音声対話を実現した英会話用知的 CAI システムの構成

山本秀樹† 田川忠道† 宮崎敏彦†

近年，英語を母国語とする人と実際に会話をしているような環境の提供を目的とした環境型の CAI システムが開発されている．しかし，これらのシステムは英会話の教育において重要な音声による会話文の入力機能を持たないため，十分な臨場感を学習者に提供できていない．またこれらのシステムのいくつかは会話の目的や状況を理解していない学習者を対象としていない．そこで，本論文では，音声による会話文の入力機能とビデオを対話的に操作する機能の両方を持つことにより，会話の臨場感を高めること，および会話の目的や状況を理解していない学習者をも対象とすることを可能にした英会話用知的 CAI システムについて述べる．まず，対話的に操作可能なビデオと音声による対話機能により，目標のある会話を行う際に必要な英会話能力を段階的に習得可能な指導方法を提案する．次に，音声による会話のシミュレーションを実現するために，音声認識システムをどのようにシステムに統合したかについて述べる．さらに教材の作成が容易な会話の知識表現を提案する．本システムは，ワークステーション上に実装されており十分な速度で音声による会話を可能としている．本論文で述べた方式は他の音声対話システムにも応用可能である．

ICAI System for Conversational English Based on Spoken Discourse Simulation

Hideki Yamamoto,† Tadamichi Tagawa† and Toshihiko Miyazaki†

Computer assisted instruction (CAI) systems have attracted a lot of attention as a suitable method for language learning. A number of such CAI systems have already been proposed and implemented. However, nevertheless of its importance, there seems to be no voice interactive intelligent CAI system. The purpose of this paper is to compare some of the techniques used to implement language learning CAI systems and to propose a more extensive intelligent CAI system that we integrate those techniques and speech recognition technique into. We introduce a teaching method using conversation simulation. This method enables to acquire communicative competence for anyone, who may not know typical sentences of conversations, gradually. Moreover, on the current speech recognition technique, to fill the lack of the recognition reliability, special care has been taken so that the system respond appropriately as a tutor. The method we propose may be applicable to the other domain of spoken language discource applications.

1. はじめに

近年，英語を母国語とする人と実際に会話をしているような環境の提供を目的とした環境型の CAI システムが開発されている[1)~4)]．文献1),2)は，会話の場面が録画されたビデオを，計算機を介して対話的に操作する IVD (interactive videodisc)[5),6)]により，学習者に疑似的な英会話の環境を提供している．学習者は，システムを対話的に操作して，ビデオ中の発話の内容やヒントを参照したり，次の発話にふさわしい表現を選択肢の中から選択したりすることを通して学習を進める．システムに対する学習者の入力手段は，ポインティングデバイスによるメニューの選択や英単語のキーボード入力といったように制限されているため，システム側が指導のために必要とする情報を十分に確保できない．例えば，システムは学習者が英会話の表現を正しく生成し発声できているかどうかを正確に把握できない．これらのシステムは，英会話のための基本的な会話を理解していない初心者に複数の会話文を例示し選択させることで会話を理解させることはできるが，すでに基本的な会話文を理解している学習者に，さらに応用力をつけさせるための指導を行うことは難しい．

一方，文献3),4)のシステムは，キーボードを用いたある程度自由な会話文の入力を受け付けることで，ある場面における会話をシミュレートする．これらのシステムは，入力された会話文を構文解析・意味解析することによって学習者の入力に含まれる誤りを発見し，その誤りに対して細やかな指導を行う．そのた

† 沖電気工業株式会社研究開発本部関西総合研究所
Kansai Laboratory, R. & D. Group, Oki Electric Industry Co., Ltd.

め，基本的な会話文を理解しているが応用力は十分でないレベルの学習者に対しても，適切な指導をすることができる．しかしながら，外国語会話の学習にとって重要である音声による対話の機能がないため，音声を使ったコミュニケーションに必要な聞き取り能力の訓練と発話能力の訓練に十分とはいえない．また，これらのシステムは，学習者が会話をしようとしている場面の状況を十分に理解した上で会話文を入力することを暗に仮定しているので，文献1),2)が対象としているようなレベルの学習者，すなわち会話の状況を理解できていない学習者や，その場面における典型的な会話を習得できていない学習者に適さない．

一般的な学習者の能力を考えてみると，一つの教材の中でもある場面についてはそこで用いるべき典型的な表現を理解していない場合や，また別の場面については典型的な表現は理解しているが応用力は十分でない場合がある．したがって，高度な個別教育機能を持ったシステムを実現するためには，会話の中で用いる典型的な表現を教える機能と応用を養成する機能の両方をうまく備えることが重要である．

さらに，音声を用いたコミュニケーションに必要な発話能力の訓練には，音声認識機能を備え音声入力を受け付けることが必要である．音声による会話シミュレーション機能を実現するためには，不特定の話者が連続的に任意の文を発話した場合でも完全に正しく認識できるような音声認識システムが望まれる．しかしながら，現状の技術では，このような音声認識システムを実現できていない[7]．そこで，既存の音声認識システムを英会話用知的 CAI システムにいかに統合するかが重要な課題になる．

本論文では，典型的な英会話文の習得および応用力の養成を目的とした英会話用知的 CAI システムの機能と実現方式について述べる．システムは，IVD と音声認識技術をシステムに統合することによって，それぞれの場面で用いられる典型的な表現を教える機能と応用力を養成する機能を備え，さらに音声による会話機能を有している．

本論文の構成は以下のとおりである．第2章ではシステムの指導方法について，第3章ではシステムの構成について述べる．第4章では英会話用知的 CAI システムに音声認識システムを統合する方法について述べる．第5章では知識表現について，第6章では誤りの同定と誤りに対する指導について述べる．第7章では実験的に開発した教材によるシステムの実行例を説明する．第8章では試作したシステムに関する考察を行う．

2. 英会話の知識の習得過程に基づいた指導方法

本システムの教育目標は，学習者が目標指向の会話モデルに基づいた会話を英語を使って行えるようになることである[4]．目標指向の会話モデルは，「会話には目標がありそれを達成するために会話者は情報を交換する」というモデルである．学習者が目標指向の会話モデルに基づいた会話を行えるようになるためには，単語，熟語といった英語の基本事項や会話特有の表現の理解，リスニング力の向上，および発音の正確さの向上といったことだけではなく，現在進行している会話の状況を理解したり，会話の流れを理解したり，英語圏の文化を理解したりすることが必要である☆．会話に必要なこれらの個々の要素とそれらを計算機を使って指導する方法を**表1**にまとめる．

従来の英会話の環境を学習者に提供しようとする英会話用 CAI システムは，表1に示した指導方法の一部を個別に実現している．例えば，文献1),2)のシステムは，IVD を用いているため，会話の状況や会話の流れをまだ理解していない学習者には適しているが，音声による会話文の入力を受けつけないので，学習者が英語表現を正しく生成し発声できているかどうかをシステムは正確に把握できない．また，これらのシステムは，英会話のための基本的な会話文を理解していない初心者に複数の会話文を例示し選択させることで会話文を理解させることはできるが，すでに基本的な会話文を理解している学習者に，会話の流れに即して発話するといった応用力をつけさせるための指導を行うことは難しい．

一方，文献3),4)のシステムは，自由な会話文の入力が可能であるが音声で入力できないという問題がある．また，すでにある程度の会話の流れなどを理解して会話文を入力できるレベルの学習者を対象にしているため，会話の状況を理解できていない学習者や，その場面における典型的な会話の表現を習得できていない学習者に適さない．

英会話用 CAI システムを用いて学習しようとする

☆ ここで，会話の状況とは，例えば，電話の会話の場合で，「相手が不在の場合」，「電話中の場合」など会話の参加者がおかれているある一つの世界の有り様をいう．また会話の流れは，ある状況のもとでの発話と発話の続き具合，すなわち会話の文脈をいう．

表 1 英会話教育に必要な要素
Table 1 Teaching items for conversational English.

