Pergamon

PII: S0360-1315(97)00037-7

Computers Educ. Vol. 29, No. 4, pp. 163–180, 1997

Printed in Great Britain
0360-1315/97 $17.00+0.00

CONVERSATION REBUILDING: FROM THE FOREIGN LANGUAGE CLASSROOM TO IMPLEMENTATION IN AN INTELLIGENT TUTORING SYSTEM

ALESSANDRO MICARELLI[1]* and PATRICK BOYLAN[2]

[a]Dipartimento di Informatica e Automazione, Università di Roma Tre, Via della Vasca Navale 79, 00146 Roma, Italy

[2]Dipartimento di Linguistica, Università di Roma Tre, Via Castro Pretorio 20, 00195 Roma, Italy

(Received 15 May 1997; accepted 12 August 1997)

Abstract—This paper shows how an innovative "communicative" technique in teaching foreign languages—Conversation Rebuilding (CR)—readily lends itself to implementation in an Intelligent Tutoring System (ITS). Classroom language teachers using CR get students to formulate acceptable utterances in a foreign idiom by starting from rough approximations (using words the students know) and gradually zeroing in on the utterance which a native speaker of that idiom might produce in a similar setting. The ITS presented here helps students do the "zeroing in" optimally. It lets them express themselves temporarily in an "interlingua" (i.e., in their own kind of French or English or whatever they are studying), as long as they make something of their communicative intent clear, that is, as long as the System can find a semantic starting point on which to build. The ITS then prods the students to express themselves more intelligibly, starting from the "key" elements (determined by a heuristic based on how expert classroom teachers proceed) and taking into consideration the students' past successful or unsuccessful attempts at communication. To simplify system design and programming, however, conversations are "constrained": students playact characters in set dialogs and aim at coming up with what the characters actually say (not what they could possibly say). While most Intelligent Computer Assisted Language Learning (ICALL) focuses the attention of students on norms to acquire, the ICALL implementation of CR presented in this paper focuses the attention of students on saying something—indeed, almost anything—to keep the conversation going and get some kind of meaning across to the other party. It sees successful language acquisition primarily as the association of forms with intent, not simply as the conditioning of appropriate reflexes or the elaboration/recall of conceptualized rules (which are the by-products of successful communication). Thus, in espousing this hard-line communicative approach, the present paper makes a first, non-trivial point: ICALL researchers might usefully begin by investigating what the more able teachers are doing in the classroom, rather than by building elaborate computer simulations of out-dated practices, as happens all too often. The paper then goes on to describe the architecture of a prototype ITS based on CR—one that the authors have actually implemented and tested—for the acquisition of English as a foreign language. A sample learning session is transcribed to illustrate the man–machine interaction. Concluding remarks show how the present-day limits of ICALL (and Artificial Intelligence in general) can be partially circumvented by the strategy implemented in the program, i.e. by making the students feel they are creatively piloting an interaction rather than being tested by an unimaginative machine.

INTRODUCTION

This paper presents an innovative technique typical of the recent "communicative revolution" in language learning[1]—Conversation Rebuilding (CR)—and shows how it readily lends itself to implementation in an Intelligent Tutoring System (ITS) designed to help Italian students master English as a foreign language.

Traditional foreign language teaching tends to view students as generators of isolated utterances, each one immediately labelled as "correct" or "incorrect" by an omniscient and implacable Tutor. The "communicative approach", on the other hand, views verbal exchange as a collaborative relationship: the sense of a particular utterance and the acceptability of a particular linguistic construction are something which the parties involved negotiate to some extent (this also holds among native speakers when they create or use slang and, as conversational analysts have shown, in everyday talk as well [8,9]). Language acquisition is seen as an evolutionary process: students acquire "degrees" of English or French or Russian, each "degree" constituting an intermediary version of the target language (or "interlingua"), with its own internal consistencies to be taken into consideration.

These two principles may be illustrated by the following exchange in a typical English-as-a-Foreign-

* To whom all correspondence should be addressed. E-mail: micarel@inf.uniroma3.it

[1] For a fuller explanation of the "communicative revolution" of the 1970s, see: van Ek [1], Brumfit and Johnson [2], Corder [3], Littlewood [4], Johnson [5], Oller and Richard-Amato [6], and Ellis [7].

Language classroom, where the teacher is using the "communicative approach". STUDENT: "That's intelligenter." TEACHER: "Huh?" Or, if a more explicit prompt is required: "Huh? ... More what?" Or if a still more explicit prompt is called for: "'Intelligenter?' Ah, you mean 'smarter', huh? ... Or 'more intelligent?'" To which the student may reply "Yes, smarter", "Yes, more intelligent" or simply "Yes", meaning "You got the idea", and go on. (The degree of explicitness called for is a function of the history of the interaction up to that point.) Thus, in "communicative" teaching, grammaticality is seen as a means of making oneself understood, not an end in itself; the only really unpardonable "error" is to fail to get what one means across. This does NOT mean that grammatical sloppiness goes unchallenged; it is indeed challenged, but as an information-flow (not normative) bug, similar to the bugs that characterize sloppy programming.

Thus, the "communicative approach" sees language teachers who playact or dialog with their students (and conversationalists in general) as intelligent pattern matchers, people who try to assign meaning to utterances on the basis of instances of similar constructions stored in memory, taking into consideration context, probable intent, and what has been said previously. When they encounter utterances that "jar" linguistically or sense-wise, they make a stab at a possible meaning and check it out, calling into question whatever "jars". This prods their interlocutors into eliminating the discrepancies from their utterances, in order to make themselves understood as intended. If simple prodding proves insufficient, more explicit forms are used.

"Communicative" methods have proven to be invaluable since they offer—albeit through simulation—real-time construction of situated intentional utterances, the key to language acquisition [10] (see [11] for a counter argument[2]). Most communicative approaches to language instruction hold, in fact, that learners, to make progress, must alternate between experimenting instinctively with language to "get something done"—which is considerably more than just rote practice of "rules"—and reflecting on the consistencies discovered in the samples of language encountered—which is much more than simply studying a series of norms [2,10,12,13]. The question is how much emphasis to give each of these two phases and what kind of constraints to put on the students' "free" experimentation with and discovery of linguistic forms. Although the hard-line communicative approach advocated here favors the first (experimental) phase, it does not eliminate all guided reflection and conscious application. To do so would indeed smack of behaviorism and unaccountably ignore the findings of cognitivism and three decades of psycholinguistic experimentation.[3] This is particularly true with students who are studying a language which embodies a culture radically different from theirs.[4]

Conversation Rebuilding (CR), a recently developed "communicative technique" in language teaching, attempts to strike a happy medium between experimentation and control [16–18]. It avoids the old fashioned discrete-point approach, which led students to focus on single phonemes, single morphemes or lexical items, and single grammatical or pragmatic rules at a time [19]. Instead, CR "causes the learner to process sequences of elements in a language" which, while ideally conforming to "the normal contextual constraints of that language", constitute—first and foremost—purposeful "pragmatic mappings" onto "extralinguistic context" [19].[5]

Although CR requires considerable skill on the part of the teacher, the greater satisfaction reported by students has induced Dilit-International House, a language school located in Rome (Italy), to retrain all of its teachers in the technique. CR has also become the heart of the school's certification program for language teachers working in both private and public schools in Italy. Dilit-International House staff report that CR is now used in seven European countries and is rapidly spreading to other parts of the world as well.

The CR technique—and the linguistic and pedagogical principles underlying it—clearly lend themselves to implementation in an ITS and specifically in one based on an Artificial Intelligence (AI)

[2] Philip Swann, a professional language teacher as well as ICALL system designer, makes, in our view, the only possibly convincing case for rejecting "communicative" teaching and going back to traditional grammar-translation methodology: the latter is, on occasion, what the market seems to demand. See a full discussion of the pros and cons at the end of the section "Literature and Related Work".

[3] See Krashen [14] for the strongest arguments in this direction.

[4] Students of languages and cultures radically different from their own must first learn to "feel" differently, in order to begin to want to say things differently from the way they would have said them in their home culture—and this implies accepting the other culture. To speed up the process of acceptation, a contrastive study of the two cultures—eliminating stereotypes and relating the students' own culture—can be of help. See for example Sercu [15].

