

日本認知科学会第20回大会 ワークショップ W-1 知的情報処理を活用した外国語学習¹

企画責任者： 原田康也

1. 発表・参加予定者 (50音順)

井佐原均(isahara@crl.go.jp)：独立行政法人通信技術総合研究所自然言語グループ

内山将夫(mutiyama@crl.go.jp)：独立行政法人通信技術総合研究所自然言語グループ

佐野洋(sano@tufs.ac.jp)：東京外国語大学

中條清美(chujo@mmm.cit.nihon-u.ac.jp)：日本大学生産工学部

中村隆宏(takahiro@shogakukan.co.jp)：小学館マルチメディア局

原田康也 (harada@waseda.jp)：早稲田大学法学部

宮崎佳典(yoshi@ssu.ac.jp)：静岡産業大学国際情報学部

宮田高志 (miyata.t@carc.aist.go.jp)：科学技術振興事業団戦略的基礎研究推進事業(CREST)

山田玲子(yamada@atr.co.jp)：ATR 人間情報科学研究所

渡辺隆行(nabe@twcu.ac.jp)：東京女子大学現代文化学部コミュニケーション学科

2. はじめに (ワークショップ企画趣旨)

電子式コンピュータの実現以来、機械翻訳がその応用の一つとして構想されてきた。コンピュータを利用した複数言語間のコミュニケーション支援としては、機械翻訳・音声翻訳通信などに加え、インターネットの普及に伴って多言語検索なども普及しつつあるが、こうした研究で蓄積された要素技術を外国語学習に応用する研究にはまだ大きな発展の余地が残されている。

外国語学習における情報処理技術の応用は、コンピュータとインターネットを中心としたマルチメディア教室の構築やCAIなどに集約されるように考えられる傾向にあるが、e-commerceにおけるSDI(選択的情報配信)をweb-based trainingに応用し、音声認識・音声合成・情報抽出などの知的情報処理の要素技術を応用することにより、外国語学習の質を飛躍的に高めることが可能となっている。

本ワークショップは、コーパス言語学・知的情報処理・計算機科学などさまざまなバックグラウンドの研究を現実の外国語学習に直接的に応用しつつある研究者があつまり、「知的情報処理を活用した外国語学習」のあり方について、その可能性と限界を探ろうとするものである。

3. 外国語学習におけるシステム支援

外国語学習の形態には、教室で一人の教員を中心に学習を進める一斉授業方式、一人の教員と小人数の学習者で対面式に行う運用練習、PCなどを利用して個人で行う自習など、さまざまなスタイルが考えられる。学習内容についても、知的発達段階と学習到達度に応じて、語彙・構文などの言語材料の習得、文章構成の基本的方略や有効な説得のための対話方略の習得など、さまざまなレベルが想定される。デジタルメディアの利用についても、メールや遠隔会議システムなどを利用して習得中の言語による実際のコミュニケーションを実体験することを重視する「コミュニケーション誘発型の学習」と音声・文字の提示やキーボード・マウス・音声などによる反応を中心とした「基本事項のドリル学習」とでは、必要となる技術的支援も異なる。

選択的情報配信を外国語学習に応用すると、学習者の到達度と学習履歴から、ネットワークにアクセスした時点で特定の語彙や表現の学習を促す画面を提示するといったシステムや、その学習者がwebのブラウザを使用して外国語の文書を見ようとしたとき、到達度を超える表現について例文を示したりパラフレーズを提示したり辞書の記述を表示することによってその読解を支援するといったシステムの介入が考えられる。

これは外国語をまだ不完全にしか習得していない学習者に対して、システムの介入によってその外国語の理解を補佐しつつその利用を支援することによって結果的に学習を成立させるという考え方であり、到達学習レベルに応じて補助的な語彙学習を促す、音声文字化して表示する、文字テキストを音声化して提示する、テキストを書き換えて提示する、内容の理解に不可欠な文化的背景の説明を加えるなど、さまざまな実現形態が構想できる。

