With these experiments, [Schustack and Anderson 79] wanted to address the issues of *how* and *when* an analogy has its effects. They found that a famous person's name was only beneficial when facts in the biography were true for that person. This suggested that prior knowledge improves remembering and that a cue can induce the use of analogies to stored information.

[Douglas and Moran 83] show how text editor novices rely on knowledge of typewriting in order to understand the "semantics of text editor operators"⁸. They show that this knowledge partially accounts for the learners' performance errors. In their experiment the users were computer-naive people who were familiar with typewriting. Their analogy to typewriting was evoked primarily because of the similarities of the keyboards, the similarity of the screen to a typed page, and the similarity of the task in editing and typing. The teachers also prompted the typewriter analogy by using remarks such as "It works just like a

⁸ [Douglas and Moran 83]p. 102

regular typewriter."⁹ The author has found similar effects on users moving from one text editor to another. [Schuster 83]

4.3. Our Approach to Second Language Learning

The work presented here has been developed under the lines of the CA theory. Since interference errors reflect the grammar of the student's native language but are errors only because the same rules are not operative in the second language, they are often predictable from a contrastive analysis of the two grammars and, possibly, a knowledge of the student's overall proficiency. This work focusses on one problem that appears on the syntactic level of L₂ acquisition, that is, the use of the complex construction, verb plus preposition/particle, in the English of non-native speakers.¹⁰

Prepositions in English are a frequent source of errors in the speech of non-native speakers. [Scott and Tucker 74] analysed the use of prepositions in the English of Arab students. Prepositions seldom have a one to one correspondence between English and Arabic. An Arabic preposition may have several alternative English translations, while an English preposition may have several alternative Arabic translations. This work shows that approximately two thirds of the errors can be accounted for interference of the native language. It should be noted that a larger proportion of errors attributable to native language interference occurred

⁹ [Douglas and Moran 83],p. 100

¹⁰Note that we are dealing only with second language acquisition, not third or fourth, where the learner generalizes only from his/her native language.

at the end of the training session as opposed to the beginning. This suggests that the students were not able to overcome what seems to be a persistent problem of native language interference. With the belief that errors can be attributed to native language interference, the work described here was developed to teach the use of verbs and preposition/particle in English to native Spanish speakers.

Our system, *VP²*, helps overcome what [Cassidy 83] has identified as one of the major sources of instructional error: "the failure or inability to take into account appropriate learner circumstances",¹¹ by which she means "the learning experience must be on target in terms of an individual's profile of prior knowledge or skill"¹². This prior knowledge we take to be their native language.

¹¹ [Cassidy 83]p. 15

¹² [Cassidy 83]p. 15

CHAPTER V

Overview of VP₂

5.1. Identifying the Problem

Familiarity with a wide range of idiomatic expressions, and the ability to use them appropriately in context, are among the distinguishing marks of a native speaker of English. Expressions such as **go over** (to review, to be received), **look on** (to watch), and **get ahead** (to make progress, to succeed) are part of the common coin of everyday conversational exchanges, and the tendency, especially in everyday use, to prefer these combinations over their single equivalent --review, watch, succeed-- helps to explain the widely-held view that idioms such as these are among the most characteristically "English" elements in the general vocabulary [Cowie and Mackin 82].

To say that such expressions are used widely does not necessarily imply that their meanings are always self-evident: A Spanish speaker would most likely understand **watch** better than **look on**, while a native English speaker may have some difficulty in explaining the sense of the combination in terms of its constituent parts. In fact, a close study of various kinds of idiomatic items brings to light many curious anomalies of form and meaning. While we can easily talk of

looking on or **breaking in** and replace the verbs by their equivalents, it may be awkward to speak of **the president making up his decisions** instead of **making up his mind**, where the same verbs appear.

As we have seen, the CA theory of second language acquisition predicts that non-native speakers of a language will tend to use forms that exactly match those of his/her native language. For example a native speaker of Spanish may talk of

- (1-1) **thinking *in* buying a house**
- (1-2) **pensando *en* comprar una casa**

instead of

- (2-1) **thinking *of* buying a house**
- (2-2) **pensando *de* comprar una casa**

or

- (3-1) **thinking *about* buying a house**
- (3-2) **pensando *acerca de* comprar una casa**

In Spanish, it is more common to say sentences such as 1-2, while in English it is sentences like 2-1 and 3-1 that are grammatical. The 'direct translation' of the Spanish preposition *en* is *in*, while the 'corresponding' English prepositions are *of* and *about*. Hence we have the Spanish sentence and its corresponding English translations:

- (4-1) **pensando *en* comprar una casa**
- (4-2) **thinking *of* buying a house**
- (4-3) **thinking *about* buying a house**

The native speaker of English picks up all of these forms and they become part of his/her speech. The foreign student, however, many times resorts to his/her knowledge of direct translations from his/her native language. This problem has

been addressed in *VP²*: we want to show the ways in which a tutoring system can assist a non-native speaker of English in his/her learning of the usage of verbs and prepositions and/or verbs with particles.

5.2. Scope of the Phenomenon

Traditionally, constituents of sentences in a given language have been divided into **open** and **closed** class items. Open class items include noun phrases, verb phrases, adjective phrases, adverbial phrases, as well as nouns, adjectives, adverbs, and main verbs. Closed class items include prepositions, particles, conjunctions, determiners, quantifiers, complementizers, possessives, pronouns, and auxiliary verbs. Open class items are new items that come easily into the language, they usually denote "objects" or "concepts", while closed class items are usually items that do not denote specific objects or concepts. They are grammatical markers and are usually unstressed.¹³ It seems that it is easier to learn new words (which we do all the time in L_1) and set just set them up into syntactic configurations of L_1 .

Prepositions and particles are considered closed class items and they are shown to be particularly problematic [Kean 79]. For example: it is common for native Spanish speakers to talk about

...*dreaming with leaving the USA*

instead of

¹³Note that in principle, what is open class in one language may be closed class in another.

...*dreaming of leaving the USA*

or for native Yiddish speakers to say

...*went in school*

instead of

...*went to school*

The system described here focusses on teaching the use of these closed class items in verb phrases. They do not seem to be learned the same way as open class items are, as attested to by the frequency of incorrect usages observed among non-native speakers of English. *VP²* looks at these closed class items in the native language and their correlates, if any, in the second language in order to predict errors in their use.

For the purpose of this system we will characterize English verbs and preposition/particles as follows:¹⁴

Verb-Particle - Sometimes called two-word verbs, these consist of a lexical verb followed by a modifier. These verb-particle forms cannot be "created" by arbitrarily combining any verb with any preposition or particle. For example:

Vanessa *ran into* Moris in the street.
Raquel and Tania will *talk over* our proposal.
She [filled up] the container.

In a sentence like *She filled up the container*, *up* serves as a particle. It can appear after the verb or after the object as in *She filled the container up*. If the

¹⁴*The Key to English: Two-Word Verbs*, New York: Collier-Macmillan, 1977

object is a pronoun, it must precede the particle as in *She filled it up* – not *She filled up it*.

A verb-particle combination is a grammatical unit which fulfills normal verb functions in English sentences. Like other verbs, it may be transitive, intransitive or bi-transitive, depending on whether it is followed by a noun object or not. It is also a semantic unit that has a meaning which often differs from the sum of its parts.

Verb-Preposition - These are verb-prepositional phrase sequences. For example,

Vanessa *ran* into the street.

Raquel and Tania usually *talk* over the fence.

This construction allows syntactic movements -- that is, the prepositional phrase can be fronted as in the following:

Into the street ran Vanessa.

Over the fence Raquel and Tania talk.

5.3. Goals of the Approach

VP² is a system that has been implemented to tutor non-native speakers of English in the use of English verb-particle and verb-preposition combinations. By maintaining a model of the user's knowledge -- in this case, a model of the user's native language -- the system can tailor its responses to the user's needs and facilitate his/her success in learning the use of English verb-particle and/or verb-preposition combination.