[5] We recall that the term linguistic context refers to "the physical stuff of language which is organized into ... verbal elements", while the extralinguistic context "has to do with the world, outside the language, as it is perceived by language users" [19].

diagnosis of user needs. The CR technique permits much more gratifying and didactically effective man–machine interaction than what one typically finds in the Intelligent Computer Assisted Language Learning (ICALL) systems described in the literature.

Most proposals for ICALL programs, in fact, remain anchored to the teaching philosophy that their authors apparently remember from school days (and which, in fact, may still be found in many schools and universities, in spite of the inadequate results such teaching is known to produce—cf. the Council of Europe report 1976 [20] and the report of the Department of Education and Science Working Group on Modern Foreign Languages 1990 [21]; see also Micarelli and Boylan [22]). For example, out of 20 articles on ICALL in the extremely stimulating volume edited by Swartz and Yazdani [23], 15 offer language learning support that is discrete-point as opposed to holistic and either rote or rule-oriented as opposed to intentional and *ad hoc* (locally-situated).[6] This does not mean that the authors—or the ITS, AI and ICALL communities in general—fail to recognize the value of truly communicative language teaching and learning.[7] Most do, but simply point out the difficulty of inventing communicative tasks for a computer to do that are programmable for present-day schoolroom PCs or Macs using currently available resources.[8]

The situation does not seem to have improved much a year later if we judge from the special issue on ITS research appearing in *Computers & Education* [44] or the year after that if we look at the Selected Papers from the EUROCALL Conference published by *Computers & Education* [45]. In [44], Jaspers *et al.* claim they espouse communicative principles but in fact aim at getting students to check their grammar and vocabulary more than their communicative effectiveness; Sciarone and Meijer rightly present words in whole texts, but opt for deliberately artificial ones which cannot function as vehicles of intentionality (the basis of communication). Whereas in [46], only six out of 20 articles describe an ITS based on helping students manage holistic, intentional, situated, on-going discourse (Farrington, Thomson, Hayet, Grezel and Sciarone, Rézeau, Little).

The ITS presented in this paper is an attempt at doing just that. It aims at simulating an "authentic performance environment", by exploiting the technique of Conversation Rebuilding. Developed in Common Lisp and running on a MacIntosh platform, the ITS is admittedly far too rigid and limited to qualify as truly "communicative"; nonetheless it has, we feel, the merit of moving in the right direction.

This paper first describes the CR teaching procedure as used in the better classrooms and then the implementation of a more restricted variety as a prototype of a future full-blown system. Subsequently the paper gives details on how student utterances are evaluated, presents a sample interaction with the system and concludes with general remarks on the significance of this project and on intelligent tutoring in general. The project presented here, in fact, raises an interesting theoretical issue. Does man–machine interaction have to be "authentic" to promote authentic learning of a foreign language? Or is it enough to give students the IMPRESSION they are engaging in authentic interaction in a foreign language to get them to link situated linguistic forms to states of intentionality (what we described earlier as the key to learning a foreign language)?

CONVERSATION REBUILDING AS PRACTICED IN THE CLASSROOM

Real conversations are dynamic processes in which two (or more) subjects pragmatically map extralinguistic data onto linguistic data, and vice versa [19]. Put another way, in a two-way conversation

[6] Here is a breakdown of the kind of language-learning provided by the programs described in Swartz and Yazdani [23]. Out of 20 contributions, fifteen focus on discrete-point learning: five deal with learning lexical items [24–28] and ten deal with learning grammatical items [29–38]. The remaining five systems described seek to offer the kind of holistic, intentional, situated, communicative learning advocated in the present paper: Legenhausen and Wolff [39] (story writing with STORYBOARD), Hamburger and Hashim [40] (getting animated figures to obey one's orders), Sussex [41] (obtaining appropriate help from the system), Wilks and Farwell [42] (dialoguing with stick people, writing letters), and Criswell *et al.* [43] (contextualizing reading passages). See bibliography for full references.

[7] It is interesting to note that most of the articles in Swartz and Yazdani [23]—together with the excellent in-depth introduction written by the editors themselves—endorse the kind of communicative learning that only a small minority of the contributors have tried to implement. Frederiksen *et al.* [32] sum up the contradiction. While the system they describe concentrates largely on a single point of grammar (modal verbs), the authors nonetheless maintain that a proper ICALL system ought to "be built around an authentic performance environment in which learners use and produce discourse in the second language in realistic situations that reflect the natural situations of discourse communication" (p. 118).

[8] Handke [27] explain why their implementation—and most of the others in Swartz and Yazdani [23]—do not aim at creating an "authentic performance environment" using AI technology: "It is an extremely complex task. Theoretical issues, programming demands and hardware requirements ... constrain the implementation to a large degree" (p. 304).

both subjects, as speakers, translate their perceptions, feelings, thoughts and intentions into appropriate and culturally authentic communicative behavior centered on utterances (linguistic realizations). Inversely, both subjects, as listeners, retranslate the utterances they hear into a set of perceptions, feelings, thoughts and intentions which they then ascribe to the other party [46,67].

Thus, the specific form that utterances take reveals a multitude of simultaneous, multifaceted and interrelated expressive choices which attempt to exploit the communicative potentiality of the language being used and the historical situation of the utterance.

In Conversation Rebuilding the students are led, slowly and purposefully, through the same selection process that native speakers perform instantaneously and subconsciously in conversing. This technique (as reported in Aiello *et al.* [47], and Aiello and Micarelli [48]) may be described as follows:

(1) The teacher begins with a target utterance in mind: if the students are viewing a film, it can be the response that one character is about to give another character at the point in which the teacher stops the film; if the students are reading an anecdote (e.g., a humorous dialog), it can be the tag line or the lines immediately preceding it. The teacher directs the attention of the class to all relevant extralinguistic (contextual) information available to the speakers. If the teacher is using a film full of subtle cultural references or if the anecdotal story takes place in a non-obvious setting, the extralinguistic information can be evoked by means of a combination of mime, pictures and language.

(2) The teacher invites the class to impersonate the character who is about to speak and to guess what, given the context, that character may WANT to say, e.g., what utterance effectively translates the character's probable perceptions, feelings, thoughts and intentions. Any utterance, however ungrammatical, can constitute a suitable starting point for CR.

(3) The first student to speak is asked to repeat the utterance she/he hypothesizes so that everyone can hear it. This marks the beginning of the learning process: the attention of the class is focused on the relationship between communicative intent and linguistic form.

(4) The teacher mentally compares the student's hypothesis with what the character in the film or anecdote is actually about to say. The task now is to get the whole class to modify the utterance hypothesized, one trait at a time, until it conforms to what the character actually says or, if the teacher decides to give free rein to the students, could conceivably say. Traits are normally examined one at a time (not in clusters, as would be preferable, since students are generally unable to work with simultaneous sets of variables in an unfamiliar phonemic/graphemic system).

The CR technique leaves it up to each teacher to decide whether or not students must come up with the exact words that appear in the film or anecdote. In point of fact, most teachers prefer demanding the very same words since this makes the elicitation process easier for them to handle (they must envisage only one possible target answer). And, as we will see when we turn to the implementation of the CR technique in an ITS, opting for only one acceptable answer clearly facilitates the task of programming the system.

But for the theoretical reasons given above regarding the way humans actually communicate, a truly resourceful teacher (and a truly humane ITS) ought to be willing to accept any response suggested by students, so long as it fits in perfectly well with the context and "does the job" that the original utterance does. In other words, a resourceful teacher and a humane AI system should aim at eliciting utterances from students in function of a target model that can vary continually as the students shade what they want to say. The only constraint is that the students' final utterance must correspond to the overall communicative intentionality required by the story line. Some teachers object that "variable response" CR (in a classroom, on a computer) may disorient students, since there would be no single "correct answer". And they are right. But that is the price students must pay to acquire linguistic flexibility, i.e. the capacity to deal with (and produce) unpredictable utterances. A teacher (course designer) who creates a safe world where students can predict all the utterances they meet and need to produce, does not prepare students for the real world.