番号	教育要素	教育内容	システムによる指導方法
(1)	英語の基本事項の理解	単語，熟語，構文，文法を理解すること	伝統的 CAI による演習問題
(2)	会話の状況の理解	電話の会話の場合で「相手が不在の場合」,「電話中の場合」など，おかれている状況を理解すること	状況ごとの会話の見本のビデオ，各状況ごとの会話のシミュレーション
(3)	会話の流れの理解	ある会話の状況でどのように会話を進めていけば良いかを理解すること	状況ごとの会話のビデオ，状況ごとの会話のシミュレーション，状況をあらかじめ示さない会話のシミュレーション
(4)	会話特有の表現の理解	例えば，電話で自己紹介をするときは，“This is 名前”のようにいうことなど	会話特有の表現の指導，演習問題，その表現を使う状況での会話のシミュレーション
(5)	リスニング	会話で起きる音の連結（リエゾン） 強弱や抑揚の違いによる意味の違いを理解すること	ネイティブの音声を再生する機能，学習者の音声を録音再生する機能
(6)	発音	個々の母音，子音を正しく発音できること 正しい抑揚，強弱をつけること	音声による入力とそれに対する発音指導
(7)	英語圏文化の指導	英語圏の歴史・文化を理解すること	伝統的 CAI による英語圏の歴史文化の指導

学習者は，一つの会話の中のある瞬間，すなわち会話のある場面*についてはそこで用いるべき典型的な表現を理解していなかったり，また別の場面については典型的な表現は理解しているが応用力は十分でなかったりする．このような学習者が目標指向の会話モデルに基づいた会話を行えるようになるためには，システムは会話の中で用いる典型的な表現を会話の状況に即して教える機能と，会話の流れの中での応用力を養成する機能の両方をうまく備えることが重要である．

そこで，IVD および音声認識技術を用いて，一つの会話の教材を段階的に教育する方法を以下に詳述する．なお，すでにある程度のレベルに達している学習者のモチベーションを低下させないために，本方式では学習者が練習の種類を自由に選択できるようにしている．**図 1** にこの教育方法に基づく教育の流れを示す．

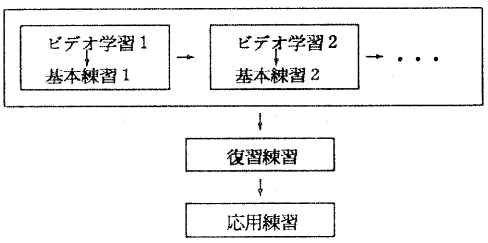


図 1 英会話用知的 CAI システムを使った学習の流れ
Fig. 1 A training method of ICAI for conversational English.

☆ 前述の会話の流れは，挨拶の場面，提案の場面，確認の場面など複数の会話の場面の続き具合をいう．

2.1 会話の状況と典型的な会話の流れの理解（ビデオ学習）

ビデオ学習は，目標指向の会話における，典型的な状況，個々の状況における会話の流れ，個々の会話の場面で使われる典型的な会話の表現，を学習者に理解させることを目的としている．すなわち，表 1 の(2)～(5)の項目を学習者に十分理解させることを目的としている．

そのためシステムは学習者に対し，(a)会話の目標とその会話で起こりうるいくつかの状況を，**図 2** のように画面上に提示する，(b)学習者にその中から一つの状況を選択させる，(c)選択した状況における模範的な会話を録画したビデオを見せる，といったことを行う．これらによって，その会話の流れを理解させる．

映像はランダムアクセス可能なレーザディスクに記録されているため，システムは，学習者の要求に応じて任意の時点に瞬時に移動しそこからビデオを再生すれことできる．場面間の移動を学習者が簡単に操作で

目標 面談したい相手に電話をかけて食事の約束をする
状況 (1) 相手と直接電話で話をし約束をする
(2) 相手が外出中なので秘書に伝言を残す
(3) 相手が外出中なので秘書に外出先の電話番号を尋ねる
(4) BCD 社に間違い電話をかける
(5) 一般家庭に間違い電話をかける

図 2 電話の会話におけるさまざまな会話の状況
Fig. 2 Examples for the substory of telephone conversation.

きるようにするため，マウスで操作できるスライダーを画面上に用意している．スライダーは映像の進行に従って左から右に進むようになっており，学習者は，そのスライダーをマウスでドラッグすることで任意の場面に移動できる．また，学習者は，必要に応じてビデオ中で用いられる発話文やその日本語文を，ビデオと同期してテキストとして見ることができる．

2.2 典型的な会話の練習（基本練習）

基本練習は，ビデオ学習で使用した映像に現れる発話文を学習者に実際に話させることを目的としている．そのため，ビデオ学習の映像に現れる適当な一人の役割を学習者に演じさせ，あたかも映像の中の会話相手と実際に会話をするように音声で会話をさせる．

基本練習の具体的ステップを以下に示す．

1. ビデオ学習と同じように，システムは学習者に会話の状況を選択させる．
2. システムは選択された状況のビデオを再生し，学習者が発話すべきところで映像を静止させて，学習者の入力を持つ．
3. 学習者は，ビデオ学習で学習したとおりの会話文をシステムに対して，音声で入力する．
4. 学習者の入力中に誤りが含まれていた場合は，システムはその誤りを記録し，それに対する指導を行う．

学習者の発話に誤りが含まれていた場合，その誤りに対する指導を随時行うかあるいは後でまとめて行うかといった教授戦略は，あらかじめ学習者が設定できるようにしている．誤りの同定およびそれに対する指導については第6章で述べる．誤りがなかった場合および指導を行った後は，システムは次の場面に進む．学習者は，何を発話したらよいかわからない場合やシステムの発話した内容がわからない場合には，システムに対してヒント要求のメニューやボタンを選択することでヒントを見ることができる．この機能は以下の復習練習および応用練習でも使える．

2.3 臨場感を高めた会話練習（復習練習）

復習練習は，すべての状況に対してビデオ学習，基本練習を繰り返し行い，各状況で発話するべき文を十分暗記した後の練習である．この練習の目的は，学習者がここまでに練習した会話の状況を正しく識別し，状況に即した会話を行えるようにすることである．そのため，学習者が会話の状況を選択するのではなく，システムが基本練習で用いた会話の状況の一つを選択する．システムはどの会話の状況を選択したかを学習者に提示しないので，学習者はシステムの発話を正しく聞きとり，聞きとった内容からどの状況の会話の練習をしているのかを判断し，適切な応答をしなければならない．この練習によって，システムおよび学習者は先の練習での達成度，すなわち，表1の(2)～(6)の達成度を再度確認できる．どの状況での会話練習を行うかは，基本練習での各会話の状況ごとの，誤りの回数，ヒントの使用状況，練習回数などに基づいてシステムが選択する．