Tutoring systems which do not retain user models might behave inadequately in teaching different subjects. Without such a model, the system may not realize the cause of errors and hence it may be unable to provide the most adequate explanation for their occurrence. Consider the following interaction between a tutor for English and its student:

Tutor: TRANSLATE THE FOLLOWING SENTENCE:
 Pensaba en estudiar Inglés.
 Student: I thought in studying English.
 Tutor: Incorrect. The correct answer is
 'I thought of studying English.' or
 'I thought about studying English'.
 In English, you can use the verb <to think>
 as in the following cases:
 - 'There is little opportunity to <think out>
 what the long-term solution may be'
 - 'With the beginning of the new day we
 had to <think> things <over> again'
 - 'She has to <think up> a catchy name for this system'

The system here provides the correct answer but it fails to recognize that the user has translated directly from Spanish. The system's assumption is that there is a correct answer and that answer is the one the user expects without any further issues being considered. Here, the tutor could have provided a 'more adequate explanation' by informing the user why the error occurred. The system could have added:

Note that <en> is often translated as <in> in English but here it is not.

In order to provide such an explanation, the tutor must be able to retain some

model, no matter how simple, of the user's knowledge, especially of the user's knowledge of related domains that may *interfere* in their learning. This is the point of departure of this work: to see what role users' familiarity with one language (Spanish) plays in their learning a new language (English). This study provides the basis for the main features that must be included in the development of tutorial systems in order to (1) detect errors when they occur, (2) adequately correct errors when they occur, and (3) tailor the response to the user.

For a student of any tutorial system, there are several possible results: The student may successfully learn what s/he wants or needs, s/he may learn only part of it, or s/he may not learn anything and walk away frustrated and disappointed. The latter we want to avoid. Our goal is to maximize the student's understanding and minimize the time spent acquiring it even if the student only learns part of the information. We also want to prevent the student from learning incorrectly and to encourage him/her to reason about the errors s/he has made.

By tailoring the information to the student, we hope to make the learning process more efficient and therefore more satisfactory. [Farrell et al. 84] have claimed that human tutors can give good tutorial assistance because they can infer a model of the student's knowledge. In this work, we hope to use the user model to actually achieve the most satisfying and "good" results.

5.4. User Model in VP²

Instead of having a stereotyped user model or individual user models for each user, VP² has a canonical user model, that is, a more general model describing the grammar of the student's native language. The user model consists of a Spanish Grammar. This information is assumed to be standard for all Spanish speakers. One may argue that there are several dialects of Spanish and that the grammar represented in the system does not *exactly* correspond to the one the user has. I have considered a grammar of standard Spanish to be the universal for this system because the issues addressed in this work do not focus on idiomatic expressions but on specific language structures: the use of verb and prepositions, and verb with particles. These constituents do not seem to vary as much from one idiolect to another.

A student's knowledge of Spanish may either *enhance* or *hinder* his/her current learning of English. The latter is of primary interest here: because of his/her knowledge of Spanish, the student tends to draw inferences that may cause incorrect English forms.

5.4.1. Knowledge Facilitates Performance

First let us consider the case in which the student's knowledge of Spanish (and basic English words) facilitates the student's performance:

TUTOR: TRANSLATE THE FOLLOWING SENTENCE:

Yo escribo con el lápiz.

STUDENT: I write with the pencil.

TUTOR: Correct! Notice that the preposition <con> corresponds

to the English <with> and it appears in the same position in the sentence as it does in Spanish.

In this example, the student provides the correct answer and the tutor responds with additional information within the context of what the student **knows**. *VP²* uses the information in its user model in order to provide more adequate and "customized" answers.

We cannot tell whether the student has translated from Spanish or not in this sentence. All the system can do is point to the fact that indeed if direct translation was used for this sentence, it worked. [Dommergues and Lane 76] claim that "the knowledge that the speaker brings with him facilitates rather than hampers his second-language performance."¹⁵

5.4.2. Knowledge Hinders the Performance

There are times when the student's reasoning influenced by his/her previous knowledge hinders the interaction and causes the student to answer incorrectly.

When the student's previous knowledge affects the learning of new information the tutorial system should be able to recognize the error and correct it. Consider the following:

TUTOR: TRANSLATE THE FOLLOWING SENTENCE
 Moris pensó en comprar un carro.
 STUDENT: Moris thought in buying a car.
 TUTOR: You used the incorrect preposition <in>.

¹⁵p. 121

In English you can use <think of> or <think about> in this sentence. Note that the direct translation of <think of> --<pensar de>-- does not exist in Spanish. In English you can also use <think up> (an excuse, invent); <think over> (review); <think out> (consider, examine).

A response as in the previous example is not only more helpful but it reassures the student that the tutor is responding to his/her particular error. This situation requires the system to maintain a model of the student's native language or at least of the way verbs and prepositions are paired in that language. Tutorial systems that do not have user models may leave the students to determine the relationship between the use of certain structures or patterns in the new domain (or language) by themselves. We have also seen [Schuster 83] that users of some help systems pose their questions in the framework of their previous knowledge leading us to believe that the learning experience could be enhanced by a system that explicitly describes similarities and differences.

Maintaining a user model in VP^2 can help the tutor anticipate the kind of errors the student may make. Certain prepositions go with certain verbs in Spanish while different prepositions go with the corresponding verbs in English. For example, <pensar en> in Spanish translates directly to the English <think in> but corresponds to <think of> or <think about>. Therefore, a Spanish sentence that has <pensar en> translates to English as <think of> or <think about>. The same goes for <soñar con> which directly translates to <dream with> but corresponds to the English <dream of> or <dream about>. There is

also the problem of one verb in Spanish corresponding to a verb with particle in English, as in <recoger> corresponds to <pick up>; <escoger> to <pick out> and so on.

SPANISH		TRANSLATED TO		ENGLISH
<i>pensar en</i>	-->	<i>think in</i>	-->	<i>think of</i> <i>think about</i>
<i>soñar con</i>	-->	<i>dream with</i>	-->	<i>dream of</i> <i>dream about</i>
<i>recoger</i>	-->	?	-->	<i>pick up</i>
<i>escoger</i>	-->	?	-->	<i>pick out</i>

The user model in *VP²* also encompasses the possibility that the student may overlook the differences and use a one-to-one correspondence mechanism. In general, if the student uses this kind of one-to-one correspondence, user modelling can still recognize the error and deal with it in a satisfactory manner.

5.5. Form of Instruction

VP² is designed to interact with students who have acquired most of the English vocabulary and syntax and who have a basic knowledge of grammatical terms. All instructional information is given in English.

VP² presents the student with translation exercises. These exercises are not designed to substitute for classroom instruction because they are limited in their scope. They have been designed to provide additional practice in using two-word verbs in English.

In this set of exercises, the student is presented with a sentence in Spanish and is asked to translate it into English. After the student translates the sentence, the tutor looks for errors in the use of two-word verbs, i.e. the use of a wrong verb-preposition combination or incorrect particle. If the tutor finds such an error, it examines its model of the student's native language to locate the source of the error. Given this information it can then correct the error and explain to the student its probable cause. It may also make general comparisons between the verb forms used in Spanish and English. If an error is not seen as coming from the native language, the tutor provides the correct answer and proceeds to the next exercise. Note that the system expects the student to use forms such as <look over> instead of <review> because it focusses on the use of verbs with particles.

The following are some examples of the exercises. This is an actual session which has been annotated with comments preceded by semicolons. Input from the student appears in italics.

**;System presents sentences to
;translate from Spanish to English**

: exercise.

TRANSLATE THE FOLLOWING SENTENCES:

1.- Vanessa pone el disco. ;First sentence

: vanessa puts the record. ;Student types the answer

;System responds

You did not use the particle that goes with

<puts> in this sentence.

The corresponding complete translation of the verb is
<puts on>.

Even though the verb can be translated
into a one-word verb in Spanish,
in English you need both the verb and the particle

More? (Type 'yes.' or 'no.')

;Student repeats the exercise

or type 'repeat.' to repeat the exercise: *repeat.*

1.- Vanessa pone el disco.

: *vanessa puts the record on.*

;System's Response

Correct! Notice that the particle <on> can appear after the verb
or can be separated and moved after 'the record'.