Another objection is that "variable response" CR raises the question of who or what determines the grammaticality of a sentence; exercises in which "anything goes" would be pure anarchy, it is claimed. We beg to differ. While grammaticality is admittedly not the issue in communicative teaching, this does not mean that "anything goes". Any sentence that a student utters (or encounters, say, in a pop song) is "right" if it produces the intended effect: but it must produce that effect. When speaking with the Queen, one learns to use the Queen's English or else one gets cold shouldered—and the same holds when speaking with young people, bureaucrats or north coast villagers. "Correctness" exists, but only as a negotiable sociolinguistic variable. This is precisely the point students learn during a CR session in the classroom or in the computer lab: if the teacher winces or fellow students laugh or the computer balks, it is clear that the sentence uttered didn't work as intended, even if apparently constructed by the rules in the grammar book; and the opposite is just as true. Repeated attempts will teach students, not hard-and-fast rules, but rules of thumb (which is what they need in life). The teacher's role is to get them to support the "anxiety of indetermination" until they acquire the necessary rules of thumb.

In any case, the superiority of "variable response" CR is a purely academic question at present. Although we are firm believers in hard-line communicative teaching, we felt it opportune in our ITS work to use the "fixed response" kind of CR. We will explain our reasons later on but they can be summarized here in a sentence: although "correctness" is a sociolinguistic variable, linguists have only just begun to describe languages in such terms; and even if all data were available, the task of accounting for all sociolinguistic (pragmatic) parameters in a program would be staggering. Our response has therefore been to get students to want to imitate a character of a film or anecdote: the deviant sentences they produce can thus be handled not as "correct" or "incorrect" but as "close enough" or "not close enough" to what the character actually says in the film to be recognizable by a computer. We will claim later on that this constriction goes somehow unnoticed by many students: they feel THEY are the authors of the sentences they finally come up with and, in a certain sense, they are. (Thus our system promotes the marriage of speaker intentionality and authentic, sociolinguistically dense language.) The present discussion should make it clear, however, that we consider our "fixed response" prototype system as a stopgap measure, something that has simply permitted us to evaluate the benefits of CR, when automated. (We will discuss those benefits—and limitations—in our conclusion). In other words, we consider our prototype system simply a tiny first step in (what we consider to be) the right direction: a program that avoids the computational explosion of foreseeing all possible student responses and, at the same time, gives something more than the illusion of creativity.

(5) Having compared the students' hypothesis with a model response (fixed or variable), the teacher rapidly decides which trait to deal with first and, on the basis of this decision, comments on the hypothesis in such a way as to lead the students to modify it. Let us look at an example taken from an actual CR session recorded in an Italian language school. The students' syntax is typical of the intermediary language ("interlingua") that most Italian speakers go through in acquiring English.

Context:	New York–Rome late evening flight; magazine rack; Italian male passenger addresses American hostess
Intent:	Discover her interests through her magazine preferences
Target utterance:	What kind of magazine do YOU like reading?
Students' initial hypothesis:	You like read this... or this?
Differences:	(a) missing SUBJECT/AUXILIARY inversion—thus, the hostess hears a grammatical affirmation instead of a question and may find the contrast syntax/intonation confusing; (b) the base form of "read" instead of the gerund—the anacoluthon is disrupting; moreover it suggests an action instead of an activity, which is what in fact the passenger wants to inquire about; (c) use of deixis ("This or this?") instead of a generalization ("kind of magazine")—thus, the hostess hears an intrusive closed question instead of a more discreet open one.

The teacher would probably decide to deal with the first point first, since the OVERALL COMMUNICATIVE INTENT of the passenger is to signal an inquiry, not formulate an observation. This intent may become blurred if the SUBJECT/AUXILIARY inversion is missing.

Possible elicitation:

The teacher writes a question mark (?) and a period (.) on the blackboard. She points to one and then to the other, saying with an exaggerated intonation:

Teacher (puzzled):	"You like??" or "You like."
Students:	You like??
Teacher:	Ah! And so?
Students:	Uhmmm... Do you like?
Teacher:	Uh huh, do I like what?

This is an example of how "latent knowledge" is elicited: the students know how interrogatives are constructed and require only a little prodding to recall the form. If such is not the case, however, the teacher's elicitation has to be more direct:

Teacher (puzzled):	"You like??" or "You like."
Students:	You like??
Teacher:	I see. (Writes "... you like" and points to dotted line) A question word... here, please!
Students:	(puzzled silence)
Teacher:	"Do".
Students:	"Do you like?"
Teacher:	Yeah, if you're asking me a general question, say "do"!

The teacher does not explain the general concept of inverting subject/auxiliary word order in interrogatives, nor does she explain the need to use dummy auxiliaries ("do", "did") when there are no others. She simply treats "do" as a "question word" that she seems to want to hear when asked a generic question—at least in a semi-formal situation such as the one described. Systematic reflection will come later. If teaching were truly flexible, moreover, students would have previously had the opportunity to clarify their language-learning goals. If they didn't really care how they sounded and only wanted to speak the foreign idiom in order to "get by" on a vacation trip, the teaching heuristic could conceivably accept all instances of MISSING SUBJECT/AUXILIARY INVERSION as tolerable and move on to the next point (BASE FORM/GERUND ANACOLUTHON). Finally it should be noted that the students do not have to recognize the specific items in the teacher's patter ("I see... Uh huh... Yeah, say 'do'!") to gather that she is musing, agreeing or exhorting. Probably the most serious tactical error of novice teachers (and courseware designers) is to think that everything students read or hear must be immediately understandable to them on the basis of past work (the "building block" learning theory), i.e., that their linguistic knowledge must develop incrementally in a linear fashion.

(6) The teacher compares the students' revised hypothesis with the target utterance and again decides which divergent trait is the most appropriate to deal with.

(7) This procedure is repeated until the students' final utterance conforms with the target utterance, after which the teacher, using the same procedure, presents the next utterance from the source dialog to be reconstructed. The CR process goes more quickly as the dialog proceeds, since students key into the aims of the speakers. It remains, however, a psychologically demanding, time-consuming activity that teachers are wise to limit to 40 min per hourly lesson, one lesson out of five.

COMPUTER-BASED CONVERSATION REBUILDING

The greatest problem in automating CR lies, of course, in the emulation of the teacher's thought processes:

1. determining the nature of the various differences, if any, between the target utterance and the student's utterances;
2. deciding the order in which the differences should be dealt with, given the history of the student–teacher interaction up to that point.

These thought processes are neither trivial nor random. To appreciate their complexity, one has only to consider a possible Artificial Intelligence architecture capable of replicating them. Our working hypothesis is that tutorial advice provided to the student in the form of questions, requests and information can be viewed as "operators" that are intended to help a student reduce the difference between the utterance currently hypothesized and the (fixed or variable) target utterance. Let us examine the case of a fixed target utterance, since that is what we have actually undertaken to implement. We postulate that any differences between the student's utterance and the target utterance must be one of only four possible types: lexical, syntactic, semantic or pragmatic. This last category is, in reality, an ambiguous catch-all: it attempts to account for everything from illocutionary and perlocutionary differences (for example, questions that turn out to be requests) to stylistic and cultural differences such as appearing circumspect while, in reality, being intrusive. Unfortunately, present-day linguistic accounts of English come nowhere near describing the full complexity of a speaker's intentionality and

Weltanschauung.[9] This fact accounts in part for the shortcomings of the AI model we have built. Even though it is based on fixed target utterances, it still is not able to foresee and handle all deviant utterances when "pragmatic" issues are at stake. Hopefully our model, through its shortcomings, will encourage linguists to give greater emphasis to research into speech as *polis* (the expression of a way of life) and not simply as *logos* (the expression of universal ideas)—see Boylan [46,67].

During a session with the Language-Learning Support System, the hypothesis that the student is working on in a given moment may be viewed as the "current state" of a problem solver, while the target utterance may be seen as the "target state". Thus the AI implementation of Conversation Rebuilding requires the definition of algorithms which are able to:

1. compare the student utterance with the target utterance;
2. identify and express the "distance" between the student utterance and the target utterance;
3. associate a "tutorial advice scheme" with each distance type.