2.4 さまざまな表現，さまざまな会話の流れを使う練習（応用練習）

応用練習は，学習者がこれまでの練習で身につけた知識の応用力を高めることを目的としている．そのため応用練習では，ビデオ学習，基本練習，および復習練習などで使用した会話の内容と会話の目標は同じであるが，内容の異なった会話をシミュレートする．ただし，学習者には，基本練習や復習練習と同じ役割を演じさせる．応用練習のシステムの発話文は，基本練習，復習練習で用いたシステムの発話文と(a)意味的には同じであるが，構文や単語が異なるもの，(b)会話の中で出てくる，「約束の時間」，「電話番号」など内容が異なるもの，などがある．また，学習者がビデオに収録された文と異なる文を発話した場合でも，システムは，その発話にある程度追随し会話を続ける．システムは，どのような内容の発話を行うかを，今までの練習で判明している学習者のレベルや，ヒントの使用状況，乱数などによって決める．

3. システムの構成

本章では，第2章で述べた教育機能を実現する英会話用知的 CAI システムの構成について述べる．

図3にシステムの構成を示す．本システムは，音声認識部，会話制御部，指導部，学習者モデル部，およびマルチメディア制御部の五つのモジュールと教材知識ベースおよび教授知識ベースから構成される．音声認識部は，学習者の音声入力を処理するモジュールである．詳細は第4章で述べる．会話制御部には，学習者が音声で会話文を入力した場合，音声認識部の出力する文字列が入力される．また，学習者がキーボードから会話文を入力した場合は，入力された文字列がそのまま入力される．会話制御部は，入力された文字列を解析して文法誤りや単語誤り等の誤り発見および次のシステムの発話の決定を行う．誤りが発見された場合，会話制御部は学習者の誤りを指導するための教授

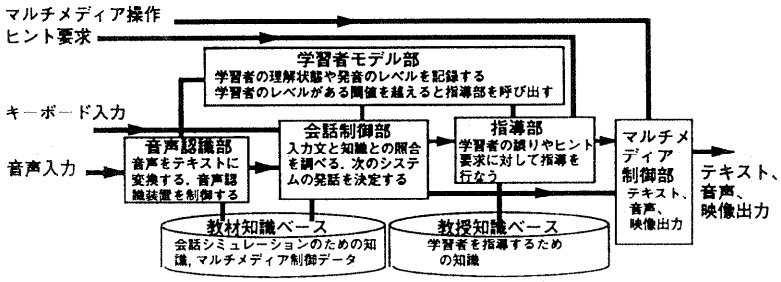


図 3 英会話用知的 CAI システムの構成
Fig. 3 An architecture for ICAI system for conversational English.

知識を指定して指導部を呼び出す．指導部は，学習者モデル部や会話制御部から指定された教授知識を用いて推論を行い，その結果，学習者モデルの変更および学習者への指導を行う．学習者モデル部は，学習者モデルの管理を行う．学習者モデルの詳細については，第6章で述べる．マルチメディア制御部は，会話制御部または指導部の指定または学習者からの要求によって，音声，映像，およびテキストの出力を行う．本システムでは，映像および音声の一部は，計算機でランダムアクセス可能なレーザディスクに格納し使用する．教材知識ベースは，本システムの対象とする英会話に関する知識を格納している．教授知識ベースは英会話の指導に関する知識をルール形式で格納している．教材知識ベースおよび教授知識ベースの詳細については第5章で述べる．

4. 英会話用知的 CAI システムに音声認識システムを統合する方法

本章では，第2章で説明した基本練習，復習練習，および応用練習といった会話シミュレーション機能を実現するために，本システムの音声認識部においてどのように既存の音声認識システムを統合するかについて述べる．

4.1 音声認識システムの概要

音声認識システムの研究は大別して二つに分けられる[7]．一つは，極めて大きな語彙（例えば，2×10^4語）で，文型も話題も限定しないが，単語または文節を孤立発声することや対象とする特定話者の音声による訓練が必要なものである．もう一つは中程度の大きさの語彙（1,000語程度）で文型・話題に制約を設ける代りに，連続音声，不特定話者を対象とするものである．

英会話用知的 CAI システムでは，会話の臨場感を保つことが重要であるので，学習者が連続音声で発声できる後者の方式のシステムを採用する．この方式のシステムは，あらかじめ作成された文法規則（後述）によって生成される文（候補文）の中で，入力された音声と一致度の高い文とその一致度（評価点）の組を出力することができる．出力された組の中で最も高い評価点を持つ組の文を認識結果文と呼ぶ．もし，評価点がある閾値より低い場合，音声認識システムは認識結果なしを出力する．一般に評価点の平均は個人によって異なる．

音声認識の候補文を文法規則を使って記述した例を**図4**に示す．図において，“_s” は開始記号，“_” で始まる他の識別子（例えば，“_AGREE” など）は非終端記号，そして文字または数字で始まる識別子（例えば，“goodbye”）は終端記号である．

```
_s → _THANKS _AGREE goodbye
_s → Ok goodbye
_s → Goodbye

_THANKS → Thank you
_THANKS → Thank yon mr doe

_AGREE → so am i
_AGREE → i am too
```

（a） 音声認識システムで用いる文法規則の例
（a） Examples for grammar for a speech recognition system.

```
Thank you, so am I. Goodbye.
Thank you, I am too. Goodbye.
Thank you. Mr. Doe, so am I. Goodbye.
Thank you, Mr. Doe, I am too. Goodbye.
Ok. Goodbye.
Goodbye.
```

（b） （a）から生成される候補文の例
（b） Sentences generated from (a).

図 4 音声認識システムで用いる文法規則と候補文の例
Fig. 4 Examples for grammar and sentences for a speech recognition system.

4.2 会話の臨場感を保つための音声認識の制御

4.2.1 学習者とシステムの特性に基づく制御

文型・話題を制限した音声認識システムも認識精度は完全ではなく認識誤りを犯すことがある．英会話の学習者ができるだけ自然に会話文を入力できるようにするためには，このような音声認識システムの特性を考慮し会話シミュレーションを実現しなければならない．音声認識システムの認識誤りの症状としては，

1. 学習者が候補文の一つを入力したにもかかわらず認識結果なしが出力される
2. 学習者が候補文以外の文を入力したにもかかわらず候補文の一つが閾値より高い評価点をとり認識結果文として出力される
3. 学習者が候補文の一つを入力したにもかかわらず，別の候補文が閾値より高い評価点をとり認識結果文として出力される

といったものが考えられる．これらの認識誤りが起きる原因としては，(a)学習者の発音が悪いことや学習者が言い淀んだこと，(b)閾値の設定が適切でないこと，(c)音声認識システムの能力の限界，が考えられる．システムがこれらの原因を区別することは不可能であるが，音声による会話シミュレーションを実現するためには何らかの処置を講ずる必要がある．

症状1が生じた場合，システムは一つの発話の場面で最高3回までの再入力を学習者に要求する．もし，3回目の入力に対しても音声認識システムが認識結果なしを出力したときは，4回目の入力は，キーボードを使うように要求する．この制御は特に原因(a)の場合，つまり学習者が入力した発話に対して，「自分は発音があまりよくない」とか「たった今の入力は言い淀んでしまった」ということを自覚している場合に効果的である．また，その自覚がない人においても再入力させるという制御は，教育システムという本システムの性格から自然なものであると考えられる．また一方では，原因(b)の閾値が高すぎる場合も考えられるため，一つの発話で4回以上の入力を行ったときは閾値を所定量だけ下げる制御を行い，本症状が継続することを防止する．また，「自分は発音があまりよくない」と思っている学習者には，発音練習モードが用意されている☆．