システム構築のための構成要素には、学習の基礎原理と学習対象のモデル化、学習用基礎資料の収集・選択・開発、コンテンツの作成・展開、教材の配信と提供、学習履歴の分析と保存、学習到達度の測定などがあるが、以下ではこれに関連して語彙学習とコーパスについて、また言語資料を教材として加工する際に利用可能なさまざまな基礎的研究について紹介する。

¹ 本稿は 2.はじめに と 3. 外国語学習におけるシステム支援 を原田康也が、4. 語彙学習 を中條清美が、5. 学習資源としてのコーパス を井佐原均と内山将夫が、6. コーパスの普及と活用技術の向上を目指して を中村隆宏が、7. Global Document Annotation とその応用 を宮田高志が、8. 日英2ヶ国語音声化システムBEPと語学学習 を渡辺隆行が、9. 9. 教員のニーズに基づく学習履歴の分析と表示 を宮崎佳典が担当し、全体を原田康也が取りまとめた。

4. 語彙学習

4.1. 語彙学習の重要性

一般に学習者は語彙の知識を内容理解の第一の拠り所とする(Huckin & Bloch [29])ため、語彙不足は学習者にとって重大な障害になる(Ulijin [36])。従って、このような問題点を克服するには、教材選択の際、学習者の語彙レベルと教材の語彙レベルが乖離しないような配慮が必要である。そのためには、語彙レベルを計測するための信頼性の高い尺度を設けた上で、1) 学習者の語彙レベルはどの程度か、2) 当該英文素材を理解するのに必要な語彙レベルはどの程度かを計測する必要がある。

4.2. 学習者の語彙レベルの計測

現在までのところ、学習者の到達語彙レベルを測る標準化された語彙力判定テストはないようである。望月[19]、投野他[12:151]によると、現在、考案されているテストに関して、「語彙力の測定は語彙サイズと語彙知識の深さの2面からなされるべきものであるが現時点では両者を同時に測定するテストは開発されていない」(望月[19:28])といわれている。前者の語彙サイズを測る代表的なテストにはNationによるVocabulary Levels Testがある[31]。このテストの対象者は2,000語レベル以上であるので、初級学習者の語彙レベル測定には適さない。後者の語彙知識の深さを測るテストも開発されているが、上級レベルに達しない日本人英語学習者には、語彙サイズのテストの方がすぐれていると考えられている(望月[19:28])

初級学習者から中級学習者の語彙力判定の試みとして、竹蓋他[9:4]は「Keyword System 5000」(竹蓋1994)の認知レベル1~5の5,000語データベースからランダムに抽出したサンプル語を利用して学習者の語彙力を推定している。上記語彙リストは公開されているが、テスト自体は公開されていない。

4.3. 学習教材の語彙レベルの計測

難易度が「易」から「難」レベルまで広範囲にわたる英文素材の語彙レベルを計測するには、比較基準として十分なサンプル量に基づいた大規模な語彙リストが必要となる。100万語のコーパスでは出現頻度7,000位になると、語の出現回数は10回程度に激減してしまい、十分な資料が得られない。その点、2000年にEU圏外にも公開された1億語の電子コーパスであるBritish National Corpus (BNC)の場合、出現頻度38,000位で語の出現回数が100回あり、現在使用可能なコーパスの中で最も詳細な資料が得られるものと言える。

このBNCの出現頻度順リストを尺度に利用して中條[10]、中條他[11]では入試問題、各種テスト、英文雑誌、新聞、ニュース、学校教科書、英文論文等の語彙レベルを計測している。現在、広範囲にわたる英文素材の語彙レベルを自動的に計測できるシステムの開発中である。

平易な英文素材の場合は、イー・キャストが教科書語彙データベースを利用して、中1レベルからセンター試験レベルまで6段階のレベルチェックを行なうソフトウェアを発売している。