A data flow chart showing these technical aspects is given in Fig. 1, where "tutorial advice" represents an operator for the reduction of the difference between the current state and the target. The computational model we have used for the implementation of the prototype is represented in Fig. 2, where the main cycle of a Conversation Rebuilding session is shown. The approach used has similarities with "means–ends analysis", i.e. the search technique used in GPS [49]. In the architecture of Fig. 2, the distance between the student hypothesis and the target utterance is represented in the working memory Δ, constituted by a set of differences δ. Π is a set of production rules establishing a correspondence between a set of possible differences present in Δ and a set of operators Π for the reduction of the corresponding differences. A rule has the following form:

IF the difference Δ_i is in Δ

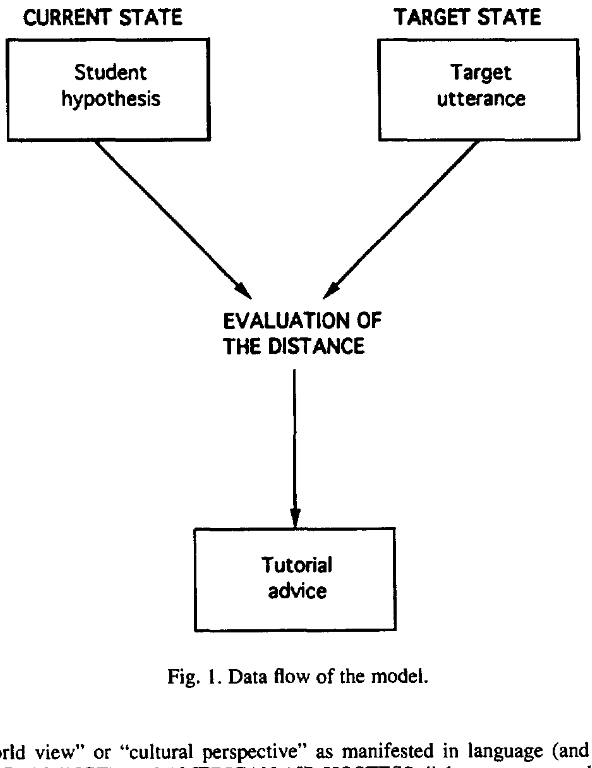

Fig. 1. Data flow of the model.

[9] *Weltanschauung* means "world view" or "cultural perspective" as manifested in language (and other symbolic activity). For example, in the ITALIAN PASSENGER and AMERICAN AIR HOSTESS dialog reconstructed earlier, the passenger's native views on gender and authority will, among other cultural determinants, tend to show up in the English he uses. To appear to the hostess to speak "better" English (i.e., more like hers) the passenger could try to reflect the hostess' views on gender and authority in his language. As we mentioned earlier (note 4), learning to speak a language authentically means learning to "feel" things differently, and that includes relationships such as masculinity–femininity and authority–subalternation.

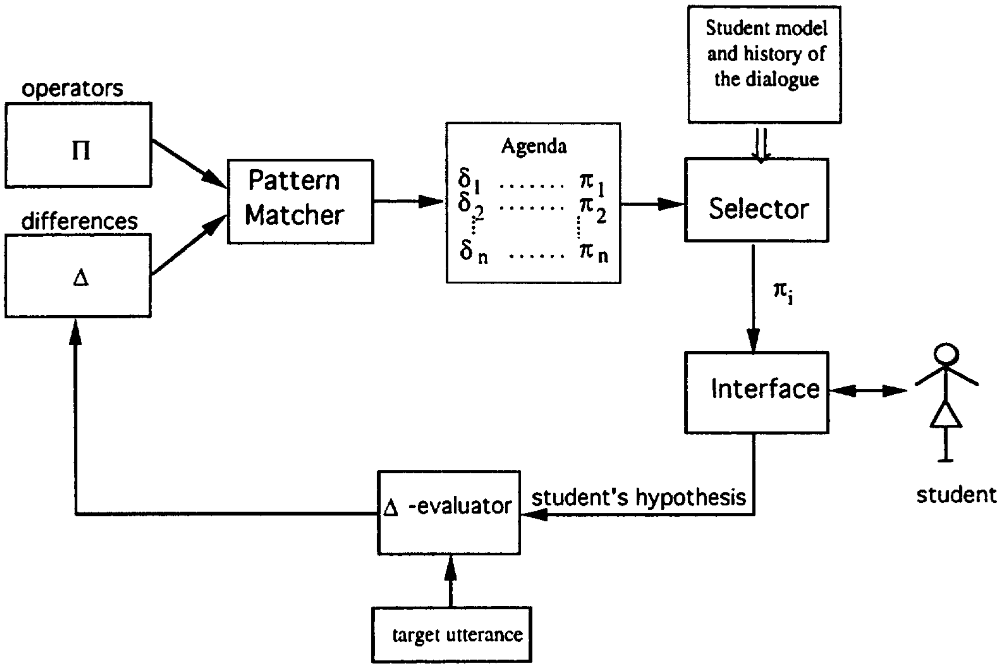

Fig. 2. The main cycle of a Conversation Rebuilding session.

THEN execute operator Π_i.

We are now ready to have a look at the behavior of the proposed architecture. At the beginning of a teaching session, students input their initial hypotheses. A given hypothesis is compared with the target utterance by a module called Δ-evaluator. The differences determined are then placed in data base Δ. At this point the "pattern matcher" is activated. During this phase, all the operators corresponding to the differences present in Δ are singled out and put onto an agenda (also called a conflict set). Those operators (instances of production rules) constitute the teaching actions applicable to a specific situation. In general, more than one teaching action may apply. In the next phase (the "conflict resolution" phase) the Selector chooses the best instantiation of the operators present on the agenda, according to a given heuristic which takes into account the history of previous interactions with the student, i.e. the content of the Student Model database built up by the system during the session. The application of the selected operator produces a message which "prods" the student into revising his/her hypothesis and producing another utterance. The new utterance is compared with the target, and so on. Thus, a Conversation Rebuilding session built on a fixed target utterance can be viewed as a certain number of Evaluation–Match–Select–Execute cycles.

In the following sections, a description of the anatomy of the various modules (Δ-evaluator, Working Memory Δ, Production Memory Π, Pattern Matcher, Selector) that compose the architecture of the system will be presented.

To be fully functional, our computational model should be able to deal with the following situations: (1) the student generates no utterance (= "I don't know"), (2) the student generates an utterance which, while understandable, is "unauthentic" (i.e., "deviant from the target utterance"); (3) the student generates two or more utterances to cover all bases, (4) the student asks a question, (5) the student generates an utterance so much at variance with the target utterance that it cannot be defined in terms of specific differences.

The Evaluation–Match–Select process described above appears robust enough to handle cases (1), (2) and (3), as any mental experiment will readily show. Case (4) could probably be handled by the process described, only so long as the student's question is definable in terms of (2). Case (5) admits no easy solution. It is precisely in order to handle such a case that we have opted for the more restrictive "fixed target" version of CR. We thus count on the students' willingness to accept the rules of the game: they must want to play the character in the film or anecdote to the point of wanting to guess exactly what the character says. Case (5) can therefore be handled with only two kinds of standard answers:

(a) if the student's utterance yields no semantic correlations, it is rejected without comment and a different line of research is suggested.
Situation: Hostess, taking out cigarette, asks passenger something.
Student (Hostess): "Could you?"
[Program finds no correlation between utterance and target ("Have you got a light?").]
Tutor: No, ask me if I HAVE something.
(b) if there are any semantic correlations, the Tutor can motivate rejection of the student's utterance:
Same situation.
Student: "Got a match?"
[Program finds GOT in the SEMANTIC AREA of "HAVE" (+COLLOQUIAL), and "A MATCH" in the AREA of "A LIGHT" (+HYPONYM).]
Tutor: "No, 'got' is too colloquial for your character.
'Match' is too specific."

These semantic correlations are possible in our prototype, of course, only because they have been hard-wired into the system. In other words, our system has not been constructed to handle any dialogs but only the ones for which the appropriate semantic networks have been constructed by hand. This is obviously a major defect from the standpoint of information science—but not from the standpoint of didactics and thus CAI/ICALL. Indeed, a hard-wired Intelligent Tutoring System is no less desirable than an audio-lingual cassette or CD-ROM encyclopedia: all three are designed to be used unaltered by different groups of students. Hard-wiring has permitted us to experiment with a finalized version of computerized CR, something that would have been far out of our reach had we felt obliged to come up with semantic nets capable of handling any dialog input. This shortcut has, in turn, permitted us to gather precious information on what CR can and cannot do for students and how it should be implemented. We might compare our prototype to the ones created in an inflexible "demo language" by software houses before getting down to writing the final program in a full-blown language.

At the present stage of our project, the AI tutor can end the CR session only when the student has generated an utterance equal to the one initially targeted.