症状2は，原因(b)の閾値が低すぎる場合に発生する．ここで，学習者の評価点平均と個々の評価点を観察すると，一般に学習者が候補文の一つを発話した場合の評価点は，その学習者の評価点が平均より高く，候補文以外を発話した場合は，評価点平均より低い．そこで，学習者ごとの評価点の平均に沿って閾値を変化させることで，ある程度の認識誤りを回避することができる．閾値および評価点の平均は学習者モデルに記述する．

症状3の場合，学習者の発音が悪いとき（上記原因(a)）は上記のように閾値を高くすることで別の候補文を誤って認識することを避けることができる．一方，学習者の発音が悪くないときは，その主な原因は原因(c)であると考えられる．しかしながら，音声認識システムそれ自体の能力の改善は本論文の範囲外なので，その代替として認識結果の表示機能と直前の学習者の発話場面に会話を戻す機能をシステムに設ける．学習者が会話の流れに違和感を覚えた場合には，これらの機能を用いて画面に表示された前回の発話の認識結果と自分の意図した発話との比較を行う．そして意味的にも異なっていた場合だけ会話の場面を戻して再入力する．システムが学習者の発話を一字一句正しく認識していなくても，同義文を認識しているならば，学習者は再入力せずにそのままの状態で会話を継続しても構わない．会話の場面を戻す機能は，画面上のメニューを選択するだけでなく，できるだけ人間同士の会話に近付けるために音声コマンド（例えば，“roll up” など）によっても起動できるようにする．

4.2.2 音声認識の認識候補文の動的変更

本システムで採用した音声認識システムの方式は，文法規則で生成される候補文の数が多いほど認識誤りが多くなる．そこで，候補文の数をできるだけ少なくし，認識精度を高く保つために，本システムでは会話の場面ごとに文法規則を変更している．

5. 教材知識の知識表現

本章では，第2章で述べた段階的な会話シミュレーションに基づく教育機能を実現するための教材知識の知識表現について論じる．

5.1 教材の概念構造

知的 CAI システムの教材知識は，教育しようとする領域の問題を解決するための知識であり，また知的 CAI システムが学習者に伝達する対象となる知識である．本システムは英会話教育を対象としているので，教材知識は，設定された目標に沿った，種々の状況，会話の流れ，使われる会話文を表現できる必要がある．そこでこれら四つを**図5**に示すように層状の構

☆ 本論文では，発音練習モードの詳細については省略する．

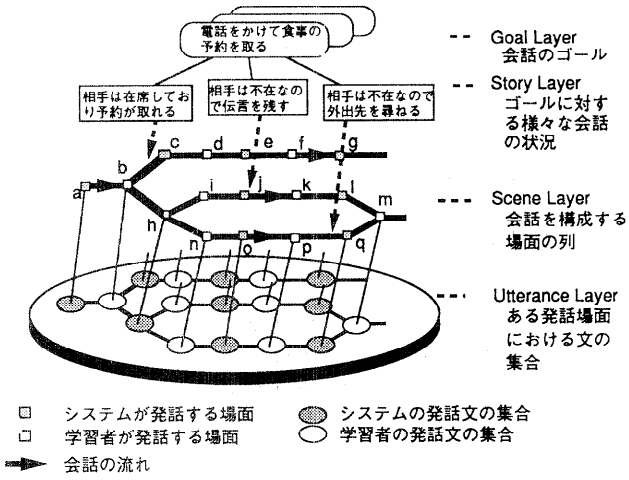


図 5 会話シミュレーションのための教材の概念構造
Fig. 5 Conceptual architecture for courseware.

造でモデル化する.

Goal Layer: Goal Layer は, この構造の最も上位の層であり,「面談したい相手に電話をかけて食事の約束をする」や「ホテルのフロントで部屋の予約をする」といった会話の目標を表現する層である. Goal Layer のノードは教材の種類に対応する.

Story Layer: Story Layer は Goal Layer で設定した目標のもとで会話を行う場合に実際に起こりうる種々の状況を表す層である. 例えば,「面談したい相手に電話をかけて食事の約束をする」という目標に対して, 図2に示すようないくつかの状況が考えられる.

Scene Layer: Scene Layer は各状況における会話を構成する場面の列を表す層である. Scene Layer は図5に示すように, 場面を頂点(ノード)とし, 場面間の遷移を有向辺で表した有向グラフとして表現される. この有向グラフは, 会話の始まりを示す一つの始点と, 会話の終りを示すいくつかの終点を持つ. 始点から一つの終点までの経路が会話の状況 (Story Layer のノード) に沿った一つの会話を表す. 例えば,「相手は在席で予約がとれる」という状況の会話は,「電話を受けた先が会社の名前を名乗る場面」を始点として,「相手を呼び出してもらう場面」,「挨拶をする場面」,「食事の申し込みをする場面」, といった会話の場面を表す頂点を結ぶ経路として表現される. 図5では, 会話の場面を英文字のついた頂点として表している. さらに Story Layer の頂点「相手は在席しており予約がとれる」,「相手は不在なので伝言を残す」, および「相手は不在なので外出先を尋ねる」が表す状況は, それぞれ Scene Layer の会話の場面の列 $\langle a, b, c, d, e, f, g\rangle$, $\langle a, b, h, i, j, k, l, m\rangle$, $\langle a, b, h, n, o, p, q, m\rangle$ という列から構成されることを示している. a および b は上記三つの状況に共通して現れる場面であり, それぞれ,「電話を受けた先が会社の名前を名乗る場面」,「相手を呼び出してもらう場面」に対応するものである.

Utterance Layer: Utterance Layer は, Scene Layer の頂点に対応した各場面における発話文の集合を表現する層であり, 図5に示すように Utterance Layer の頂点は Scene Layer の頂点と一対一で対応している. システム側の発話文の集合には, その場面におけるふさわしい文として, 丁寧さなどの異なる文, 難易度の異なる文といった同じ意味を表す異なる表現の文が含まれる. 学習者側の発話文の集合には, システム側の発話文の集合に含まれる文と同じようにその場面にふさわしい同じ意味を表す異なる表現の文だけでなく, 学習者が犯すと予想される種々の誤りを含んだ文も含まれる. 後者の文は誤りを含んだ学習者の発話を理解するために使用される.

このモデルに沿うと, ビデオ学習, 基本練習, および復習練習における会話の状況の選択は, Story Layer の頂点の選択に対応する. これらの練習時のシステムと学習者の間の会話は Story Layer の一つの頂点が指す Scene Layer の始点から終点までの経路をたどることに対応する. システムの発話文は, Scene Layer の頂点に対応した Utterance Layer の頂点に含まれている一つの文に対応する. 誤りのある発話文も含めた学習者が入力可能な発話文は, Utterance Layer の頂点の文の集合に対応する. 一方, 応用練習における会話は, Scene Layer の始点から終点までの経路のうち Story Layer の頂点とつながっていないものも含まれる. 応用練習の場合のシステムの発話文は, Scene Layer の頂点に対応した Utterance Layer の頂点の文の集合に対応する.

表 2 学習者ノードの形式
Table 2 Learner node frame.