More? (Type 'yes.' or 'no.')

or type 'repeat.' to repeat the exercise: *yes.*

;Student requests more
;information
;System provides examples
;of sentences in English
;(the Spanish translations)
;with the particular verb
;and their corresponding
;particles

Examples of uses of the verb <TO PUT> in English,

- 'Vanessa <put down> the suitcase.' or
- 'Vanessa <put> the suitcase <down>.'
(Vanessa <dejó> la maleta)
- 'Vanessa <put away> the old record-player.' or

- 'Vanessa <put> the old record-player <away>.'
(Vanessa <guardo> el tocadisco viejo)
- 'I would <put> the weight <at> about 10 pounds.'
(Yo <pondría> el peso <en> 10 libras)
- 'Please <put> the book <back> in its place.'
(Por favor <devuelve> el libro a su lugar)
- '<Put down> that weapon before you hurt somebody.'
(<Deja> ese arma antes de que hieras a alguien)
- '<Put on> some clothes.'
(<Ponte> una ropa) or (<Vístete>)
- 'How much money can he <put up>?'
(Cuanto dinero puede <contribuir>?)
- 'He <put off> the exam until tomorrow'.
(El <pospuso> el examen hasta mañana)

;Next exercise
;Student's response

2.- El libro que desapareció consta de dos capítulos.

: *the book that disappeared consists of two chapters.*

Correct! Notice that the preposition <of> appears after the verb <consists> and cannot be moved to the end of the sentence. It can be separated from the verb by an adverb as in the following:

- 'The book <consists> only <of> two chapters'

More? (Type 'yes.' or 'no.')

or type 'repeat.' to repeat the exercise: *yes.*

Examples of the use of the verb <TO CONSIST> in English:

- 'Education does not <consist> simply <in> learning many facts'
(La educación no <consiste en> sencillamente aprender muchos hechos)
- 'The full set <consists of> 32 glasses'
(El juego completo <consta de> 32 vasos)

Note that <to consist of> corresponds to <constar de> in Spanish while <consist in> in English corresponds to <consistir en> in Spanish. In general, they can be easily confused.

Appendix B contains an extended sample session.

5.6. User of VP²

It must be noted that the system is designed for relatively sophisticated users -- that is, for non-native speakers of English that have acquired most of the English vocabulary and grammar. As shown in the previous examples, the sentences are not extremely complicated but they have been designed to extract the essence of the problem of dealing with verbs and prepositions in English: the processing of verbs and prepositions/particles in a new language.

Because the system is geared to fluent speakers, it can handle acceptable translations whose word order differs from that of the exercise sentences. This is basically what "syntactic freedom" consists of, in allowing the user to type sentences such as:

TUTOR: TRANSLATE THE FOLLOWING SENTENCE:
Vanessa pone el disco.
STUDENT: The record vanessa puts on.

Sentences that have particles or prepositions that have been "illegally" placed are recognized as having those particles moved to the wrong position and the appropriate explanation is provided to the user.

5.7. Implementation

VP² has been implemented in Prolog, using the University of New Hampshire version 1.3 Prolog. Prolog was chosen as an implementation vehicle because it provides a good pattern matching mechanism and parsing techniques.

In general, in Prolog a grammar for any language can be described as a set of rules that specify the sequences of words that are acceptable as sentences in that language. The rules specify how the words must be grouped together into phrases and what the acceptable orderings of phrases are.

Prolog has a convenient grammar rule formalism: Definite Clause Grammars (DCG) [Pereira and Warren 80]. DCGs not only provide a description of a language but an effective means for analyzing strings in that language. *VP²* uses two grammars: One grammar for Spanish and one for English. These grammars include the dictionary of words that are part of the system. The grammars with their dictionaries appear in Appendix A. The Spanish Grammar forms the user model. Figure 5-1 shows the components of *VP²*.

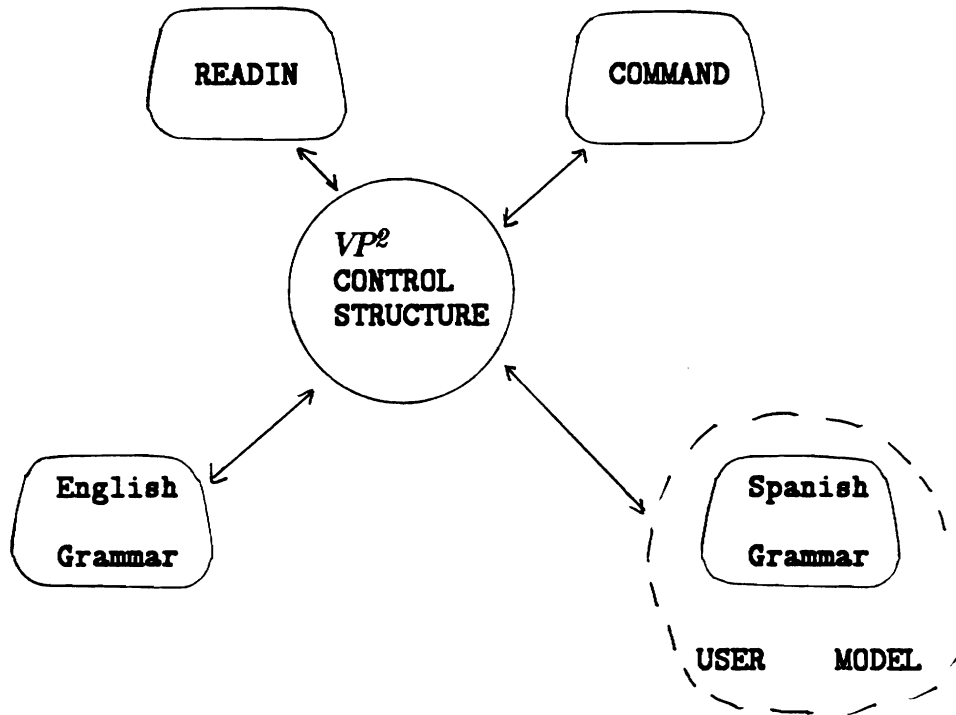


Figure 5-1: Components of *VP²*

Many verbs in English enter into special combinations with certain prepositions. In some cases, a verb cannot be used without the appropriate preposition. In other cases, the absence or presence of a preposition changes the way in which the central participants (subject and object) are interpreted. Based on such ideas, *VP²* parses the English sentence that the student types and extracts its verb and preposition (if it appears) or prepositions in order to check for the correct translation and checks it against the information provided in the user model about those verbs and prepositions in Spanish.

5.7.1. Components of VP^2

The three components of VP^2 are shown in Figure 5-1: READIN, COMMAND, and VP^2 CONTROL STRUCTURE.

READIN (adapted from Clocksin and Mellish¹⁶) reads in a sentence typed at the terminal and converts it into a list of Prolog atoms. The program knows when one word of the input ends and when one begins. The end of a sentence is recognized by the program to end when one of the following appears: '.', '!' or '?'. Upper-case characters are automatically converted to lower-case characters.

COMMAND contains a set of rules to parse commands that the user can type instead of the answer such as:

- 'help', which provides the student with a list of verbs and prepositions;
- 'example' which generates a sample of an exercise in the session;
- 'skip', which allows the student to skip over information or exercises;
- 'repeat', which allows the student to repeat an exercise;
- 'exercise', initiates the exercise session with the tutor.

VP^2 CONTROL STRUCTURE controls the interaction with the student in the following way: It presents the student with an introduction message to the system and explanation of the possible commands that s/he can type. Once the student requests the initiation of the exercise session, the system presents the student with

¹⁶ [Clocksin and Mellish 81]p. 87-88

an exercise sentence and reads in the student's response. The system parses the English sentence and extracts its verb and preposition, assuming that the sentence the student types is correct. If the student misspells a word, or provides an answer that the system cannot understand, *VP²* gives him/her another chance to answer. If the English parse fails, due to an incorrect or missing preposition, the system traces its parse and looks at the verb phrase to try to extract the verb (and the preposition, in the case of an incorrect one). Then, it looks into the user model to obtain information about the Spanish verb. From there, it may deduce the source of the error. If it believes that the error is caused by interference from Spanish, the system provides the correct answer along with information that indicates the error caused due this interference. If the system does not believe the error to be caused by interference, it provides the student with the correct answer and proceeds to the next exercise.