4.4. 学習者のレベルに応じた語彙学習支援システムの可能性

学習者の語彙レベルと学習教材の語彙レベルの両者が共有データベースを使用して自動的に計測可能になれば、学習者の到達語彙レベルを超える学習教材の語彙を支援するシステムが可能になる。そのシステムを考える際には、瑣末なように思えるかもしれないが、固有名詞・数字等の処理、lemmatizeできない語の処理をどうするか、といった問題点を解決する必要がある。

5. 学習資源としてのコーパス

5.1. コーパスを用いた英作文支援

大規模なテキストデータ(コーパス)は、それ自身、英語学習のための資源として利用できる。たとえば、読売新聞は、日本語記事だけではなく、英文の記事(Daily Yomiuri)も研究用に利用でき、1年分ごとのCD-ROMが市販されている。通信総合研究所では、読売新聞とThe Daily Yomiuriを用いて、18万対の日英文対応と、約10万対の日英記事対応(95,000対)のデータを作成し、公開している[6]。これは研究利用が可能な日英対訳データとしては最大のものである。日英対応のついた新聞記事(コーパス)を人間にとって有効な情報源として用いて、私たちが英語の文章を書く場合に有効な情報を得ることが出来る。ここでは「いかがなものか」という辞書では調べられないような表現を、まず日本語記事から検索し、得られた日本語文と、それに対応する英語文から、それぞれ日本語と英語の特徴語を表示する。日本語の特徴語から「いかが」を、英語の特徴語から「I」を選んで絞込検索(双方の語を含むようなものだけを表示する)を行うことにより、「いかがなものか」の訳として「I doubt」という表現を候補として得ることができる。このような手法を用いれば、最近の新聞記事を通して、新しい単語や表現の適切な訳を選択することが可能となる。

5.2. 学習者コーパス

大規模な英語学習者コーパスとしては、通信・放送機構と通信総合研究所が作成した、ACTFL/OPIに準拠したインタビュー形式の英語能力判定テストであるSST(Standard Speaking Test)のインタビューを書き起こしたSST Corpusがある。日本人の被験者1200名(300時間)の英語によるインタビューを書き起こし、言い直しやフィルター等の情報、誤り情報を付与しており、世界最大級の学習者コーパスとなっている。通信総合研究所では、このコーパスに誤りタグを付与することにより、日本人の英語発話の誤り分析を行う[3]と共に、誤りの自動検出[7]、被験者の英語習熟度[35]の自動判定の研究を行っている。また、補助的なコーパスとして、母語の干渉を調べるため、このコーパスを日本語に翻訳した「日本語コーパス」と、母語話者の発話との比較のための「正解コーパス」を作成している。

6. コーパスの普及と活用技術の向上を目指して

6.1. コーパスの普及

従来、コーパスは一部の研究者や辞典執筆者だけが使用する特殊な言語資源と考えられてきた。現在、インターネットの普及によって、質や内容さえ問わなければ、誰もが簡単に外国語のテキストデータを集めて二次加工できる環境となっている。しかし、その一方で、信頼性のある BNC や BankOfEnglish のようなバランスコーパスこそ、インターネット検索と同じように簡単に利用できないかというニーズが高まっている。また、ESP 教育に対応して、科学技術論文や専門誌をソースにした ESP コーパス([33], [32])の需要も非常に高い。こうした背景から、小学館では、中学、高校の英語教育者から、一般の英語学習者まで、目的に応じて簡単に使えるコーパスの検索サービスを本年度中に実施する予定である。文法問題の作成、辞書では確認できない語法の調査、CALL 教材とコーパスとの連動など、コーパスの利用範囲は非常に広いと考えられるため、様々な場面に応じたコーパスの利用方法をサービスサイト上で啓蒙してゆく。