EVALUATION OF STUDENT HYPOTHESES

The Δ-evaluator module plays a key role in the architecture. The distance Δ between a hypothesized utterance H and the target utterance T is defined as a set of instances of difference types:

$$\Delta(H,T)=\{\delta_1,\delta_2,\ldots,\delta_n\}.$$

Natural Language Processing (NLP) techniques are used by the Δ-evaluator module to determine the various difference types. During the lexical analysis the system checks if the words that compose the student's sentence are present in the system dictionary. If the search fails, the analyzer determines the list of the unknown words, useful for the generation of the lexical differences. Presently, the system dictionary contains about 3600 lexical items.

Syntactic differences

During the syntactic analysis of the student's input, the grammatical category of each word is identified, a derivation tree is generated (whenever possible), and student "mistakes" are singled out. "Mistakes" are unaccountable variants, definable as the syntactic differences between the student utterance and the target utterance. The parser, based on the chart formalism [50], is constituted by 124 grammar rules, and is described in some detail in Aiello *et al.* [51]. It deals with a wide range of sentences, i.e., declarative, interrogative and imperative.

A Working Memory element δ_i expressing a syntactic difference has four components: the first two represent the type descriptors of the difference; the third and fourth are the attributes of the mistake made by the student and its correction. Here is an example:

(syntax-error *subj-verb* error correction).

This difference is relative to the discrepancy between subject and verb. Here is an instance of that difference:

(syntax-error *subj-verb* (I wants) (I want)).

Semantic differences

During the semantic analysis the ideational consistency between the student's sentence and the content of the target utterance is verified. To this end the system uses both an internal representation of extralinguistic information and a representation of the system expectations which constitute, in substance, the content of the single target sentences. The frame formalism has been used for representing relevant knowledge.

Figure 3 shows an example of a frame—specifically, the predicate "want to send" in the target sentence of the sample CR session transcribed in Section 6. This structure is composed of four elements:

phrase / meanings / translation / syntax.

The phrase identifies the represented object. Meanings consist of a set of slots, i.e., attribute–value pairs, describing various properties of the object. The translation represents the Italian translation of the word(s) representing the object. The syntax lists the syntactic characteristics of the object. Extralinguistic knowledge is represented as a semantic net the nodes of which are frames containing the relevant objects and relationships present in the scene being considered.

The system is also endowed with an explicit representation of the target utterances (i.e., the system expectations). For each target there is a structure representing the peculiar aspects of the sentence, called CLASS. Figure 4 shows an example of CLASS relative to the target "Can I help you?" For the sake of brevity, details of the representation will be omitted here. Suffice it to say that the representation contains information about predicates, agents, objects, temporal forms, possible synonyms (according to context), etc. relative to the target and used by the system for determining possible differences with the student hypothesis. There are various kinds of semantic differences determined by the system and put in Δ when the student utterance type differs from the target utterance type. They are constituted by a type descriptor and a variable number of attributes. Let us see some examples:

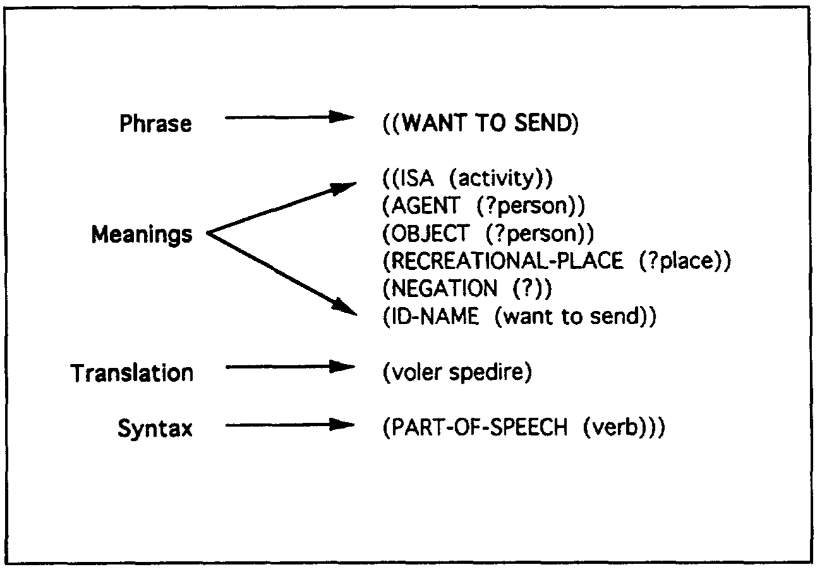

Fig. 3. An example frame.

(wrong-sentence-type

correct-type	target sentence type (affirmative, interrogative, etc.).
wrong-type	student sentence type if divergent.
expected-aux	expected auxiliary missing.
negation	possible negation used by the student.
missing-final-point	no end punctuation given by the student.
wrong-context-words	possible words unsuitable to the context.

)

The following difference is created if the student did not use a necessary predicate:

(necessary-predicates-not-used

list-nec-pred-not-used
nec-pred-used

)

This is an important difference since it shows that the student is presumably on the wrong track. Nec-pred-used is a list of the expected predicates used by the student. List-nec-pred-not-used is an association list: an expected predicate is used as a key and the Italian translation of the predicate is associated with it as a value which can be used to put the student back on track.

Other differences determined by the system: the agent of a predicate is wrong, the tense of a predicate is different from the expected tense, the student used synonyms of expected words, etc. Presently the system is capable of determining 56 difference types. Since even the abridged version of any communicative grammar lists over 800 semantic areas covered by English grammar, it is clear why the

```
                          (((((HELP) ((SHOP-ASSISTANT) (noun))
                           ((CAN + BARE PRESENT INFINITIVE) NIL)
                            ((I) (pronoupers))
                          (OBJECT AUS NEGATION)
                           ((YOU) (pronounpers))
                          ((PETER) (proper-name))

                          (((SHOP-ASSISTANT) NIL (S) NIL NIL)
                           ((PETER) NIL (S) NIL NIL))

                          NIL

                          NIL

Synonyms         ------>  (((HELP) ((want) (need) (wish) (aid) (want to
                          help) (wish to help) (want to aid) (wish to aid))))

Pronouns-guide   ------>  (((SHOP-ASSISTANT) ((person-reference1)
                          (maperson-reference)))
                          ((PETER) ((person-reference2)))
                          ((MOTHER) ((feperson-reference))))

                          1

Target question  ------>  (CAN I HELP YOU _QUESTIONMARK_)
                          (COSA DICE IL COMMESSO A PETER?)

                           (((CAN) (verbo)) ((I) (pronome personale))
                           ((HELP) (verbo)) ((YOU) (pronome personale))
                           ((_QUESTIONMARK_),(punto interrogativo)))

Sentence type    ------>  ((INTERROGATIVA) NIL (CAN) (_QUESTIONMARK_)))
```

Fig. 4. CLASS relative to the target "Can I help you?"

system insists on only one acceptable target utterance and why the dialogs are carefully chosen to limit possible student hypotheses as much as possible. Nonetheless, once the student accepts to play the game the system proposes—which is to second guess what a film character is about to say—the illusion of creativity is maintained, since the student forgets that he is working within a very limited semantic–syntactic area.

PRODUCTION MEMORY, PATTERN MATCHER AND SELECTOR

The Production Memory of the computational model of Fig. 2 is made up of a set of production rules, implementing the teaching operators that are intended to reduce the distance between the student sentence and the target. In our implementation the production rules assume the following form:

(*rule-name* (**if** pattern ⟨pattern⟩)
(**then** (lambda-variable ⟨lambda-variable⟩) (expression ⟨expression⟩)))

The condition part of the rule is a pattern which describes the data configuration for which the rule is appropriate. The action part is relative to the corresponding teaching action capable (hopefully) of reducing the difference between the student's proposal and the target. Figure 5 gives an example rule: the condition part is constituted by only one condition element. The selection of the best teaching action to apply is driven by heuristics that capture practiced teachers' expertise. To this end, expert language tutors were interviewed and observed as they taught English to native Italian speakers. For each student that uses the system, a database is maintained which keeps information on which syntactic differences have been determined. We call such a database the Student Model. The Π-Selector is also driven by the content of the Student Model during the conflict resolution phase. In general the heuristics used in this phase can be summarized as follows:

```
(syntax.undet
 (if     syntax-error undet (? error) (? correction)))
 (then  (error correction)
        ((let ((s " "))
        (setf *standard-output* messages)
        (ask messages
           (window-select))
        (let   ((str-corr (_words->string_ correction))
           (str-err (_words->string_ error)))
        (dotimes (v 1 1)
           (terpri))
        (format t "~%Non va bene l'uso dell'articolo indeterminativo)
        (format t "~%Non si scrive <<~a>> ma <<~a>>"
                str-err  str-corr)
        (format t "~%Riscrivi la frase e usa <<~a>>"
                str-corr))
        (setf *standard-output* sent)
        (ask sent
          (window-select))
        (terpri)
        (terpri)
        (print 'Input-> )
        (setf s (read-line))
        (setf s (_string-not-nil_s))
        (_Cycle_s *TARGET* *CLASS*))
)))
```

Fig. 5. An example rule.