VP² generates the answer and explains possible reason for the error, especially in the context of Spanish and its differences and similarities to English. Whether or not errors are found, *VP²* looks at the user model to provide additional information that may prove helpful to the student.

While the user model contains a Spanish grammar, it does not use it all in its execution. It looks mainly at the Spanish verbs and their prepositions and the way Spanish verbs combine with those prepositions in the sentences. The grammar gives the system this information. We have chosen to use a relatively rich Spanish Grammar in *VP²* for extensibility: other tutors can then use other parts of the

grammar in tutoring other aspects of syntax such as correct verb-subject agreement, correct tense and number agreement, that are faulty due to interference [Taylor 75].

5.8. Procedure Used to Handle Responses

VP² has information about the verbs in each language and their prepositions. In English, the entries for verbs have six arguments, indicating form, root, tense, number, aspect, and prepositions. In Spanish, they have an additional entry indicating their regularity. For example:

IN ENGLISH:

```
is_verb(dream,dream,inf,_,_,of).
is_verb(dreamed,dream,past,_,_,about).
is_verb(pick,pick,inf,_,_,at).
is_verb(picked,pick,past,_,_,up).
is_verb(put,put,inf,_,_,in).
is_verb(put,put,past,_,_,away).
is_verb(put,put,present,_,_,out).
is_verb(puts,put,present,sing3,_,_).
```

IN SPANISH:

```
is_verb(soñe,soñar,past,sing1,intran,_,con).
is_verb(recoger,recoger,inf,_,_,_).
is_verb(pone,poner,present,sing3,tran,_,_).
```

In addition, a parse tree is stored for the correct answers to each exercise. These parse trees are matched against the user's answers in order to locate any errors in those answers. For now, we have chosen to store these parse trees because the system only has a few exercises.¹⁷ The system also has a table of direct

¹⁷If the system is expanded to work with a larger number of exercises, we will allow the system to generate the correct parse trees instead of having them stored.

translations of verbs and prepositions.¹⁸

VP^2 goes through the following procedure procedure in order to figure out the answer and provide the most adequate response.

- It presents the student with the Spanish sentence to translate.
- It accepts as INPUT the student's English translation, which it attempts to parse. The following will cause the parse to fail:
 - unknown words and misspellings.
 - unknown constructions outside the verb phrase.
 - missing preposition or particle when one is required.
 - wrong preposition or particle when one is required.
 - extra preposition or particle when one is not required.
- If the parse succeeds, it compares it with the stored parse tree for this sentence in order to make sure that it is indeed the appropriate translation. If the translation that the student provides can be parsed but it is not the translation of the given sentence, the system allows the student to try again. If the parse has succeeded and corresponds to the stored parse tree, VP^2 takes its verb phrase, looks up the direct translation of the verb and its preposition and matches this

¹⁸Remember that by 'direct translation' we mean the most common translation used for a certain word, the one one would find in looking up that word in a dictionary. This is different from the 'corresponding translation' which is the actual translation of the words and which vary in the context of the sentence. For example, the direct translation of <con> is <with> while the corresponding translation of <con> when used with the verb <soñar>, <to dream>, is <of> or <about>.

translation against the verb in the user model (i.e. the Spanish lexicon). Based on this it can indicate to the user any differences in the translations and any additional information about the verb in English or Spanish. For example, if the student types:

TUTOR: TRANSLATE THE FOLLOWING SENTENCE:

Yo *soñé* con los angeles.

STUDENT: I dreamed of the angels.

the system provides an answer such as:

Correct!

Note that the direct translation of <soñar con>
--<dream with>--does not exist in English.

In English you can also use <dream about>
in this sentence.

- If the parse fails, it matches the verb phrase from the failed parse against the previously stored correct parse to figure out the differences between the correct parse and that of the student's. From this comparison, it finds the wrong prepositions and/or missing particles. If it finds an incorrect preposition with a correct verb, it looks up the direct Spanish translation of the incorrect preposition. For example, if in the previous example the student typed:

I <dreamed with> the angels.

the system obtains the incorrect pair <dream with>, looks it up in its table of direct translations and finds it to be <soñar con>. *VP²* proceeds to its user model and finds this pair in it. Therefore it is able to explain the incorrect preposition. It generates an error message indicating the error as being due to interference from Spanish. It then goes to the English grammar and looks up the correct preposition(s) that go with the verb and provides the student with more

information, e.g. other prepositions that can be used and so on. For example, *VP²* provides the student with information about other preposition(s) that may be used with a verb:

**In English you can use <dreamed of> or
<dream about> in this sentence.**

5.9. Error Handling

The errors that students produce in using verbs with their prepositions in English, can be subcategorized into four different types:

5.9.1. Missing Particle

Where Spanish uses a single verb and English uses a verb particle combination, it is likely for native speakers of Spanish to translate the verb from Spanish without including the particle. For example

**TUTOR: Vanessa pone el disco.
STUDENT: vanessa puts the record.
TUTOR: You did not use the particle that goes with <puts>. The corresponding complete translation of the verb is <puts on>. Even though the verb can be translated into a one-word verb in Spanish, in English you need both the verb and the particle**

In this example, the system fails to parse the English sentence that the student has typed. Once the parse has failed, the system obtains the stored parse tree of the correct sentence, and matches it with the parse of the student's sentence in order to figure out the differences. From here, the system can notice that the particle that goes with the verb is missing -- that is, it obtains the pairs

<put>- _ and <put>-<on> and notices the differences. *VP²* then looks up in its translation table the direct translation of <put>, finds it to be <poner>. Next it looks at the information about the verb <poner> and preposition in its user model, finds that the <poner> verb in Spanish has no particle for this meaning. It then deduces the cause of the error, that there is a missing particle.

5.9.2. Additional Preposition

In Spanish, the verb <ir> --<to go>-- requires a preposition after it, e.g. <a> and <con>, which correspond to <to> and <with> respectively. Sentences that include the verb <ir> with <a> usually imply 'going to do something' as in the following example:

Se fue a correr.
He went to run.
'He went running.'

It may be the case that a native speaker of Spanish would use the additional preposition <to> and produce sentences such as:

*** He went to running.**
? He went to run.

In this case, the system checks the failed parse of the student's input, compares it with its own stored parse, and obtains the pairs <went>-<to> and <went>-_. It looks up the translation of <went> and <to> in its table, finds them to be <fue> and <a> respectively and looks this pair in the user model to deduce the occurrence of the additional preposition <to>. It then provides an appropriate answer indicating the redundancy.

5.9.3. Incorrect Preposition

Where both English and Spanish use a verb + preposition, but the prepositions don't correspond, another set of errors occur. For example, <pensar en> is translated as <think in>, <sonar con> is translated as <dream with>.

VP²'s approach to handling these errors is again by first figuring out if the sentence is correctly translated into English. Note that a sentence may be correct in English but it may not be the correct translation. If it is not, it obtains the 'correct' translation for the sentence, matches it against the one the student types and notices the differences. *VP²* then looks at the user model to obtain information about the verb and preposition in Spanish and then matches that information with the 'incorrect' English sentence from which it can find the errors caused by direct translation of the preposition from Spanish to English.

5.9.4. Spelling Errors or Unrelated Answers

When the system encounters a spelling error, a word that does not exist in its dictionary or a sentence whose meaning does not correspond to the one presented to the student, it responds by telling the student to try again. For example:

TUTOR: TRANSLATE THE FOLLOWING SENTENCE:
 El siempre trabaja en un problema hasta que lo
 resuelve.
 STUDENT: he always works in a prlem untl he solves it.
 TUTOR: Sorry, I cannot help you, please try again.

Clearly this could be improved with the inclusion of a spelling correcting module. Given that the student's response is highly constrained, such a module would have strong expectations to suggest possible misspellings.

CHAPTER VI

Conclusions

Based on observations and experiments done elsewhere, we have seen that people often rely on their knowledge of one domain when learning a new (similar) domain. That seems to be the case for language learning where students learning a second language use much of the knowledge they have of their native language.