スロット名	内　容
Node-ID	ノードの識別子 (例). GREETING 1, RESERVATION 1, MEETING-TIME 2, …
Input-Sentences	システムが認識する文の集合に関する記述 (例). (図6参照)
Hints	学習者に与えるヒントを決めるためのルール (例). **IF** ヒントの要求回数が1回以上である **THEN** (*sentence*) Good afternoon, Mr. Jack Doe, this is Taro Tanaka of XYZ. **ELSE** (*hints*) 挨拶をして自分の名前を名乗りましょう. **ENDIF**
L-Rules	次のパートナーノードを決めるためのルール (例). **IF** 現在の練習モードは，応用練習である and ***greeting*** は2以上である **THEN** GREETING 2 というノードに進む **ENDIF**

このモデルは，英会話に必要な知識である，会話の目標，会話の状況，会話の流れ，および会話文の関係を簡潔に表現しているとともに，第2章で述べた教育方法とこれら英会話に必要な知識の関係を簡潔に表現している．そのためこのモデルを用いることで教材作成者は，教材知識を上位の層から段階的に設計していくことができる．以下にこのモデルに沿った具体的な教材知識の知識表現形式を示す．

5.2 会話知識の表現

システムの知識表現は，上記のモデルのうち具体的な場面の遷移を表している Scene Layer と個々の発話文に関する知識を表している Utterance Layer を表現できる必要がある．そこで，本システムでは会話の場面を一つの知識の単位とするように教材知識を表現する．すなわち，会話シミュレーション時に，学習者が演ずる発話者の発話文に関連する知識と，システムが演ずる発話に関連する知識表現を知識の単位にする．学習者が演ずる発話者の発話文，すなわち学習者が入力する文に関連する知識の単位を学習者ノード，システムが演ずる発話者の発話に関連する知識の単位をパートナーノードと呼ぶ．

学習者ノードは，学習者が発話する文を認識するための知識（Input-Sentences スロット），その発話に関する学習者へのヒント知識（Hints スロット），および学習者ノードからパートナーノードを結び会話の流れを形成するための知識（L-Rules スロット）から構成される．学習者ノードの形式を**表2**に示す．Input-Sentences スロットの知識の形式は，音声認識部に与える文法規則の形式を拡張して，誤りの指導のための教授知識の呼び出しを記述できるようにした形式である．Input-Sentences スロットのデータの例を**図6**に示す．図6中の(1)(2)(3)の '{' と '}' に囲まれた部分は，誤りに関する指導のための教授知識モジュールが記述されている．Hints スロットには，学習者に与えるヒント情報が記述される．L-Rules スロットには学習者ノードからパートナーノードへの遷移を決めるためのルールが記述される．このルールは，

```
          _s → _GREETING_INTRODUCTION
          _s → _GREETING
                 {remedy (MissingIntro);}…(1)
_GREETING → good afternoon_Jack
      _Jack → Mr. Jack Doe
      _Jack → Jack Doe
                 {remedy (UsageHisName);}
                                        …(2)
_INTRODUCTION → _PRO_NAME_COMPANY
       _PRO → this is
       _PRO → my name is
                 {remedy (UsageMyName);}
                                        …(3)
      _NAME → taro tanaka
   _COMPANY → of xyz
```

図 6 学習者ノードの Input-Sentences スロットの知識例
Fig. 6 Knowledge example for Input-Sentences slot in Learner node frame.

表 3 パートナーノードの形式
Table 3 Partner node frame.

スロット名	内　　容
Node-ID	ノードの識別子 (例). GREETING 2, MEETING-TIME 1,…
Output-Sentences	システムが発話する文と，それに関連するマルチメディア制御情報 (例). 1 *sentence*: Friday is fine. *video*: 0130 to 0154. *voice file*: audio0130. 2 *sentence*: On Friday? Yes, That's fine. *video*: 0155 to 0174. *voice file*: audio0155.
Hints	学習者に与えるヒント情報 (例). 1 *sentenec*: 金曜日は大丈夫です. 2 *sentence*: 金曜日ですか．かまいません.
P-Rules	次の学習者ノードを決めるためのルール (例). **IF** 現在の練習モードは，応用練習である and *meeting-time* は 0 である **THEN** MEETING-TIME 2 というノードに進む **ENDIF**

学習者モデルや現在のシステムの状態を参照して次のパートナーノードを決める.

パートナーノードは，システムが発話する発話文に関する知識とその場面に関連したマルチメディア制御データ (Output-Sentences スロット)，システムの発話文に関するヒント知識 (Hints スロット)，およびパートナーノードと学習者ノードを結び会話の流れを形成するための知識 (P-Rules) から構成される．パートナーノードの形式を**表 3** に示す．Output-Sentences スロットには，システムが発話する文とそれに関連するマルチメディア情報を複数組記述することができる．システムは，複数組のうちのどれを選択するかを，学習者のレベルに応じて動的に決定する．P-Rules スロットには，次の学習ノードを決めるためのルールが記述される.

5.3 会話の制御

会話制御部は，学習者ノードとパートナーノードを使って，次の(1)〜(3)のステップを繰り返すことでシステムと学習者の会話を行う.

(1) システムがパートナーノードに沿って学習者に対して発話を行う.

(2) 学習者の発話を認識した認識結果文を受けとると，それが会話の流れに沿った学習者ノード，すなわち，(1)で発話したパートナーノードの P-Rules で決められた学習者ノードの Input-Sentences スロットの知識とマッチするかどうかを調べる.

(3) もしマッチしなければ会話の流れの誤りとして学習者モデル (第 6 章参照) に記録し，指導部を起動する．もしマッチすればその学習者ノードの L-Rules スロットに沿って次のパートナーノードを決定する.

6. 学習者の誤りの同定と誤りに対する指導

6.1 誤りの同定

システムは，学習者の発話中の単語や文法の誤り[3]，言語機能に関する誤り[8),9)]，および学習者がシステムに対して要求するヒントの内容と頻度を学習者モデル部に記録する．学習者モデル部は学習者のレベルを判別し学習者のレベルがある閾値を越えると指導部を呼び出す.

本システムでは学習者の誤りの同定は，あらかじめ誤りを発見するための知識を用意しておくバギーモデルに基づいて行う[10]．学習者がキーボードを使って入力した場合，システムは，構文解析のルールに誤り発見のためのルールを追加しておきそれを使って文法誤りを発見する[3]．一方，学習者が音声で入力した場合は，学習者ノードの Input-Sentences スロットに文法誤りや言語機能に関する誤りを発見するための文法規則を追加しておき，それによって誤りを発見する (図 6 の(1)(2)(3)).

6.2 学習者のレベルの表現

本システムでは学習者のレベルを学習者モデル部で管理している変数の値によって表現している．学習者モデル部で管理する変数には，(1)誤りに関するもの，(2)ヒント要求に関するもの，(3)音声認識システムの制御に関するもの，(4)学習者の学習履歴に関するもの，がある．誤りに関しては，言語機能誤り，文法誤り，スペル誤り (キーボード入力時のみ) などがある．**表 4** に本システムが扱う誤りに関する変数の一部を，**表 5** にそれ以外の変数の一部を示す．変数は，教材作成者が教材に応じて新たに追加できるようになっている.

学習者モデル管理部では，学習者のレベルを示す変数名，変数の値，変数の閾値，およびその閾値を越えると学習者に対する指導や他の変数の値の変更を行う

表 4 誤りを記録するための学習者モデル部の管理する変数
Table 4 Predefined student level variables for errors.

変数名	意味
reserve_judgement	態度を保留する
expect	自分の期待を表明する
thank	相手に感謝する
greeting	相手に挨拶をする
spelling	スペル誤り
pronounciation	発音誤り
miss_order_subject	主語の位置が正しくない
verb_illegal_clause	この動詞は節を目的語としてとらない

表 5 誤り以外の学習者モデル部の変数
Table 5 Predefined student level variables except for errors.