- Lexical Δ have higher priority than syntactic and semantic Δ.
- Semantic Δ have higher priority than syntactic Δ.
- The specificity principle [52] is applied. This principle states that if one instantiation rule matches the same working memory elements as another instantiation rule but contains additional condition elements, it dominates because the corresponding teaching action is more specific.

These criteria may be modified according to the contents of the Student Model. For instance, the system increases the priority level of an operator if it concerns a difference already present in the Student Model.

SAMPLE INTERACTIONS

Our prototype enables students to rebuild a selected number of complete authentic conversations, based on anecdotes. Two translated (and thus partially adapted) excerpts of Conversation Rebuilding sessions follow: the System's messages, in fact, are in Italian to ensure understanding but appear here in English. They consist of pre-constituted sentences with "slots" filled in with specific words at run time. The first dialog (Version I) was obtained with the fully functioning prototype based on the earlier, overly grammatical approach which focuses the student's attention on his utterances; in addition, the student played all roles. The second (Version II) is from the revised program which we are still working on and represents the kind of dialog it can produce: here the student plays only one role (to keep in character) while the tutor focuses attention on the meaning that the student communicates and the effect that the student makes. In trial runs this technique proved sufficient for signalling to the student the remedial work she/he must undertake to make her/his utterance understood as intended.

The flower shop anecdote

Context: A young Londoner, Peter, is thinking of his mother on vacation in Italy. It is her birthday. He sees a flower shop that is just opening and goes inside.
Intent: Clerk: assist client (sell flowers). Peter: Send mother flowers (buy flowers).
First two target utterances to reconstruct:
Clerk: Can I help you?
Peter: I want to send some flowers to my mother in Italy.

Interaction—Version I

System: What does the clerk say to Peter?
Student: Good morning.
System: Please pay attention.
They are talking about *help*.
Try again using *help*.
Student: May I help you?
System: The tense of the predicate *help* is correct.
Try with a synonym of *may*.
Student: Can I help you?
System: That's right.
Now, what does Peter answer the clerk?
Student: I wants to buy any flowers.
System: Your proposal seems *negative* while I was expecting an *affirmative* sentence.
Try to use an affirmative form.
Student: I wants to buy some flowers.
System: There is a discordancy between subject and verb.
Student: I want to buy some flowers.
System: Try to use the word *mother*.
Student: I want to buy some flowers for my mother.
System: *Want to buy* is "similar" to the expression I was expecting.
Student: I want to send some flowers for my mother.
System: *For my* is not the appropriate *preposition* and *possessive*.
Student: I want to send some flowers to my mother.

System: ...They are talking about a touristic place.
Have a look at the extralinguistic information.
Student: I want to send some flowers to my mother in Italy.
System: That's right.

Interaction—Version II

Tutor: Now you're Peter and I'm the clerk. Good morning.
Student: Good morning.
Tutor: Can I help you?
Student: I wants to buy any flowers.
Tutor: "Any?" Is this a question?
Student: Some flowers.
Tutor: What do you mean? Say something more.
Student: Some flowers, please.
Tutor: Who are they for?
Student: for my mother.
Tutor: Peter says "send".
Student: I wants to send some flowers for my mother.
Tutor: "For?"
Student: to my mother.
Tutor: "I wants" sounds terrible.
Student: I like to sound terrible.
Tutor: Could you ask your question again? And tell me where your mother is.
Student: I wants to send some flowers to my mother in Italy.
Tutor: Ok—here's what Peter says in the conversation.

THE LITERATURE AND RELATED WORK

One of the earliest attempts in applying Artificial Intelligence techniques to foreign language learning is the work of Weischedel *et al.* [53]. Their system is capable of detecting syntactic and agreement errors in student utterances, even if they do not provide students with explanations of why the forms singled out are to be considered mistakes. Our ITS, on the other hand, specifically addresses this question—if anything, it has too much of a metalinguistic bent, especially in the first versions which used grammatical terminology as prompts ("Is your verb singular or plural?").

Imlah and du Boulay have tackled the problem of robust parsing for educational purposes [54]. Their system, named FROG (French RObust Grammar checker) and implemented in PROLOG, checks sentences typed in by students for a limited number of grammatical errors in French; it can only handle declarative sentences, however, whereas our system provides for all major sentence types. Barchan *et al.* [55] have taken FROG a step further with their FGA (French Grammar Analyzer) system; it adds to the FROG parser an explicit taxonomy of common misconceptions among students, something our system might admittedly profit from having.

Other AI language teaching systems have grown out of a critical appraisal of FGA. One is the Definite Clause Grammar (DCG) parsing system [56] which has, however, proved unsatisfactory. Yazdani [57,58] and O'Brien [59] have tried to get a DCG to perform like a chart parser in their latest versions of an English tutor. The prototype of an Authoring Environment capable of providing the student with correction and explanation of grammatical errors for the Dutch language has also been developed [60]. Schwind [61] has presented an eclectic tutoring system for German, implemented in PROLOG II, which makes no explicit claim to any particular theory of language or language learning (ours, based on the "communicative approach" as defined by Brumfit and Johnson, is described in the first section) or student modeling. Schwind asserts that his system is instead based on an "objective" knowledge base about German, sufficiently complete to permit analyzing sentences proposed by students, detecting and explaining errors, and even answering students' queries (something we feel we would be able to implement only in the restricted form mentioned at the end of the third section).

A common characteristic of all the above mentioned systems is that they deal with isolated sentences only. None of them has been used as the main module for a more sophisticated language learning support system based on purposeful conversation—but that is what, by all accounts, language is supposed to be for. Instead, they encourage students to manipulate words gratuitously for the sole purpose of checking or reinforcing rote knowledge of some grammatical rule—something that we feel can be obtained through purposeful conversation as a by-product.

One system that does attempt to deal with discourse is described in Yamamoto *et al.* [62]. Their conversation simulation is "based on a goal-oriented discourse model", an excellent premise. But the authors go on to explain that their system "advises the student of errors" without specifying whether appropriateness errors are taken into consideration. As Hymes pointed out long ago, "without rules of use, the rules of grammar would be meaningless" [63]. Our work on a Support System using Conversation Rebuilding, on the other hand, has been an attempt to relate "grammar" to specific instances of language use. Students must get an interlocutor (the "Tutor") to understand their intended messages which they frame to handle a specific setting with specific sociocultural connotations, within the context of the history of the interaction up to that point.

Some authors feel that it is simply not worth the effort to create an ITS dealing with discourse because it will never get used anyway, given the didactics practiced in most foreign language classrooms. As we mentioned in Note 2, Swann [11] confesses that, although he is convinced that "descriptive grammar and drill and practice exercises may be neither necessary nor sufficient for foreign language learning", he implemented them in the ICALL software development project he conducted in an Italian secondary school because otherwise the courseware would not have been used (p. 252). To his surprise he then discovered, after creating and installing the system, its "success and popularity with at least some types of students" (p. 252). The question is, of course, what "type of student"? Swann goes on to note that the teachers assigned to work with him in developing and testing an ICALL program, were traditionalists convinced that the only way to learn a foreign language is by grammatical analysis followed by drill and practice. It is therefore not hard to imagine that these highly defensive teachers had transmitted their highly defensive attitudes towards language learning to their students, in a process akin to the self-fulfilling prophecies of failure that some teachers unconsciously make when they single-out and brand certain students as "slackers". The lack of success that Swann's teachers had previously encountered when they had attempted to use "communicative" methods in the classroom (p. 263) could have very well been a key moment in transmitting to their students a distrust of the instinctual processes underlying "communicative" learning and a need for the kind of defensive, rationalized hold on language, or rather on the formal aspects of language, that grammatical analysis and drill and practice undoubtedly give.