VP² has addressed the question of how we can represent the knowledge of certain aspects of a language in a computer system and use this knowledge to provide the student of a second language with information tailored to his/her knowledge. It has also focussed on the problem of how correspondence of grammars of two languages can provide a sufficient basis for explaining the possible origin of grammatical errors made by native speakers of Spanish.

VP², provides some insights into how instruction in a second language might be augmented. With the system able to figure out the errors in the usage of verbs and prepositions and/or verbs with particles and explain them in terms of the student's native language, it is hoped that the learning process will become more efficient and therefore more satisfactory for the student.

6.1. Further Research

There are some issues that can be addressed by the work presented here. One is how systems like this can be adapted to students of different linguistic backgrounds. This may be achieved by installing into *VP²*'s user model the grammar of a language and seeing the different ways in which the student's native language might affect his/her learning of a second one. It will be important to notice how the different languages are processed by the system and how it handles its possible explanation of errors due to the native language. For example, in Hungarian¹⁹ the prepositions are attached at the end of the noun or verb as in the following example:

ENGLISH		HUNGARIAN
space	=	'ür
to dream	=	'álmodni

ENGLISH: I dream about space.

HUNGARIAN: 'ürröl 'ámodok'
about space I dream

Do these prepositions, with their very different dispositions, interfere in the English speech of Hungarians? Can a model such as the one developed in *VP²* explain and account for errors due to L_1 being Hungarian? *VP²* should behave as a domain-independent system and handle these problems too. It should also be as modular as possible due to the variation of its user models.

¹⁹I thank G. Herman for this example.

VP² has addressed errors due to interference of a native language. It has focussed on the incorrect use of prepositions and particles in English. One interesting issue to address next will be the "illegal movement of particles" in English. That is, once the student has indeed acquired the verbs with particles and overcomes the missing particle error due to interference, it will necessary to integrate into the system a mechanism for handling particles that appear in 'incorrect' positions. For example:

TUTOR: Vanessa pone el disco.
STUDENT: on vanessa puts the record.

Here, the tutor cannot just disregard this answer as incorrect. It must realize that the student has problems in placing the particles, maybe because s/he is not used to them. The system should be able to provide an answer such as the following:

TUTOR: You used the particle <on> in the wrong position.
 In this sentence you can place the particle
 after the verb or after the object
 ('the record') as
 in the following:
 - ' Vanessa puts <on> the record.'
 - ' Vanessa puts the record <on>.'

A next step in *VP²*'s enhancement would be to focus in other error correction aspects such as tense agreement in subjects, verbs and objects, and different grammatical forms such as questions and negative responses that may be incorrect due to interference.

Further work will be needed on how one's native language interferes with the

learning of a second language. Much of this work depends on what is similar and what is different between the two language systems. There may be other issues that do not fall in these two categories, and that may provide other approaches to viewing errors made in second language learning.

APPENDIX A

Grammars

Both the grammars for English and Spanish were written using the standard DCG notation provided by Prolog. These DCG rules stand as "shorthand" for ordinary Prolog code and their notation makes the code easier to read because it suppresses information that may not be interesting.

The notation is built around the "standard" notation for Context-free grammars. When these rules are passed on to Prolog, they are automatically translated into Prolog clauses. The translation basically consists in changing every predicate into a predicate with two arguments. Also, whenever a grammar rule indicates that there are rules that follow, the translation must also indicate that they arguments will match. The system also knows how to translate those rules introduce actual words into the lists forming the arguments of the predicates. For example,

```
determiner(D,Num) --> [D],{is_det(D,Num)}.
```

translates into:

```
determiner(D,Num,[D|S],S).
```

```

/* TWOGRAM: DCGrammar for the VP2 tutorial system */
/* The following are the rules for a grammar that will
be used to parse a sentence. */
/* Defines the rules for the determiners and nouns */
determiner(D,Num) --> [D],{is_det(D,Num)}.
noun(N,Num) --> [N],{is_noun(N,Num)}.
proper_noun(N,Num) --> [N],{is_prop_noun(N,Num)}.
/* Defines the possible noun phrases */
noun_phrase(N,Num,P) --> proper_noun(N,Num).
noun_phrase(N,Num,P) --> noun(N,Num).
noun_phrase((D,N),Num,P) --> determiner(D,Num),
noun(N,Num).
noun_phrase((D,N,X),Num,P) --> determiner(D,Num),
noun(N,Num),rel_clause(X).
noun_phrase((D,N,X),Num,P) --> determiner(D,Num),
noun(N,Num),prep_phrase(X,P).
noun_phrase((D,Adj,N),Num,P) --> determiner(D,Num),
adjective(Adj),noun(N,Num).
noun_phrase((Adj,N),Num,P) --> adjective(Adj),noun(N,Num).
/* Defines the adjective phrases */
adjective(Adj) --> [Adj],{is_adj(Adj)}.
/* The rules for the verb and verb_phrases */
aux_verb(Aux,Num,Tense) --> [Aux],{is_aux(Aux,Root,Tense,Num,_)}.
trans_verb(V,Num,Tense) --> [V],{is_verb(V,Root,Tense,Num,tran,P)}.
intrans_verb(V,Num,Tense) --> [V],{is_verb(V,Root,Tense,Num,intran,P)}.
verb_phrase((V,X),Num,Tense,V,P) --> trans_verb(V,Num,Tense),
noun_phrase(X,Num1,P),prep_phrase(X,P).
verb_phrase((V,_),Num,Tense,V,P) --> trans_verb(V,Num,Tense),
noun_phrase(X,Num1,P).
verb_phrase((V,X),Num,Tense,V,P) --> intrans_verb(V,Num,Tense),
prep_phrase(X,P).
verb_phrase((V,_),Num,Tense,V,P) --> intrans_verb(V,Num,Tense).
/* The prepositional phrases */
prep_phrase((P,X),P) --> prep(P),noun_phrase(X,Num,P).
prep_phrase(P,P) --> prep(P).
prep(P) --> [P],{is_prep(P)}.
/* The Relative clauses */
cl_head(C) --> [C],{is_cl_head(C)}.
rel_clause((C,VP)) --> cl_head(C),verb_phrase(VP,Num,Tense,V,P).

```

```

rel_clause((C,NP,VP)) --> cl_head(C),noun_phrase(NP,Num,P),
    verb_phrase(VP,Num,Tense,V,P).

/* And the rules for the sentences */

sentence(sentence(noun_phrase(NP),verb_phrase((V,X))),Num,Tense,V,P) -->
    noun_phrase(NP,Num,P),verb_phrase((V,X),Num,Tense,V,X).
sentence(sentence(noun_phrase(NP),verb_phrase((V,X)),
    prep_phrase((P,N1))),Num,Tense,V,P) -->
    noun_phrase(NP,Num,P),verb_phrase((V,X),Num,Tense,V,X),
    prep_phrase((P,N1),P).
sentence(sentence(noun_phrase(NP),verb_phrase((V,X)),
    prep_phrase((P,N1))),Num,Tense,V,P) -->
    noun_phrase(NP,Num,P),noun_phrase(NP,Num,P),
    verb_phrase((V,X),Num,Tense,V,X),
    prep_phrase((P,N1),P).
sentence(sentence(noun_phrase(NP),aux_verb(Aux),
    verb_phrase((V,X))),Num,Tense,V,P) -->
    noun_phrase(NP,Num,P),aux_verb(Aux,Num,Tense),
    verb_phrase((V,X),Num,Tense1,V,X).
sentence(sentence(noun_phrase(NP),aux_verb(Aux),verb_phrase((V,X)),
    prep_phrase((P,N1))),Num,Tense,V,P) -->
    noun_phrase(NP,Num,P),aux_verb(Aux,Num,Tense),
    verb_phrase((V,X),Num,Tense1,V,X),prep_phrase((P,N1),P).

/* These rules are not used at the moment, they may be used when
    expanding the system */

sentence(sentence(aux_verb(Aux),noun_phrase(NP),
    verb_phrase((V,X))),Num,Tense,V,P) -->
    aux_verb(Aux,Num,Tense),noun_phrase(NP,Num,P),
    verb_phrase((V,X),Num,Tense1,V,X).
sentence(sentence(wh_word(Wh),verb_phrase((V,X))),Num,Tense,V,P) -->
    [Wh],{is_wh(Wh)},verb_phrase((V,X),Num,Tense,V,X).
sentence(sentence(verb_phrase(verb(Cop),noun_phrase(NP),Adj)),
    Num,Tense,V,P) -->
    [Cop],{is_verb(Cop,be,Tense,Num,tran)},
    noun_phrase(NP,Num,P),adjective(Adj).
sentence(sentence(noun_phrase(NP),verb_phrase(verb(Cop),Adj)),
    Num,Tense,V,P) -->
    noun_phrase(NP,Num,P),[Cop],{is_verb(Cop,be,Tense,Num,tran)},
    adjective(Adj).
sentence(sentence(verb(Cop),noun_phrase(NP),noun_phrase(NP1)),
    Num,Tense,V,P) -->
    [Cop],{is_verb(Cop,be,Tense,Num,tran)},
    noun_phrase(NP,Num,P),noun_phrase(NP1,Num,P).