6.2. 活用技術の向上

コーパスの普及と同様に、コーパスの活用技術の向上も求められている。注意深く設計され収集されたコーパスは、言語現象、言語活動の多様性を良く再現していると考えられるため、語彙の頻度情報のような単純なものから、動詞のサブカテゴリの頻度分布など、コーパスから得られる二次的な言語知識の活用が広く期待されている。そもそもコーパスが開発された直接の目的は、個別の言語研究よりも語学教材と辞典の開発であった。コーパス開発から辞典執筆の過程は、辞典という知識データベースに言語素材をコンパイルするデータドリブンなフローとして捉えなおすことが可能である。コーパスに別種の言語知識や自然言語処理の技術を適用させたコーパスマイニング[30] は、今後様々な言語上の知見をもたらしてくれるだろうし、そこで得られる二次的な言語知識自体の公開サービスも必要であろう。

このようにコーパスの普及と活用技術の向上は、外国語学習に携わる人々に役立つ情報基盤に不可欠な要件となるだろう。

7. Global Document Annotation とその応用

Global Document Annotation (GDA) [14]は XML のインスタンスであり、タグの構造とそれらの属性によって、文書の統語や意味に関する構造を明示するための枠組である。具体的には係り受けや省略の補完、語義、スコープ、照応・共参照、発話行為などを明示することができる。

文書群に対して予めこのような情報を付与しておく、質問の構造と文書中で合致した各部分の構造の両方を考慮することで、システムは単にシソーラスを引いて類義語を提示するだけではなく、類義語の優先度を決められるようになる[34]。例えば、「住宅を安く作る」という質問に対して「作る」の類義語を提示することを考える。質問グラフ

において「作る」は「住宅」と隣接していることから、「作る」の類義語の中でも文書中で対応する部分グラフにおいて「住宅」と隣接する語の優先度を高くする。このようにすれば、文書中で「住宅を建設する」のような部分が存在すれば、「建設」の優先度が高くなり、より適切な類義語を提示できる。

また要約についても、変換後のグラフに対して活性拡散を行なうことによって重要な要素を見つけ出し、それらを含むようなできるだけ小さい部分グラフから文を生成するという方法で高品質な要約を行なうことができる[6]。活性拡散を含めて何らかの方法で文書中の各要素の重要度を計算して重要度の高いものだけを残す、という方法で文書を要約する方法はすでに様々なものが実現されているが、単に重要度の高いものだけを残すだけでは意味の通る文章にはならない。先行詞が削除された代名詞をもとの名詞と置き換えたり、主節が削除された従属節の時制や相を整えたりといった細かい後処理もタグで付与された情報を使えば可能である。

現在、新聞記事の他に国語辞典の語釈文や和英辞典の例文に対してアノテーションを行なっている。とくに和英辞典の例文に対しては、見出し語と対応する語句および対訳文間での語句の対応関係を明示したタグを付与している。これらのデータと上記の検索や要約の技術を用いることで、与えられた語に対する訳語がどのように使い分けられているかを網羅的に調べたり、原文と構造を指定した部分的な翻訳から残りの部分を自動的に生成するといった、簡単な英作文の学習支援や問題の半自動生成を行なうことが可能となる。

8. 日英 2ヶ国語音声化システム BEP と語学学習

8.1. 日英 2ヶ国語音声化システム BEP

BEP (Bilingual Emacspeak Platform) [37, 22] は、日本の視覚障害者用の日英 2ヶ国語音声化システムであり、Unix で標準的に使われている高機能エディタ Emacs を音声化する。BEP は、米国の全盲の視覚障害者である Raman が自身の学位論文執筆に使えるシステムとして開発した Emacspeak を基にしており、(1) 英語以外に日本語も扱える、(2) Linux と Windows の両 OS で使える、(3) テキストエディタである Emacs の上で、電子メール、Web ブラウズ、プログラム開発、辞書検索、文書作成などの様々な機能を利用できる、といった特徴を持つ。

BEP は日本人用の日英 2ヶ国語音声化システムとして、以下の 3つのバイリンガル音声化モードを持っている[4]。

(1) ネイティブ英語モード: ASCII 文字はすべて英語と判定し、英語の音声合成を用いて読み上げる。1文中に英語と日本語が混在していた場合も、その都度音声合成の言語を切り変えて正しい発音で読み上げる。このモードなら、'r' と 'r', "cup" と "cap" の違いも聞き分けることができるし、イントネーションも正しい。