変数名	意味
Hint_for_partner	システムの発話に対するヒントの要求回数
Hint_for_learner	学習者の発話に対するヒントの要求回数
SR_threshold	音声認識システムの評価点の閾値
SR_Average	音声認識システムの評価点の平均
SR_count	現在の場面における音声入力の回数
Basic_exercise_1	会話の状況1の基本練習の回数
Basic_exercise_2	会話の状況2の基本練習の回数
Basic_exercise_n	会話の状況nの基本練習の回数
Semi_advanced_exercise	復習練習の回数
Advanced_exercise	応用練習の回数
Current_exercise_mode	現在の練習モード

教授知識名の四つ組を管理している.

6.3 教授知識

本システムでは,教授知識を,(1)学習者の誤りに対する指導,(2)ヒントの制御,(3)会話の流れなど教授内容の変更,および(4)学習者のレベルを表す変数の値の変更,を行うために用い,if-then ルール形式で記述する.

誤りに対する指導を行う教授知識は,「電話の応対で相手の名前の呼び方を誤った場合」,「名前の告げ方を誤った場合」などのように誤り別にモジュール化されており名前(教授知識名と呼ぶ)で参照される.教授知識名は,学習者ノードの Input-Sentences スロットや学習者モデル管理部に記述され,それぞれ,学習者の誤りが発見された時点,学習者のレベルを表す変数が閾値を越えたときに呼び出される.誤りの指導は動画,音声,テキストを組み合せて行うことができる.

ヒント知識は学習者ノードおよびパートナーノードの Hints スロットに記述される.本システムは,さまざまなレベルのヒントを会話の場面ごとにあらかじめ用意しておき,学習者のレベルや要求に応じて適切なヒントを学習者に提示する.

教授内容の変更のための知識は,学習者ノードの L-Rules スロットやパートナーノードの P-Rules スロットに記述され,復習練習時の会話の状況の選択や,応用練習時の会話の流れの選択に使用される.

7. 実現例

7.1 実装

本システムは,ワークステーション OKITAC-S 4[☆1]に音声認識システム DS 200[☆2]とレーザディスク再生装置 LDP-2000[☆3]を接続したハードウェア上にC言語で実装されている.DS 200 は,35,000 語の語彙を持ち,文型・話題に制約を設ける代りに不特定話者の連続音声を認識する[11].DS 200 に与える候補文は,有限状態文法と呼ばれる文法で記述される.有限状態文法は,文脈自由文法の形式で記述される文法であり,有限の長さの文しか生成しないように,言語の生成の過程で一つの文法規則の使用回数を有限回に制限している文法である[☆4].レーザディスクには,会話の状況の映像と音声が格納される.今回の試作では,基本練習,復習練習および応用練習のシステム側の音声はネイティブの発話の録音音声を用い,誤りの指導などの音声は日本人の発話の録音音声を用いた.

7.2 教材知識

「面談したい相手に電話をかけて食事の約束をする」という目標の教材を実際に開発した.この目標に対して,会話の状況として,図2に示す5種類の会話の状況を用意している.図6に学習者ノードの Input-Sentences スロットの一部を示す.この中で,(1)〜(3)は,学習者の発話の中の誤りを発見するための知識である.学習者の誤りを発見するための知識は,教育の専門家が収集した学習者の対話中の誤りから抽出し,それを教材開発者が本システムの知識表現形式に変換した.(1)〜(3)を使って誤りが発見されたときは,それぞれ電話の応対で自己紹介が欠けた場合の教授知識 (MissingIntro),相手の名前の呼び方を誤った

☆1 OKITAC-S 4 は,沖電気工業(株)の製品であり,Sun Microsystems, Inc. Sun Workstation の相当品である.
☆2 DS 200 は,(株)オージス総研の製品である.
☆3 LDP-2000 は,ソニー株式会社の製品である.
☆4 厳密には,有限状態文法は有限個の文しか生成しないので正則文法の部分クラスである.

スロット名	内　　容
Module_name	UsageHisName
Comment	電話の応対で相手の名前の呼び方を誤った場合の指導

Rule	**IF** (name_call_error==0)
	THEN
	message ("相手の名前には，Mr. を付けましょう.");
	ELSE
	message ("相手の名前には，Mr. を付けましょう.
	例えば，Mr. Tom Doe のようにいいましょう.");
	name_call_error=name_call_error+1;
	ENDIF

Module_name	UsageMyName
Comment	電話の応対で名前の告げ方を誤った場合の指導
Rule	audio_message ("電話での自己紹介 誤り: My name is [your name]. 正: This is [your name].", "kr002.au"); my_name_error=my_name_error+1;

図 7 教授知識の例
Fig. 7 Examples for tutoring knowledge.

場合の教授知識（UsageHisName），名前の告げ方を誤った場合の教授知識（UsageMyName）が起動される．

7.3 教授知識

この教材に現れる場面の一つである電話の応対場面で，相手の名前の呼び方を誤ったときの教授知識モジュール（UsageHisName），および名前の告げ方を誤った時の教授知識モジュール（UsageMyName）の例を**図7**に示す．この例で，name_call_error と my_name_error は，この教材のために導入された学習者モデル部の変数であり，それぞれ，電話の対応で相手の名前の呼び方を誤った回数と自分の名前の告げ方を誤った回数を記録するために使用される．message は引数の文字列を画面に表示する関数であり，audio_message は，第一引数の文字列を画面に表示し，第二引数の音声ファイルを再生する関数である．

7.4 会話例

図8に，この教材を使ったシステムと学習者の会話例を示す．これは図2中の「相手と直接電話で話をし約束をする」という状況の基本練習の会話で，学習者は面談の約束をとりつけるために電話をかける役割を演じ，システムは，交換手および約束の対象である人物の役割を演じる．図中で，Learner，Partner，および Teacher は，それぞれ，学習者の発話，会話の流れに沿ったシステムの発話（パートナーの発話），および指導などパートナーの発話以外のシステムの発話（または表示）を示している．この会話では，音声認識システムの閾値は 600 点に設定している．学習者の発話に附属している点数は，学習者の発話の評価点である．例えば，(1)の学習者の発話の評価点は 638 点であり閾値 600 点を越えている．システムは，図6および図7の知識を用いて学習者と会話を行う．

学習者の入力(3)に対して，システムは UsageHisName を使った指導(4)，および UsageMyName を使った指導(5)を行っている．指導を行った後，システムは次の発話(6)に進んでいる．(7)では，学習者は先の発話をもう一度行うために，「後戻りボタン」を選択している．「後戻りボタン」はシステムの発話を一回前の発話に戻すので，(8)は，(2)と同じ発話である．続く(9)では学習者は正しい発話を行っている．

(10)の発話の評価点は閾値に満たなかったので，システムは(11)の発話で，再入力を要求している．(12)の学習者の入力は閾値を越えている．(14)では学習者は，パートナーの(13)の発話がうまく聞きとれなかったので，「繰り返しボタン」を押している．それに対して，システムは，(15)でパートナーの発話を繰り返している．