```

```

/* ESGRAM: DCGrammar for Spanish in VP2. This Grammar
   is basically the user model representation
   in the system */

/* The following are the rules for the Spanish grammar that will
   be used to parse a sentence. */

/* Defines the rules for the determiners and nouns */

determiner(D,Num,Sex) --> [D],{is_det(D,Num,Sex)}.

noun(N,Num,Sex) --> [N],{is_noun(N,Num,Sex)}.

proper_noun(N,Num,Sex) --> [N],{is_prop_noun(N,Num,Sex)}.

/* Defines the possible noun phrases */

noun_phrase(N,Num,Sex,P) --> proper_noun(N,Num,Sex).
noun_phrase(N,Num,Sex,P) --> noun(N,Num,Sex).
noun_phrase((D,N),Num,Sex,P) --> determiner(D,Num,Sex),
    noun(N,Num,Sex).
noun_phrase((D,N,X),Num,Sex,P) --> determiner(D,Num,Sex),
    noun(N,Num,Sex),rel_clause(X).
noun_phrase((D,N,X),Num,Sex,P) --> determiner(D,Num,Sex),
    noun(N,Num,Sex),prep_phrase(X,P).
noun_phrase((D,Adj,N),Num,Sex,P) --> determiner(D,Num,Sex),
    noun(N,Num,Sex),adjective(Adj,Num,Sex).
noun_phrase((Adj,N),Num,Sex,P) --> noun(N,Num,Sex),adjective(Adj,Num,Sex).
noun_phrase((Adj,N),Num,Sex,P) --> adjective(Adj,Num,Sex),noun(N,Num,Sex).

/* Defines the adjective phrases */

adjective(Adj,Num,Sex) --> [Adj],{is_adj(Adj,Num,Sex)}.

/* The rules for the verb and verb_phrases */

aux_verb(Aux,Num,Tense) --> [Aux],{is_aux(Aux,Root,Tense,Num,_,_)}.

trans_verb(V,Num,Tense) --> [V],{is_verb(V,Root,Tense,Num,tran,_,P)}.

intrans_verb(V,Num,Tense) --> [V],{is_verb(V,Root,Tense,Num,intran,_,P)}.

verb_phrase((V,X),Num,Tense,V,P) --> trans_verb(V,Num,Tense),
    noun_phrase(X,Num1,Sex,P),prep_phrase(X,P).
verb_phrase((V,_),Num,Tense,V,P) --> trans_verb(V,Num,Tense),
    noun_phrase(X,Num1,Sex,P).
verb_phrase((V,X),Num,Tense,V,P) --> intrans_verb(V,Num,Tense),
    prep_phrase(X,P).
verb_phrase((V,_),Num,Tense,P) --> intrans_verb(V,Num,Tense).

```

```

/* The prepositional phrases                                     */
prep_phrase((P,X),P) --> prep(P),noun_phrase(X,Num,Sex,P).
prep_phrase(P,P) --> prep(P).
prep(P) --> [P],{is_prep(P,_)}.

/* The Relative clauses                                       */
cl_head(C) --> [C],{is_cl_head(C)}.

rel_clause((C,VP)) --> cl_head(C),verb_phrase(VP,Num,Tense,V,P).
rel_clause((C,NP,VP)) --> cl_head(C),noun_phrase(NP,Num,Sex),
    verb_phrase(VP,Num,Tense,V,P).

/* And the rules for the sentences                             */
sentence(sentence(noun_phrase(NP),verb_phrase((V,X))),Num,Tense,V,P) -->
    noun_phrase(NP,Num,Sex,P),verb_phrase((V,X),Num,Tense,V,P).
sentence(sentence(noun_phrase(NP),verb_phrase((V,X)),
    prep_phrase((P,N1))),Num,Tense,V,P) -->
    noun_phrase(NP,Num,Sex,P),verb_phrase((V,X),Num,Tense,V,P),
    prep_phrase((P,N1),P).
sentence(sentence(verb_phrase((V,X))),Num,Tense,V,P) -->
    verb_phrase((V,X),Num,Tense,V,P).
sentence(sentence(verb_phrase((V,X)),prep_phrase((P,N1))),Num,Tense,V,P) -->
    verb_phrase((V,X),Num,Tense,V,P),prep_phrase((P,N1),P).
sentence(sentence(noun_phrase(NP),aux_verb(Aux),verb_phrase((V,X))),
    Num,Tense,V,P) -->
    noun_phrase(NP,Num,Sex,P),aux_verb(Aux,Num,Tense),
    verb_phrase((V,X),Num,Tense1,V,P).
sentence(sentence(noun_phrase(NP),aux_verb(Aux),verb_phrase((V,X)),
    prep_phrase((P,N1))),Num,Tense,V,P) -->
    noun_phrase(NP,Num,Sex,P),aux_verb(Aux,Num,Tense),
    verb_phrase((V,X),Num,Tense1,V,P),prep_phrase((P,N1),P).

/* These rules are not used with the present system, they may be used in
   future versions                                             */
sentence(sentence(aux_verb(Aux),noun_phrase(NP),verb_phrase((V,X))),
    Num,Tense,V,P) -->
    aux_verb(Aux,Num,Tense),noun_phrase(NP,Num,Sex,P),
    verb_phrase((V,X),Num,Tense1,V,P).

```

```

sentence(sentence(wh_word(Wh), verb_phrase((V,X))), Num, Tense, V, P) -->
  [Wh], {is_wh(Wh)}, verb_phrase((V,X), Num, Tense, V, P).
sentence(sentence(verb_phrase(verb(Cop), noun_phrase(NP), Adj)),
          Num, Tense, V, P) -->
  [Cop], {is_verb(Cop, ser, Tense, Num, tran, _)},
  noun_phrase(NP, Num, Sex, P), adjective(Adj, Num, Sex).

```

```

sentence(sentence(noun_phrase(NP), verb_phrase(verb(Cop), Adj)),
          Num, Tense, V, P) -->
  noun_phrase(NP, Num, Sex, P), [Cop],
  {is_verb(Cop, ser, Tense, Num, tran, _)},
  adjective(Adj, Num, Sex).
sentence(sentence(verb_phrase(verb(Cop), noun_phrase(NP), noun_phrase(NP1))),
          Num, Tense, V, P) -->
  [Cop], {is_verb(Cop, ser, Tense, Num, tran, _)},
  noun_phrase(NP, Num, Sex, P), noun_phrase(NP1, Num, Sex, P).

```

/* The dictionary entries look like these: */

IN ENGLISH:

```

is_prop_noun(vanessa, sing3).
is_noun(record, sing3).
is_verb(dream, dream, inf, __, __, of).
is_verb(dream, dream, inf, __, __, about).
is_prep(with).

```

IN SPANISH

```

is_prop_noun(vanessa, sing3, fem).
is_noun(disco, sing3, masc).
is_verb(sone, sonar, past, sing1, intran, __, con).
is_prep(con, __).