(2) 混在モード: 一定の長さ(デフォルト 40 文字)以下の ASCII 文字の連続は日本語と判定し、カタカナ読み(注;

英単語見出しに対してカタカナで発音を表記した辞書を用意し、この辞書を用いて英語をカタカナに変換してから読みあげる)で日本語の音声合成に送る。日本語文中に現れる短い英文字列は日本語の一部として読み上げる方が自然と思われる場合が多いので、日本人用のバイリンガルシステムとして、このモードをデフォルトとしている。

(3) すべてカタカナモード: ASCII 文字列もすべてカタカナ読みとして日本語の音声合成で出力する。プログラミングなど特定の環境では、英数字のみで書かれた情報でも、すべてカタカナ英語で読み上げるほうが日本人にとってわかりやすい。

プログラム開発モードは「すべてカタカナモード」、Webブラウザは「ネイティブ英語モード」、というように、読み上げる情報のコンテキストに応じてバイリンガル音声化モードを切り替えることができる。また、コマンドひとつで、いつでもモードを切り替えることができる。

8.2. 文字情報の音声利用

目で見ることができない状況では、コンピュータが持つ情報を読み上げさせて耳で聞くことになる。音声出力には、(1) 眺望感がなく、望む情報がいつ出力されるのかを最初にざっと眺めて知ることができないためにその情報が出力されるまで待たなければいけない、(2) 再生された音は消えてしまい、聞き返すことができない、(3) 目で読むよりも耳で聞くほうが約2倍の時間がかかる、といった特徴があるので、視覚障害者用の音声化システムには以下のような工夫が必要である[23]。

8.2.1. 音声出力における文字種の区別などの工夫

コンピュータが持っている文字情報を正確に効率的に聴覚を通して取得したい場合、大文字や小文字などの文字の種類や句読点などがどこにあるかなども知りたい場合がある。

文字種を英語だけに限っても、句読点や特殊記号を読むかどうか、あるいはどの程度まで記号類を読むかなどの punctuation mode を指定できる必要がある。また空白文字を指摘できる機能、短い言葉やピーブ音や声の高低や種類を変えることなどで大文字を小文字と区別して読み上げる機能、文末で少し間を置いたり短い音を出したりする機能も必要である。その他にも単語として読むだけでなくスペルアウトして読み上げる機能、alpha などのようにフォネティックコードでアルファベットを読む機能、同じ記号が連続しているときは数に変換して読み上げる機能なども必要である。

8.2.2. 漢字音声化の工夫

日本語にはさらに多くの機能が必要となる。

(1) 同音異義語が多い漢字の読み上げには3種類のモードが必要である。

a) 詳細読み; 漢字を音声で識別するために「タンボの田」などのユニークな説明つきで読み上げる機能。カナ漢字変換時に必須。

b) 簡易読み; カーソルを移動する際の一文字読み上げ時などにできるだけ簡単に漢字を音声表示する機能。音読み(や訓読み)をする。

c) 滑らか読み; 文脈の中で構文解析をして正しく読む機能。現在は音声合成ライブラリや音声合成装置がこの機能を担ってくれるので視覚障害者用音声化システムとしては特に仕事をしなくて良い。

(2) カタカナを音声で識別する機能も必須である。声のピッチまたは男女を変えてカタカナを識別している場合が多い。

(3) ローマ字、数字、カタカナの場合、全角と半角を区別する必要がある。声の高さで識別するなどの工夫がなされているものもあるが、まだ一般的ではない。

(4) ローマ字はアルファベットで書かれているが英語の音声合成では正しい発音にならないので正しい発音に変換する必要がある。Windows で使われている音声合成ライブラリなどにはこの機能がないので、音声出力システムの中であらかじめ正しい発音のカタカナ文字に変換した上で音声合成ライブラリに読み上げさせる必要がある。