(17)では学習者は，パートナーの(16)の発話の内容を理解できなかったので，「パートナーの発話に関す

Partner:	"ABC Company. May I help you?"		
Learner:	"I would like to speak to Mr. Jack Doe please."	638	…(1)
Partner:	"Just a moment please. "		
	"Sales department. Jack Doe speaking."		…(2)
Learner:	"Good afternoon Jack Doe. My name is Taro Tanaka of XYZ."	666	…(3)
Teacher:	相手の名前には，Mr. をつけましょう		…(4)
	電話での自己紹介　誤: My name is [your name]. 正: This is [your name].		
	"電話の会話では，This is を使います．例えば 'This is Taro Tanaka' のようにいいます．"		…(5)
Partner:	"Good afternoon Mr. Tanaka. What can I do for you? "		…(6)
Learner:	<「後戻りボタン」を押す>		…(7)
Partner:	"Just a moment please. Sales department. Jack Doe speaking."		…(8)
Learner:	"Good afternoon Mr. Jack Doe. This is Taro Tanaka of XYZ."	634	…(9)
Partner:	"Good afternoon Mr. Tanaka. What can I do for you?"		
Learner:	"I'm in town for a few days and I'd like to meet you if it's convenient."	590	…(10)
Teacher:	Please repeat.		…(11)
Learner:	"I'm in town for a few days and I'd like to meet you if it's convenient."	631	…(12)
Partner:	"Yes, I'd like that. What day is good for you?"		…(13)
Learner:	<「繰り返しボタン」を押す>		…(14)
Partner:	"Yes, I'd like that. What day is good for you?"		…(15)
Learner:	"How about tomorrow?"	693	
Partner:	"I'm afraid I'm rather busy tomorrow. Umm, what about the day after tomorrow Friday?"		…(16)
Learner:	<「パートナーの発話に関するヒントボタン」を押す>		…(17)
Teacher:	I'm afraid I'm rather busy tomorrow. Umm, what about the day after tomorrow Friday?		…(18)
Learner:	<「学習者の発話に関するヒントボタン」を押す>		…(19)
Teacher:	相手の提案を承諾し，昼食をいっしょにどうかと提案する．		…(20)
Learner:	<「学習者の発話に関するヒントボタン」を押す>		…(21)
Teacher:	Yes, that's fine. Shall we meet for lunch?		…(22)
Learner:	"Yes, that's fine. Shall we meet for lunch?"	713	

図8 会話例
Fig. 8 A subdialogue between the system and a learner.

るヒントボタン」を押している．

それに対して，システムは，(18)でパートナーの発話を英文で表示している．(19)では，学習者は次に発話するべき内容についてわからないので，「学習者の発話に関するヒントボタン」を押している．それに対して，システムは(20)のように学習者にどのような内容の発話をすればよいかを提示している．

(21)でさらに学習者が「学習者の発話に関するヒントボタン」を押すと，システムはより直接的なヒントを(22)のように学習者に提示する．

8. 考　察

本システムはまだ開発途中であり，今後，実際の現場で試用しながら学習者や教育者の意見をシステムに反映させていく必要がある．現在，研究室レベルでの評価が完了している．研究室レベルの評価では，英会話の教育者と著者の所属する組織の従業員が試用した．本章では，システムを試用した際の評価とシステムの設計に関する考察について述べる．

8.1 システムを試用した際の評価

8.1.1 教育的な効果

研究室レベルの評価では，教育的に重要な結論が得

られた．一つは，学習者はシステムを楽しんで使用していたことである．理由としては，IVD の導入により学習者が会話の状況を理解する際に学習者の意のままにビデオを操れたことと，実際に音声による対話ができたことが上げられた．もう一つは，システムの教育方法は，英会話の初心者から中級者の間のいずれのレベルの人にも適当であるように観察された．

8.1.2 音声対話機能

本システムの音声対話機能を評価するために，通常の運用状態で，教材知識中に音声認識候補文として記述されている文をシステムが正しく認識できるかどうかを調べる実際を行った．ここでは，日本人の被験者5名が会話の場面10カ所*で音声認識候補文を発話するという実験を二度行った．被験者には，仕事では英会話を話すことがほとんどない人で，さまざまな英語のレベルの人を選んだ．実験では，一度目と二度目は異なる文を発話し，音声認識の閾値は600点に設定した．この発話文には文法的な誤りを含んだ文も含まれている．

個々の音声の入力を集計した結果を**表6**に示す．表6の「認識結果が誤っている」のうち9件は，“in”と“on”の認識誤りや“John Doe”と“Mr. Doe”の誤りといった一単語の認識誤りである．今後より精度の高い音声認識システムが開発された場合には，それを採用することで，これらの認識誤りを減少できると考えられる．しかしながら，今回の実験では，すべての場面において被験者の少なくとも一人は，常に正しい認識結果を得ていた．これより，本システムに採用した音声認識システムは日本人にとっても使用可能なレベルに達しているといえる．なお，4.2.1 項で述べた評価点平均を基にした閾値の制御の評価は，十分な数の発話データがなく閾値の変化量を決められないため今後の課題とする．

8.2 システムの設計に関する考察

8.2.1 学習者の入力した音声が認識されなかったときの応答

人間同士の会話を考えると，学習者の入力した音声が認識されなかった場合にパートナーが言い直しや言い換えを行うのは，学習者の発話から「学習者はパートナーの発話を誤って理解している」と判断したときといえる．一方パートナーが学習者の発話を一切聞きとれなかったとき，すなわち理解できなかったときは，パートナーは聞き直す意図で“Pardon me”と発話すると考えられる．本システムが再入力を要求するのは，学習者の発話が音声認識の閾値に達しなかった場合であり後者に相当する．したがってシステムが再入力を要求する際の発話は，前のシステムの発話の言い直しや言い換えよりも，“Please repeat”のように直接学習者に再入力を促す発話が適切だと考えられる．

表 6 音声認識の候補文を入力した場合の認識結果
Table 6 Evaluation of a speech recognition system.

項　　目	比率 (%)
認識結果が正しい	76
認識結果が誤っている	21
認識結果なし	3

* 図8の会話の場面も含んでいる．

8.2.2 学習者がシステムの発話を聞きとれなかった場合の応答

本システムでは，学習者がシステム（パートナー）の発話を聞き取れなかった場合，図8の(14)のように単に発話を繰り返すか，同図(18)のようにパートナーの発話をそのまま英文で示しているが，もし，聞きとれなかった発話にリエゾンが含まれている場合は，文献12)のように視覚的に提示し指導できる方が望ましい．このような機能の実現は今後の課題である．

8.2.3 知識表現の拡張性

従来の文献3),4)では，会話シミュレーションのための知識表現形式について記述があるが，「相手と直接電話で話をし約束をする」といった目標のある会話のための教材知識を，どのように構成すべきかといったことに対する考察が不十分である．そのため，教材知識の作成がアドホックにならざるをえないため，教材の作成や修正（拡張）が容易ではない．一方，本論文で提案した教材の概念構造に基づくと，教材作成者は，この構造の抽象的な層（上位）の構造からより具体的な層（下位）の構造へと段階的に教材を作成できるので，教材知識の開発が容易になる．階層的な構造になっているため，変更箇所を局所化できるので，以下に示すように教材の拡張も容易になる*．

1. ある場面における発話文の言い換え文を追加，変更あるいは削除したいときは，その発話文を含むUtterance Layer のノードだけを修正すればよい．すなわち学習者の発話の場合は学習者ノードを，システムの発話の場合はパートナーノードを修正すればよい．
2. 会話の流れを新たに追加することは，Story Lay-