```


APPENDIX B

Sample Session

Comments are preceded by a semicolon. Input from the user is shown in italics.

```

$ unhp
          --- UNH Prolog 1.3 ---

[ up:[prolog.library]listing consulted ]

[ USR:[ETHEL.MSTHESIS]PROLOG.INI consulted ]
| ?- [twomain].                ;TWOMAIN loads all the files
[ auto-load /up_/prolog/expand ]

[ /up_/prolog/expand loaded ]

[ twogram consulted ]          ;TWOGRAM is the English
                               ;Grammar
[ twodict consulted ]          ;TWODICT is the English
                               ;dictionary
[ readin consulted ]           ;File to read sentences from
                               ;the terminal and convert
                               ;them into Prolog atoms
[ VP2 consulted ]             ;MAIN program

[ esgram consulted ]           ;Spanish Grammar
[ esdict consulted ]           ;Spanish Dictionary

yes
[ twomain consulted ]

yes
| ?- VP2.
Welcome to VP2:
A Tutorial System for English Two-Word Verbs.

The system generates practice exercises for the use
of verbs and particles or prepositions in English

The system presents a sentence in Spanish and you are
expected to type the sentence in English with the appropriate
verb and/or verb and preposition.

VP2 *ONLY* corrects errors in verb-preposition
and/or verb-particle usage

Terminate all your answers with a period

```

Type 'end.' to exit.
 Type 'help.' for help.
 Type 'example.' for a sample session.
 Type 'exercise.' to begin the exercises in the tutorial.

: *help.* ;help

You can obtain a list of the verbs and prepositions
 that *VP²*'S dictionary has.

For a list of verbs, type 'verbs.'
 For a list of prepositions, type 'prepositions.'

All verbs include present, past tense and past participle.

: *verbs.*

[do,be,have,dream,disappear,put,pick,complain,come,
 think,start,make,repeat,consist,take,lean,try,know,
 break,reach,stay,leave,go,determine]

: *prepositions.*

[with,about,on,of,near,in,to,for,out,up,from,by]

: *example.*

TRANSLATE THE FOLLOWING SENTENCE:

El siempre trabaja en un problema hasta que lo resuelve.

- If you answer:

HE ALWAYS WORKS IN A PROBLEM UNTIL HE SOLVES IT.

VP² will respond:

You used the incorrect preposition <in>.
 In English you can use <work at> in this sentence.
 Note that the corresponding form of <work at> does not
 exist in Spanish.

- If you answer correctly, *VP²* will recognize the
 correct answer and provide relevant additional information.

- If your answer contains a typo or an unrelated
 sentence such as:

***I SEE THE SKY or
 HE ALAYS WORK AT A PROBLEM UNTIL HE ENDS.***

VP² will respond:

Sorry, I cannot help you, please try again.

More? (Type 'yes.' or 'no.'): *yes.*

Examples of uses of the verb <TO WORK> in English,

- 'While <working on> a project, ...'
- 'The painter <works in> oils'
- 'You will not get your stories published unless you <work in> a bit of human interest', as well as forms such as
- 'You will have to <work out> the problem by tomorrow' or
- 'You will have to <work> the problem <out> by tomorrow' and
- 'She is <working on> the use of prepositions for her Masters Degree'.

: *exercise.*

TRANSLATE THE FOLLOWING SENTENCES:

- 1.- Vanessa pone el disco.
: *vanessa puts the record.*

You did not use the preposition that goes with <puts>.

The complete translation of the verb is <puts on>. Even though the verb can be translated into a one-word verb in Spanish, in English you need both the verb and the particle

More? (Type 'yes.' or 'no.'): *yes.*

Examples of uses of the verb <TO PUT> in English,

- 'I would <put> the weight <at> about 10 pounds.'
- 'The manager <put away> the correspondence.' or
- 'The manager <put> the correspondence <away>.'
- 'Please <put> the book <back> in its place.'
- '<Put down> that weapon before you hurt somebody.'
- '<Put on> some clothes.'
- 'How much money can he <put up>?'
- 'He <put off> the exam until tomorrow'.

2.- El libro que desaparecio consta de dos capitulos.
: *the bok disappeared.*

Sorry, I cannot help you, please try again.

2.- El libro que desaparecio consta de dos capitulos.
: *the book that disappeared consists of two chapters.*

Correct! Notice that the preposition <of> appears after the verb <consists> and cannot be moved to the end of the sentence. It can be separated from the verb by an adverb as in the following:

- 'The book <consists> only <of> two chapters'

More? (Type 'yes.' or 'no.'): *yes.*

Examples of the use of the verb <TO CONSIST> in English:

- 'Education does not <consist> simply <in> learning many facts'
- 'The full set <consists of> 32 glasses'

Note that <to consist of> corresponds to <constar de> in Spanish while <consist in> in English corresponds to <consistir en> in Spanish. In general, they can be easily confused.

3.- El presidente recogio el libro.
: *the president picked the book.*

You did not use the preposition that goes with <To picked>. The complete translation of the verb is <picked up>. Even though the verb can be translated into a one-word verb in Spanish, in English you need both the verb and the particle

More? (Type 'yes.' or 'no.'): *yes.*

Examples of uses of <TO PICK> in English

- 'She <picked at> the food for a few minutes, then pushed the plate away'
- 'The teacher <picks on> the bad student all the time'
- 'The police <picked> the thief <up> on the street'
- 'She is quick at <picking up> any language'
- 'I went to the store and <picked up> this cheap shirt'
- 'Moris <picks up> all the gossip at the office'

4.- Moris sono con el viaje.
: *moris dreamed with the trip.*

You used the incorrect preposition <with>.
In English you can use <dreamed of> or <dream about> in this sentence. Note that the direct translation of <dreamed of>--<sonar de>--does not exist in Spanish.

More? (Type 'yes.' or 'no.'): *yes.*

Examples of sentence with <TO DREAM> in English
- 'I never <dreamed of> being an actress'
- 'For the new AT&T building, Louis Kahn has <dreamed up> the strangest structure.'

5.- El maestro se quejo del estudiante.
: *the teacher complained of the pupil.*

You used an inappropriate preposition for this sentence
<To complain of> is usually used in sentences such as
- 'The boy <complained of> acute stomach pains'

More? (Type 'yes.' or 'no.'): *yes.*

Examples of sentences using the verb <TO COMPLAIN>

- 'This student is always <complaining about> the teachers'
- 'The patient <complained of> indigestion'

From these exercises we can see that the direct translation of verbs and prepositions from one language to another is not a one-to-one correspondence.

yes
| ?- bye.
prolog execution halted

\$ log

References

- [Anderson et al. 84] Anderson, J.R., C.F. Boyle, R. Farrell, B.J. Reiser.
Cognitive Principles in the Design of Computer Tutors.
1984.
Unpublished - Advanced Computer Tutoring Project - CMU.
- [Bailey et al. 74] Bailey, N., C. Madden and S. Krasken.
Is There a 'Natural Sequence' in Adult Second Language
Learning?
Language Learning 24(2):235-243, December, 1974.
- [Bates and Wilson 80] Bates, M. and K. Wilson.
Language Instruction Without Pre-Stored Examples.
In *Proceedings Third Canadian Symposium on Instructional
Technology*. Vancouver, Canada, February, 1980.
- [Bates et al. 81] Bates, M., J. Beinashowitz, R. Ingria, K. Wilson.
Generative Tutorial Systems.
In *CBI: Association for the Development of Computer-Based
Instructional Systems*. Atlanta, Georgia, March, 1981.
- [Brown et al. 75] Brown, J.S., R.R. Burton and A.G. Bell.
SOPHIE: A Step Towards a Reactive Learning Environment.
International Journal of Man-Machine Studies, 1975.
- [Brown et al. 82] Brown, J.S., R.R. Burton and J. de Kleer.
Pedagogical, Natural Language and Knowledge Engineering
Techniques in SOPHIE I, II and III.
In D. Sleeman and J.S. Burton (editor), *Intelligent Tutoring
Systems*, chapter 11, pages 227-282. Academic Press, New
York, 1982.
- [Carbonell 70] Carbonell, J.R.
AI in CAI: An Artificial Intelligence Approach to Computer-
Aided Instruction.
IEEE Transactions on Man-Machine Systems
MMS-11:190-202, 1970.
- [Cassidy 83] Cassidy, J.E.
Using Computers to Teach Foreign Languages to Adults.
CALICO Journal 1(2):14-16, September, 1983.
- [Chomsky 80] Chomsky, N.
Rules and Representations.
University Press, New York, 1980.