What could Swann have done to gain acceptance in the school and yet implement the kind of didactics he felt would have served the students best? A practical suggestion may be found in Boylan [67]. Suffice it to say here that Swann's dilemma makes it clear that claims for the "success" of this or that Computer-Assisted Language Learning (CALL) program should be critically appraised on the basis of a careful analysis of just what the program sets out to teach. Not all that glitters is gold. As Garrett concludes in a recent paper: "Perhaps most important, all the studies of the effectiveness of CALL assess effectiveness in terms of short-term mastery of discrete-point, disconnected, and usually relatively trivial items of language, precisely the kind of testing that modern language pedagogy tries to avoid" [64].

LESSONS LEARNED AND FUTURE PLANS

The ICALL System described here takes inspiration from the "communicative learning" technique called Conversation Rebuilding practiced by some of the more innovative teachers. Admittedly, however, it implements the more restrictive "fixed response" variety. Thus its principal merit lies mainly in indicating the direction ICALL research can and (in our opinion) should go: getting students to produce *intentional situated interactive discourse.*

Yet we have also learned that a system such as ours, with all its shortcomings, can effectively promote learning. Indeed, the satisfaction students claim to get from using our program shows that, to be effective, a Language Learning Support System does not need to BE intelligent; it only needs to get the student to interact intelligently with it. Students consistently report the feeling of having created utterances that, in point of fact, have been rigidly programmed into the system; moreover, their logs (notepads and pens left next to the computer keyboard) show that their various trials-on-paper, before committing their responses to screen, are effectively a search for rendering in words—foreign words—something they want to say. It is certainly not just applying grammatical rules diligently, to honor some abstract principle—i.e., the kind of activity most often practiced with ICALL software or, in the classroom, with traditional textbooks.

Clearly the role of the Tutor remains preponderant in the man–machine interaction we have programmed and is, quite often, overbearing. We therefore feel that future development of our Language Learning Support System should go in the direction of devising an environment that permits students to take the initiative in piloting the intelligent activity: a system that, through the constraints programmed into it, lets students "freely" decide the production of discourse in a foreign language in ways that can only turn out to be authentic. This is, as we have seen, the inspiration behind Conversation Rebuilding [65]. But instead of building a Tutor that oversees the entire interaction, we are now thinking along the lines of a hypermedia environment which lets the student do the navigating (see Papagni *et al.* [66]).

CONCLUSIONS

In this paper a sophisticated classroom teaching procedure, called Conversation Rebuilding, has been described as a possible means of helping students—led by a computer-based Support System using AI technology—to learn to produce intentional utterances in a foreign language. Imagining that they are a character in a film or anecdote, students try to second guess what their character says line by line as the story develops. They type the utterances they hypothesize onto a computer screen; the system, acting as a Tutor, challenges their utterances if they diverge from the original. In particular, the system focuses on the lexical, syntactic, semantic and pragmatic aspects that are the most misleading, conceptually divergent, anacoluthic or culturally unauthentic (in that order). This forces the students to alter their utterances item by item until they sound like something a native speaker might say; indeed, the Tutor then lets the students read or hear the original conversation in which native speakers do in fact say the exact same sentences that the students have been led to reconstruct.

The Tutor's behavior has been designed using a means–ends analysis technique realized with a Production System model of computation: differences between a currently hypothesized utterance and a target utterance are gradually reduced as the Tutor makes the student aware of the differences step-by-step. The system shows both syntactic flexibility in handling non-grammatical student utterances and semantic resilience in resolving ambiguities through domain knowledge.

The development of the system has presented several computational and pedagogical challenges. The final goal of our research endeavor is, of course, to devise an adaptable system which can be reused to handle any dialog (only the knowledge base would have to be rebuilt). As a first step, however, we chose to test the validity of the CR paradigm and, at the same time, lighten the computational task by constructing a prototype which has been "custom made" for the Flower Shop dialog. And in fact the final product has shown to be didactically quite useful as a "one-shot" instructional aid, something like a textbook or video cassette which is employed as such, and only once, by each successive levy of students.

Pedagogically, we had no pretention to create a system which could rival with Conversation Rebuilding as performed in the classroom by a gifted teacher and highly motivated fellow students. There CR is a very fast-paced, dynamic, oral/visual type of creative group activity that leaves students asking for more; on a computer it can at times be simply extenuating. Indeed, test students regularly report that from classroom to computer, Conversation Rebuilding loses somewhat in appeal: a group activity becomes a solitary activity (although this permits self-pacing); a dynamic oral activity becomes a slower-paced typing activity (although this enhances the student's awareness of spelling and written syntax); and oral/visual presentation of extralinguistic information by speaking/pantomiming becomes perusal of explanatory texts (although this improves reading skills). Despite the lesser subjective appeal, however, students add that they still enjoy learning a foreign language with the Support System—if less so than in teacher-led Conversation Rebuilding in the classroom, much more so than by doing the discrete-point exercises typical of most CALL programs or language lab drills.

We have yet to make comparative tests of how much and what kind of language learning actually takes place with each method: the task is daunting since different authors' methods set out to teach different things, often to different publics. CR in the classroom tries to develop spontaneous reflexes and an intuitive grasp of successful communicative strategy in the foreign idiom. Discrete point ICALL programs try to develop a conscious reflection on the grammatical system of the foreign idiom. Our CR Language Support System is a middle road: students focus on the overall communicative intent of the characters they play, but as they recreate their dialog, the system forces them to reflect on (or even teaches them globally by example) particular applications of the lexical, grammatical and pragmatic regularities typical of the foreign idiom.

Acknowledgements—The authors acknowledge Luigia Aiello for her inspiration and constant dedication which made this project possible. We also wish to thank Marta Cialdea, Fabrizio Cianfanelli, Alessandro De Sanctis, James Spohrer and, in particular, Christopher Humphris, who have contributed to various stages of the realization of this work.