* 教材作成支援システムについては稿を改めて報告したい．

er と Scene Layer のノードを新たに追加することに対応する．新しく Story Layer にノードを追加する際は，そのノードが既存の Scene Layer のどのノードを使用するかを調べ，不足するノードだけを新たに Scene Layer に追加する．その後，新たに追加した Scene Layer のノードに対応する Utterance Layer のノードを追加し，さらにそのノードと既存の Scene Layer に対応した Utterance Layer のノードがつながるよう既存の Utterance Layer のノードの情報を変更すればよい．

この知識表現形式自体は会話の分岐数に関する制限を持たない．しかし，会話の分岐数を増やすときには，分岐のところで音声認識の候補文が増えて認識率が低下しないように注意する必要がある．

8.2.4 システムの拡張性

パートナーノードの P-Rules や学習者ノードの L-Rules スロットには学習者の発話に対する次のパートナーノードを決めるためのルールが記述できるようになっている．本論文で述べたシステムでは，この L-Rules スロットのルールは，教材の概念構造における Scene Layer のノードのリンクを会話が進む中でどのようにたどるかを決定するために用いている．この表現形式に Scene Layer のノードの遷移を記述するための固有な部分はない．そのため，このシステムを，デパートや美術館に設置する音声による案内システムにも応用できると考えられる．ただしあくまで表層の応答の記述を考慮した表現形式であるため深いモデルの必要な分野には適さない．

9. おわりに

本論文では，対話的に操作可能なビデオ（IVD）と音声認識システムを知的 CAI システムに統合することで，典型的な英会話文の習得の支援から応用力の養成まで段階的に行うことができる英会話用知的 CAI システムについて論じた．

学習者がある状況で英会話をできるようになるには，会話の目標，その会話の目標を達成するための会話のさまざまな状況，おのおのの状況での会話の流れ，および各場面における発話文を理解していることが必要である．本システムは，これらの理解を支援するために，学習者から容易に操作可能な IVD を提供している．さらに，学習者がこれらの理解度を段階的に深めることができるように，基本練習，復習練習，および応用練習の3段階からなる「音声による会話シミュレーション機能」を実現している．

本システムでは音声によるコミュニケーションにおける学習者の認識能力と発声スキルの向上のために不特定話者連続の音声認識システムを採用した．この音声認識システムは，認識結果として文のその評価点の組の集合を出力する．学習者の入力誤りや音声認識システムの認識誤りに対処するために，認識結果の評価値が閾値より低いときは学習者に4回まで再入力を要求することや，学習者ごとに評価点の平均によって動的に閾値を変化させることを行う．

本システムで用いた会話の知識表現形式は，会話の場面の遷移を自然に記述できるので，教材作成者は従来のシステムと比較して容易に教材を拡張することができる．システムは，音声認識に与える文型の中にあらかじめ誤りを発見するための知識を含ませることで，学習者の音声入力時の誤りの同定を行うことができる．ただし，誤りを発見するための知識をシステムが獲得する方式の開発は今後の課題である．

本システムは，ワークステーション OKITAC-S 4 に音声認識システム DS 200 とレーザディスク再生装置 LDP-2000 を接続したハードウェアにC言語で実装されている．応答時間は1秒程度であり会話の臨場感を保持するのに実用上十分な速度である．

本論文で述べた音声認識システムを英会話知的 CAI システムに統合する方式は，DS 200 以外の音声認識システムに対しても適用可能である．また，音声認識の認識候補文の変更のためのルールと次の発話を決定するルールを拡張することで，このシステムを，デパートや美術館に設置する音声による案内システムにも応用できると考えられる．

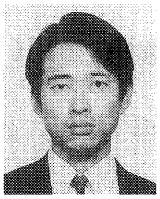
謝辞 本システムの設計にあたり熱心にご討論いただいた三重大学椎野教授，大阪ガス株式会社の宮阪信次主鑑，平山輝副課長，株式会社オージス総研の明神知副事業部長，乾昌弘主任に深謝する．また，開発にご協力いただいた株式会社オージス総研の井谷浩二氏，大場克哉氏，株式会社沖テクノシステムズラボラトリの加藤正明主任，浅野雅代主任，沢山ゆかり氏，瀬戸美枝氏に感謝する．

参考文献

1) Murray, J.: Humanists in an Institute of Technology: How Foreign Languages are Reshaping Workstation Computing at MIT, *ACADEMIC Computing*, September, pp. 34-

64 (1987).
2) DynEd International Inc.: *How to Use IIC (Intelligent Interactive Courseware)* (1989).
3) 山本秀樹，甲斐郷子，大里真理子，椎野 努：会話シミュレーションを基にした語学訓練用知的CAIシステムの構成，情報処理学会論文誌，Vol. 30, No. 7, pp. 908-917 (1989).
4) 山本秀樹，甲斐郷子，大里真理子，椎野 努：語学訓練用知的 CAI システムにおける学習者の意図の把握と会話制御方式，情報処理学会論文誌，Vol. 31, No. 6, pp. 849-860 (1990).
5) Clark, D. J.: How Do Interactive Videodiscs Rate Against Other Media?, *Instructional Innovator*, Vol. 29, pp. 12-18 (1984).
6) Matta, K. F. and Kern, G. M.: Interactive Videodisc Instruction: The Influence of Personality on Learning, *Int. Man-Machine Studies*, Vol. 35, pp. 541-552 (1991).
7) 藤崎博也：音声認識・理解の目標と将来課題，電子情報通信学会誌，Vol. 73, No. 12, pp. 1264-1268 (1990).
8) 甲斐郷子，浅野雅代，大場克哉，井谷浩二：Communicative Aproach に基づく英会話 ICAI システムの実現法，情報処理学会コンピュータと教育研究会，Vol. 13, No. 2 (1990).
9) 大場克哉，甲斐郷子：語学訓練用 ICAI システムのための言語機能情報の利用，第 42 回（平成3年前期）情報処理学会全国大会論文集 (1991).
10) Burton, R.: Diagnosing Bugs in a Simple Procedural Skill, Sleeman, D. and Brown, J. (eds.) *Intelligent Tutoring Systems*, London, Academic Press (1982).
11) 平山 輝，平島充雄：不特定話者，連続音声認識システムの開発とその応用，*Proceedings of the International Symposium: Computer World '91*, pp. 189-196, 関西情報センター (1991).
12) 岡本ほか：英会話を対象とした環境型 CAI システム，信学会研技報，Vol. 92, No. 22, pp. 29-36 (1992).

(平成4年9月14日受付)
(平成5年6月17日採録)

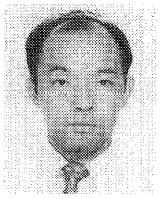


山本 秀樹（正会員）

1961年生．1984年京都大学工学部電気工学科卒業．同年沖電気工業(株)入社．データベースマシンの研究，エキスパートシステム，知識ベースシステム構築ツール，知的CAIシステム，機械翻訳等，人工知能関連の研究に従事，現在，沖電気工業(株)関西総合研究所に勤務．知識獲得と知識表現に興味を持つ．AAAI, ACL, ACM，人工知能学会，電子情報通信学会各会員．

田川 忠道（正会員）

1962年生．1987年大阪府立大学電気工学科卒業．同年沖電気工業(株)入社．現在，同社関西総合研究所に勤務．エキスパートシステム，ヒューマンインタフェース研究に従事．人工知能学会会員．



宮崎 敏彦

1981年大分大学工学部組織工学科卒業．同年沖電気工業(株)総合システム研究所入社．その後(財)新世代コンピュータ技術開発機構へ出向．現在，沖電気工業(株)関西総合研究所勤務．ヒューマンインタフェース，並列プログラミング言語等の研究開発に従事．