- [Clocksin and Mellish 81]
 Clocksin, W.F. and C.S. Mellish.
Programming in Prolog.
 Springer-Verlag, Berlin, Germany, 1981.
- [Collins et al. 75] Collins, A., E. Warnock, N. Aiello, M. Miller.
 Reasoning From Incomplete Knowledge.
 In Bobrow, D. and A. Collins (editor), *Representation and Understanding: Studies in Cognitive Science*, pages 383-414.
 Academic Press, Inc., New York, 1975.
- [Cook 73] Cook, V.
 The Comparison of Language Development in Native Children
 and Foreign Adults.
IRAL 11:13-28, 1973.
- [Cooper et al. 79] Cooper, R., E. Oshtain, R. Tucker, M. Waterbury.
 The Acquisition of Complex English Structures by Adult Native
 Speakers of Arabic and Hebrew.
Language Learning 29(2):255-276, 1979.
- [Cowie and Mackin 82]
 Cowie, A.P., R. Mackin.
Oxford Dictionary of Current Idiomatic Expressions.
 Oxford University Press, Oxford, 1982.
- [Dommergues and Lane 76]
 Dommergues, J. and H. Lane.
 On Two Independent Sources of Error in Learning the Syntax of
 a Second Language.
Language Learning 26(1):111-123, 1976.
- [Douglas and Moran 83]
 Douglas, S.A. and T.P. Moran.
 Learning Operator Semantics by Analogy.
 In *Proceedings AAAI-83*, pages 100-103. Washington, D.C.,
 August, 1983.
- [Dulay and Burt 74]
 Dulay, H. and M.K. Burt.
 A New Perspective on the Creative Construction Process in
 Child Second Language Acquisition.
Language Learning 24(2):253-278, December, 1974.

- [Farrell et al. 84] Farrell, R.G., J.R.Anderson and B.J. Reiser.
An Interactive Computer-based Tutor for LISP.
In *Proceedings AAAI-84*, pages 106-109. University of Texas at
Austin, August, 1984.
- [Flynn 83] Flynn, Suzanne.
*A Study of the Effects of Principal Branching Direction in
Second Language Acquisition: The Generalization of a
Parameter of Universal Grammar from First to Second
Language Acquisition.*
PhD thesis, Cornell University, January, 1983.
- [Gass 80] Gass, S.
Investigation of Syntactic Transfer in Adult Second Language
Learners.
In S. Krasken and R. Scarcella (editor), *Issues in Second
Language Research*. Newbury House, Rowley, MA, 1980.
- [Gazdar 82] Gazdar, G.
Phrase Structure Grammar.
In Jakobson, P. and G.K.Pullum (editor), *The Nature of
Syntactic Representation*. D. Reidel Publishing, Dordrecht,
1982.
- [Genesereth 79] Genesereth, M.
The Role of Plans in Automated Consultation.
In *Proceedings IJCAI-79*, pages 311-319. August, 1979.
- [Guitart 76] Guitart, J. M.
Markedness and a Cuban Dialect of Spanish.
Georgetown University Press, Washington, D.C., 1976.
- [Halasz and Moran 82] Halasz, F. and T.P. Moran.
Analogy Considered Harmful.
In *Proceedings CHI-82*. Gaithersburg, MD., March, 1982.
- [James 81] James, C.
Contrastive Analysis.
Longman, England, 1981.
- [Joshi 82] Joshi, A.K.
Mutual Beliefs in Question Answering Systems.
In N. Smith (editor), *Mutual Beliefs*, chapter 4. Academic Press,
New York, 1982.

- [Kean 79] Kean, M.-L.
Grammatical Representations and the Description of Language Processing.
In D. Caplan (editor), *Biological Studies of Mental Processes*, pages 239-268. MIT, Cambridge, MA, 1979.
- [Lado 57] Lado, R.
Linguistic Across Cultures.
University of Michigan Press, Ann Arbor, MI, 1957.
- [London and Clancey 82] London, B. and W.J. Clancey.
Plan Recognition Strategies in Student Modelling.
In *Proceedings AAAI-82*, pages 335-338. Pittsburgh, PA, August, 1982.
- [McCoy 83] McCoy K.
Cooperative Responses to Object-Related Misconceptions.
Technical Report MS-CIS-83-39, University of Pennsylvania, November, 1983.
- [Pereira and Warren 80] Pereira, F. and D.H. Warren.
Definite Clause Grammars for Language Analysis-A Survey of the Formalism and a Comparison with Augmented Transition Networks.
Artificial Intelligence 13(1):231-278, 1980.
- [Pica 83] Pica, T.
Adult Acquisition of English as a Second Language Under Different Conditions of Exposure.
Language Learning 32(4):465-497, 1983.
- [Pulman 84] Pulman, S. G.
Limited Domain Systems for Language Teaching.
In *Proceedings of COLING-84*, pages 84-87. Stanford University, California, July, 1984.
- [Rich 79] Rich, E.
User Modeling via Stereotypes.
Cognitive Science 3:329-354, 1979.
- [Richards 71] Richards, J.C.
A non-contrastive Approach to Error Analysis.
English Language Teaching 25:204-219, 1971.

- [Rumelhart and Norman 81]
 Rumelhart, D.E. and D.A. Norman.
 Analogical Processes in Learning.
 In J.R. Anderson (editor), *Cognitive Skills and their Acquisition*, pages 335-360. Erlbaum, Hillsdale, New Jersey, 1981.
- [Rutherford 82] Rutherford, W.
 Markedness in Second Language Acquisition.
Language Learning 32(1):85-108, 1982.
- [Schustack and Anderson 79]
 Schustack, Miriam and J. R. Anderson.
 Effects of Analogy to Prior Knowledge on Memory for New Information.
Journal of Verbal Learning and Verbal Behavior 18:565-583, 1979.
- [Schuster 83] Schuster, E.
Custom Made Responses: Maintaining and Updating the User Model.
 Technical Report MS-CIS-83-13, University of Pennsylvania, July, 1983.
- [Sciarone 70] Sciarone, A.G.
 Contrastive Analysis: Possibilities and Limitations.
IRAL 8(2), 1970.
- [Scott and Tucker 74]
 Scott, M.S. and G.R. Tucker.
 Error Analysis and English-Language Strategies of Arab Students.
Language Learning 24(1):69-97, June, 1974.
- [Sleeman 82] Sleeman, D.
 Inferring (Mal) Rules from Pupil's Protocols.
 In *Proceedings of ECAI-82*, pages 160-164. Orsay, France, August, 1982.
- [Stevens et al. 79]
 Stevens, A., A. Collins and S.E. Goldin.
 Misconceptions in Student's Understanding.
Inter. Journal Man-Machine Studies 11:145-156, 1979.

- [Taylor 75] Taylor, B. P.
The Use of Overgeneralization and Transfer Learning Strategies
by Elementary and Intermediate Students of ESL.
Language Learning 25(1):73-107, 1975.
- [Wallis and Shortliffe 81] Wallis, J.W. and E.H.Shortliffe.
Explanatory Power for Medical Expert Systems.
Technical Report, Stanford University Heuristic Programming
Project, December, 1981.
- [Weischedel et al. 78] Weischedel, R.M., W. M. Voge, M. James.
An Artificial Intelligence Approach to Language Instruction.
Artificial Intelligence 10:225-240, 1978.
- [Winston 80] Winston, P.H.
Learning and Reasoning by Analogy: The Details.
AI Memo 520, MIT, May, 1980.
- [Woods et al. 72] Woods, W.A., R. Kaplan, B. Nash-Webber.
*The Lunar Sciences Natural Language Information System:
Final Report.*
Report 2388, Bolt Beranek and Newman, June, 1972.