(5) 日本人であっても英語を扱う機会は多い。英語を正しく発音する機能も必要である。

8.2.3. その他

カレンダーのような2次元の表を聴覚に適した形で表現するためには、今表示しているのが2次元の構造を持ったデータであることを音声化システムが知り、2次元構造を1次元の音声ストリームに適した形に加工してから出力する仕組みが必要である。つまり、(1) データ以外に、データの性質を示すためのデータ、(2) データの性質に応じてデータを加工する機能、が必要になる。

画面表示に使うフォントにいろいろな種類や大きさがあるように、音声出力に使う音声にも Voice フォントという概念がある。いろいろな声、大きさ、速度、ピッチ、イントネーション、を持った声を使ってデータの性質を示すことで、耳で聞いたときにわかりやすくなることが期待できる。

8.3. BEP と語学学習

BEP のベースとなっている Emacs は、Emacs Lisp というプログラミング言語が動作するプラットフォームでもある。Lisp はもともと人工知能の研究用に開発された言語であるので、Emacs にはテキストを走査して解析する機能が豊富にある。Emacs には心理カウンセラー Eliza という、ユーザの入力に応じて会話をしているかのような振る舞いをする人口無能のプログラムもある。BEP 自体もこの Emacs Lisp で記述されており、Emacs Lisp が識別したテキストの言語属性を利用してバイリンガルの音声化を実現している。テキスト処理に優れている Emacs の特長を活かせば、Emacs をテキスト情報を処理するエンジンとしても利用できる。BEP に音声入力機能を追加すれば、音声ベースで利用できるインタラクティブなテキスト処理システムになる。そのような応用のひとつとして、語学の学習ソフトが考えられる。

BEP を英語学習に応用しようとした例として、飛岡氏の修士論文 [13]がある。この研究は Emacs を使うところまでは実装できなかったが、耳で聞いたり声に出したりするこ

とが重要な語学学習の自習システムに音声入出力インターフェースを利用する可能性を示している。すなわち学習者は、合成音声で出力された英語を聞いて Listening の練習をし、学習者が発声した英語を音声認識させることで、音声認識できるという意味で正しい発音で話す練習ができる。このシステムのバックエンドとして BEP を利用すれば、日本語と英語の両方を学習者に適した形で音声化できるし、音声認識の結果を元に正誤を判別し、それを元に新たな問題文や課題を提示することが簡単にできる。

また Eliza に音声インターフェースを追加すれば、心理カウンセラーと英会話をしているかのようなプログラムを作ることでもできる。

9. 教員のニーズに基づく学習履歴の分析と表示

近年、e-Learning システムが隆盛の時期を迎えている。それに合わせて所謂、LMS (Learning Management System) が各社から競うように開発・販売されている。ただ、多くの LMS 製品を観察した後、ひとつ共通点を上げるとすれば、それは“学習履歴表示機能・統計機能が貧弱であると言わざるを得ない”点である。学習者ごとの正解率分布が棒グラフや折れ線グラフなどで綺麗に表示されたり、学習者全体の平均点などが組み込まれている製品は多い。一方で、ちょうどその頃宮崎は CALL 用英語学習ソフトを試作しており、英語教員から“自ら教材を開発できるように”や“学習者の履歴情報が解るように”などの要望を試作ソフトに対して受けていた。しかも(前者については本発表の主旨から外れるため割愛する)、後者については、英語教師が本当に欲しい学習履歴情報というのは、上述 LMS の成績表示機能にあるようなレベルの物だけではなく、より細かく奥深いものであることを実感した。こういった背景のもと、教員の立場として見た“かゆいところに手が届く”学習履歴の収集・加工・表示法について考案し、CALL 用英語学習ソフトを現在開発している(参考文献[8, 16, 17, 18]を参照のこと)。たとえば