REFERENCES

1. Van Ek, J. A., The threshold level in a European unit/credit system for modern language learning by adults. In *Systems Development in Adult Language Learning*. Council of Europe, Strasbourg, 1975.
2. Brumfit, C. J. and Johnson, K. (eds.), *The Communicative Approach to Language Teaching*. Oxford University Press, Oxford, 1979.
3. Corder, S. P., *Error Analysis and Interlanguage*. Oxford University Press, Oxford, 1981.
4. Littlewood, W., *Communicative Language Teaching: An Introduction*. Cambridge University Press, Cambridge, 1981.
5. Johnson, K., *Communicative Syllabus Design and Methodology*. Pergamon Press, Oxford, 1982.
6. Oller, J. W. and Richard-Amato, P. A. (eds.), *Methods That Work*. Newbury House, Rowley, Mass., 1983.
7. Ellis, R., *Understanding Second Language Acquisition*. Oxford University Press, Oxford, 1985.
8. Schegloff, E. A. and Sacks, H., Opening up closings. *Semiotica*, 1973, **8**(4), 289–327.
9. Jefferson, G., Error correction as an interactional resource. *Language in Society*, 1974, **3**(2), 181–199.
10. Boylan, P., La lingua straniera. In *L'alfabetizzazione culturale e comunicativa*, eds. T. De Mauro *et al.* Firenze, Giunti Marzocco, 1992, pp. 71–157.
11. Swann, P., Computer assisted language learning for English as a foreign language. *Computers & Education*, 1992, **19**(3), 251–266.
12. Richards, J. C. and Schmidt, R. W., *Language and Communication*. Longman, London, 1983.
13. Titone, R., What is second language learning? In *id., Second Language Learning and Psychological Studies*. Milan, ISFAP Editore, 1987.
14. Krashen, S. D., *The Input Hypothesis*. Pergamon Press, Oxford, 1984.
15. Sercu, L. (ed.), *Intercultural Competence. A New Challenge for Language Teachers and Trainers in Europe*. Aalborg, Centre for Languages and Intercultural Studies. Aalborg University, 1995.
16. Urbani, S., La Ricostruzione di Conversazione. In *Bollettino Dilit*, 3, 1, Edizioni Dilit, Rome, 1981.
17. Micarelli, L., Oltre le parole: un approccio relazionale alla ricostruzione di conversazione autentica. In *Bollettino Dilit*, **2**(3), Edizioni Dilit, Rome, 1989.
18. Humphris, C. (ed.), Atti del 2 Seminario per insegnanti di lingua. In *Bollettino Dilit*, 3, Edizioni Dilit, Rome, 1990.
19. Oller, J. W., *Language Tests at School*. Longman, London, 1979.
20. Council of Europe, *The Threshold Level for Modern Language Learning in Schools*, ed. J. A. van Ek. Longman, London, 1976.
21. Department of Education and Science, Modern Foreign Languages Working Group: Initial Advice, ed. M. Harris, London, 1990.
22. Micarelli, A. and Boylan, P., Foreign language tutoring systems today: old-fashioned teaching with newfangled gadgets. *Cognitive Systems*, **5**(1), (in press).
23. Swartz, M. and Yazdani, L. M. (eds.), *Intelligent Tutoring Systems for Foreign Language learning. The Bridge to International Communication*. Springer-Verlag, Berlin, 1992.
24. MacWhinney, B., The competition model and foreign language acquisition. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani (eds.). Springer-Verlag, Berlin, 1992, pp. 39–50.
25. Miller, G. A. and Feldbaum, C., WordNet and the organization of lexical memory. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 89–101.
26. Swartz, M. L., Issues for tutoring knowledge in foreign language intelligent tutoring systems. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 219–247.
27. Handke, J., WIZDOM: a multiple-purpose language tutoring system based on AI techniques. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 293–303.
28. Psotka, J., Holland, M. and Kerst, S., The technological promise of second language intelligent tutoring systems in the 21st century. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 321–335.
29. Demaiziere, F. and Blanvillain, O., Learners' Intuitions, misconceptions and errors. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 25–36
30. Fum, D., Pani, B. and Tasso, C., Naive vs. formal grammars: a case for integration in the design of a foreign language tutor. In *Intelligent Tutoring Systems*, eds. M. Swartz, and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 51–64.
31. Abeille, A., A lexicalized tree adjoining grammar for french and its relevance to language teaching. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 65–87.
32. Frederiksen, C., Donin, J., Decary, M. and Hoover M., Semantic discourse processing and tutoring systems for second language learning. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 103–121.
33. Chanier, T., Pengelly, M., Twidale, M. and Self, J., Conceptual modelling in error analysis in computer-assisted language learning systems. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 125–150.
34. Tasso, C., Fum, D. and Giangrandi, P., The use of explanation-based learning for modelling student behavior in foreign language tutoring. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 151–170.
35. Cerri, S., Cheli, E. and McIntyre, A., Nobile: Object-based user model acquisition for second language learning. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 171–190.
36. Kempen, G., Language technology and language instruction: computational diagnosis of word-level errors. In *Intelligent*

Tutoring Systems, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 191–200.
37. Güvenir, H. A., Drill and practice for turkish grammar. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 275–291.
38. Zock, M., SWIM or sink: the problem of communicating thought. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 235–247.
39. Legenhausen, L. and Wolff, D., STORYBOARD and communicative language learning: results of the duesseldorf CALL project. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 9–23.
40. Hamburger, H. and Hashim, R., A foreign language tutoring and learning environment. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 201–218.
41. Sussex, R., Invoking and exploring HELP in a decoupled task-discussion level environment for second language learning. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 249–259.
42. Wilks, Y. and Farwell, D., Building an intelligent second language tutoring system from whatever bits you happen to have lying around. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 263–273.
43. Criswell, E., Byrnes, H. and Pfister, G., Intelligent automated strategies of teaching foreign languages in context. In *Intelligent Tutoring Systems*, eds. M. Swartz and M. Yazdani. Springer-Verlag, Berlin, 1992, pp. 307–319.
44. Collis, B., Moonen, J. and Stanchev, I. (guest eds.), Special issue, Exploring the nature of research in computer-related applications in education. *Computers & Education*, **21**(1–2), 1993.
45. Thompson, J. and Chesters, G. (guest eds.), Special issue, Emancipation through learning technology. Selected papers from the EUROCALL '93 conference. *Computers & Education*, **21**(1–2), 1994.
46. Boylan, P., Logos e polis, Tuesday Conferences. Center for Semiotics, Palazzo Caramanico al Chiatamone, Naples, 1995, mimeograph.
47. Aiello, L., Cialdea, M., Humphris, C., Micarelli, A. and Spohrer, J. C., Helping students refine their working hypotheses: a first pass in the domain of language tutoring. In *Advanced Research on Computers in Education*, eds. R. Lewis and S. Otsuki. Elsevier Science, Amsterdam, 1991, pp. 227–235.
48. Aiello, L. and Micarelli, A., A system for foreign language learning based on conversation rebuilding. In *Proceedings of the World Conference on Artificial Intelligence in Education AI-ED 93*. Edinburgh, Scotland, 1993, pp. 314–321.
49. Newell, A. and Simon, H. A., GPS, a program that simulates human thought. In *Computers and Thought*, eds. E. A. Feigenbaum and J. Feldman. McGraw-Hill, New York, 1963, pp. 279–293.
50. Satta, G. and Stock, O., Bidirectional context-free grammar parsing for natural language processing. *Artificial Intelligence*, 1994, **69**(1/2), 123–164.
51. Aiello, L., De Sanctis, A. and Micarelli, A., Computer assisted language learning: a grammar detector and corrector. In *Proceedings of the Seventh International PEG Conference*. Edinburgh, Scotland, 1993, pp. 19–28.
52. Charniak, E., Riesbeck, C., Mcdermott, D. and Meehan, J., *Artificial Intelligence Programming*. Lawrence Erlbaum, 1987.
53. Weischedel, R. M., Voge, W. M. and James, M., An artificial intelligence approach to language instruction. *Artificial Intelligence*, 1978, **10**, 225–240.
54. Imlah, W. G. and Du Boulay, J. B. H., Robust natural language parsing in computer-assisted language instruction. *System*, 1985, **13**(2), 137–147.
55. Barchan, J., Woodmansee, B. J. and Yazdani, M., A PROLOG-based tool for French grammar analysis. *Instructional Science*, 1986, **14**, 21–48.
56. Pereira, F. C. and Warren, D. H. D., Definite clause grammars for language analysis—a survey of the formalism and a comparison with augmented transition networks. *Artificial Intelligence*, 1980, **13**(3), 231–278.
57. Yazdani, M., An English tutor. Project report (1987–89), University of Exeter, 1990.
58. Yazdani, M., The linger project—an artificial intelligence approach to second language tutoring. *CALL*, 1991, **4**(2), 107–116.
59. O'Brien, P., eL: using AI in CALL. In *Multilingual Multimedia*, ed. M. Yazdani. Intellect, Oxford, 1993, pp. 85–139.
60. Pijls, F., Daelemans, W. and Kempen, G., Artificial intelligence tools for grammar and spelling instruction. *Instructional Science*, 1987, **16**, 319–336.
61. Schwind, C. B., An intelligent language tutoring system. *International Journal of Man–Machine Studies*, 1990, **33**, 557–579.
62. Yamamoto, H., Kai, K., Osato, M., Shiino, T. and Inui, M., A structure of an intelligent CAI system for training conversation of a foreign language based on conversation simulation. In *Advanced Research on Computers in Education*, eds. R. Lewis and S. Otsuki. Elsevier Science, Amsterdam, 1991, pp. 249–254.
63. Hymes, D., On Communicative competence, In *Sociolinguistics: Selected Readings*, eds. J. B. Pride and J. Holmes. Penguin Books, Harmondsworth, 1972.
64. Garrett, N., Language pedagogy and effective technology use. *Applied Language Learning*, 1991, **2**, 1–14.
65. Micarelli, A. and Humphris, C., An artificial intelligence approach to natural language tutoring. *Cognitive Systems*, 1991, **3**(1), 69–78.
66. Papagni, M. Cirillo, V., Micarelli, A. and Boylan, P., Teaching through case-based reasoning: an ITS engine applied to business communication. In *Proc. of the Eight World Conference of Artificial Intelligence in Education AI-ED 97*. Kobe, Japan, August 19–22 (in press).
67. Boylan, P., What does it mean to learn a language in today's world; what role can present-day computer technology play? In *Proceedings of the Symposium on Language and Technology*. Editrice CUSL, Florence, 1995, pp. 92–114.