各種 criteria に対するランキング (○問題別正解率 ○平均解答時間 ○テスト実施回数 ○学習日数 ○(正解率によって A,B,C などにレベル分けした) ランク分布 ○満点獲得回数 ○正解率伸び幅) 機能、また学習者の成績を総合することによる各問題の難易度(加えてランキング)の計算、学習者の誤解答の全リスト表示(または個人ごと表示)、個人別成績結果と問題別成績結果間の密なるリンク付け、(学習者の階層ごとの) クラスター分析、問題を解く際に学習者の挙動を示唆する解答軌跡などがある(太字は実装済み)。最終的にはこれらをモジュール化し、CALL システムあるいは LMS の学習履歴機能に組み込んでいくことが最終ゴールである。

参考文献

- [1] イー・キャスト「CD-ROM 単語レベルチェック Ver.4.0」
- [2] 井佐原均・投野由紀夫・平野琢也, 2000, 「日本人学習者のレベル別英語発話コーパスの作成」, 言語処理学会 第6回 年次大会(NLP2000).
- [3] 和泉絵美(CRL)・齋賀豊美・Thepchai Supnithi・内元清貴(CRL)・井佐原均, 「エラータグ付き日本人英語学習者発話コーパスを用いた学習者の冠詞習得傾向の分析」, 第9回年次大会, 言語処理学会, 2003年.
- [4] 井上浩一・切明政憲・渡辺隆行, 「日英2カ国語 Emacs 音声化システム (Bilingual Emacspeak Platform)」, Linux Conference 2002, (2002年9月18-20日, 東京).(最優秀プレゼンテーション賞受賞)
- [5] 内山将夫・橋田浩一, 「GDA タグを利用した複数文書の要約」, 第6回年次大会発表論文集, pp. 376-379, 言語処理学会, 2000年3月.
- [6] 内山将夫・井佐原均, 「日英新聞記事の対応付けと精度評価」, 情報処理学会研究報告 2002-NL-151, 2002.
- [7] 齋賀豊美・和泉絵美(CRL)・Thepchai Supnithi・内元清貴(CRL)・井佐原均, 「英語学習者の発話における誤りの検出」, 第9回年次大会, 言語処理学会, 2003年.
- [8] 芹澤正樹・宮崎佳典・南紀子, 「CALL 用英語学習プログラムに対する学習履歴分析ツールの開発」, 情報処理学会第65回全国大会論文集, pp. (4) 411-412, 2003年3月
- [9] 竹蓋幸生他, 「外国語学部における英語教育改善の試み」, 『文京学院大学外国語学部・文京学院短期大学紀要』第2号, pp. 1-13, 2003年.
- [10] 中條清美, 「高校英語教科書・大学入試問題・大学英語教科書・英語資格試験に使用される語彙レベル測定の試み」, 『第35回日本大学生産工学部学術講演会 教養・基礎科学部会講演概要』pp. 41-42, 2002年.
- [11] 中條清美・長谷川修治, 「BNC (British National Corpus) を利用した時事英語教材語彙の難易度調査」, 日本時事英語学会第44回年次大会, 大阪府立大学, 2002年.
- [12] 投野由紀夫他, 『英語語彙習得論』, 河原社, 1997年.
- [13] 飛岡正人, 「モバイル端末との音声対話による利用を想定した外国語学習システム」, 慶應義塾大学大学院政策・メディア研究科 修士論文, 2000年.
- [14] 橋田浩一. 大域文書修飾. <http://www.i-content.org/gda/>.
- [15] 原田康也, 「電話を利用した英語リスニング・スピーキング自動テスト: 早稲田大学法学部1年生のスコアからの考察」, 電子情報通信学会技術報告(信学技報) TL2002-41, pp.49-54, 電子情報通信学会, 2002年12月6日.
- [16] 南紀子・宮崎佳典, 「カスタマイズ可能な英語学習プログラムの開発」, LET 第41回全国研究大会論文集, pp. 184-187, 2001年8月.
- [17] 宮崎佳典・南紀子, 「CALL 作用問機能付英語学習ソフト及び学習履歴分析ツールの開発」, 情報処理教育研究集会平成14年度講演論文集, pp. 616-619, 2002年10月.
- [18] 宮崎佳典, 「CALL 教室用英語学習および英語教材作成プログラムの開発」, 第8回英語教育をおもしろくする会, 静岡産業大学, 2002年4月.
- [19] 望月正道, 「日本人英語学習者のための語彙サイズテスト」, The IRLT Bulletin, Vol.12, pp.27-53, 1998.
- [20] 山田恒夫・足立隆弘・ATR 人間情報通信研究所, 1999, 「英語スピーキング科学的上達法」, 講談社.
- [21] 山田恒夫・足立隆弘・ATR 人間情報通信研究所, 1998, 「英語リスニング科学的上達法」, 講談社.
- [22] 渡辺隆行・井上浩一・切明政憲, 「視覚障害者用日英2ヶ国語音声化システム BEP」, 電子情報通信学会思考と言語研究会, 信学技法 Vol. 102, No. 254, pp.15-22, 2002年7月.

- [23] 渡辺隆行, 「視覚障害者のコンピュータ利用と日英2ヶ国語音声化システム BEP」, 情報処理学会誌, 43/8 pp.873-879, 2002 年.
- [24] Alderson, J.C., 'Reading in a foreign language: A reading problem or a language problem?' In J.C. Alderson & A.H. Urquhart (Eds.), *Reading in a foreign language*, New York: Longman, 1984.
- [25] ATR 人間情報通信研究所編, 「ATR CALL 完全版英語スピーキング科学的上達法音韻編」, 講談社, 2000.
- [26] ATR 人間情報通信研究所編, 「ATR CALL 完全版英語リスニング科学的上達法音韻編」, 講談社, 1999.
- [27] Bernstein, J. & Harada, Y., "Automatic Measurement of Spoken English Skills: consistent benchmarks for English learning," 大学英語教育学会第41回全国大会, 2002.
- [28] Bernstein, J. & Townsend, B., "Computer Estimation of Spoken Language Skills," in *Proceedings of the 14th Pacific Asia Conference on Language, Information and Computation*, PACLIC 14 Organizing Committee. 2000,
- [29] Huckin, T. & Bloch, J., 'Strategies for Inferring Word-Meanings in Context: A Cognitive Model,' In Huckin, T. et al. (Eds.) *Second language, reading and vocabulary acquisition*, pp.153-180, N.J.: Norwood, 1993.
- [30] Nakamura, T. and Tono, Y., 'Lexical Profiling using Shogakukan Language Tool Box,' ASIALEX, Meikai University, 2003 (To appear).
- [31] Nation, P., *Teaching and Learning Vocabulary*, Heinle and Heinle: Boston, 1990.
- [32] Noguchi, J., Orr, T., Tono, Y. and Nakamura, T., 'Using a dedicated corpus to identify features of professional English usage,' CL2003, Lancaster University, 2003.
- [33] PERC, <http://www.perc21.org>, 2002.
- [34] Miyata, T. and Hasida, K., 'Information retrieval system based on graph matching,' in *Workshop on Knowledge Transformation in Semantic Web (ECAI2002)*, p. 109, July 2002.
- [35] Thepchai Supnithi・齋賀豊美・和泉絵美(CRL)・内元清貴(CRL)・井佐原均, 「英語学習者発話コーパスを用いた習熟度判定」, 言語処理学会第9回年次大会, 2003.
- [36] Ulijn, J., 'Conceptual and syntactic strategies in reading a foreign language,' in E. Hopkins & R. Grotjahn (Eds.), *Studies in language teaching and language acquisition*, pp.129-166, Bochum: Brockmeyer, 1981.
- [37] Watanabe, T., Inoue, K., Sakamoto, M. and Kiriake, M., "BEP: a practical bilingual speech synthesis system for Japanese," *Proceedings of Pacific Association for Computational Linguistics*, pp.307-314, 2001.