

日本語会話データベースの構築と談話分析[†]

— 中間報告 —

上村隆一
福岡工業大学
uemura@ipc.fit.ac.jp

田吹昌俊
九州共立大学
tabuki@orion.kyukyo-u.ac.jp

村野良子
国際基督教大学
murano@icu.ac.jp

近年、わが国の言語研究の分野においても、計量的分析だけでなく、データ収集および編集の道具としてのコンピュータ利用が増加してきた。一方、欧米では大規模な言語データベース（以下コーパス）を用いた語彙・文法研究、文学作品、歴史資料など特定の分野の言語資料の電子化テキスト作成が急速に進んでいるが、わが国の日本語に関する電子化テキスト等のプロジェクトは残念ながら立ち遅れた感じが否めない。

本稿では、そうした現状を踏まえながら、われわれの日本語会話コーパス構築プロジェクト（平成7年度文部省科学研究費補助重点領域研究「人文科学とコンピュータ」公募研究（課題番号07207124））について、とくに談話分析への応用の観点から研究成果を中間報告する。

Corpus Project of Spoken Japanese and Discourse Analysis

Ryuichi Uemura
Fukuoka Institute of Technology
uemura@ipc.fit.ac.jp

Masatoshi Tabuki
Kyushu Kyoritsu University
tabuki@orion.kyukyo-u.ac.jp

Ryoko Murano
International Christian University
murano@icu.ac.jp

The object of our joint project started in 1991 is to create an original hyper-media corpus of spoken Japanese. We have collected 'live' spoken data from actual conversation between experts of teaching Japanese as a foreign/second language and native / non-native speakers of Japanese, based upon a testing format known as OPI (Oral Proficiency Interview). The whole contents of experimentation recorded on high precision video tapes and magneto-optical disks were converted to digital video and/or sound data files. The sample version of our corpus is now being transferred to WWW server on our campus (URL: <http://corpus.fit.ac.jp>) with HTML-tagged texts and is expected to be available over the Internet. This project is authorized and sponsored by Japanese Ministry of Education as one of the 1995 research programs on priority areas (Project No. 07207124).

1 はじめに

本研究は、日本語母国語話者(以下NS)と非母国語話者(以下NNS)の現実発話に含まれる言い誤りの類型を比較分析することを目的として、1991年度より開始した試験研究の延長線上にある。研究当初から、分析対象となる一次言語資料の絶対量不足を痛感したため、われわれはまず、インタビュー実験形式による会話データの収集と、それに基づくコーパスの構築作業から開始することにした。

2 データ収録の実施と処理方法

本年度の研究計画では主としてNNSの会話データの収録を行うことにした。東京都内の民間日本語学校、国際基督教大学、およびアメリカ合衆国のプリンストン大学において計3回のデータ収録実験を実施し約70名分の会話データを得た。(被験者の内訳は表1のとおり。)

表1 インタビュー実験被験者の内訳

国籍	米国 26 韓国 25 中国 5 日本 3 ロシア・オーストラリア 各2 ドイツ・オーストリア・タイ 各1
性別	男 29 女 37
年齢層	20代 58 30代 7 40代以上 1

ここでは日本国内において行った実験について報告する。

会話収録の実施場所と実施時は次の通りである。

第1回 東京都内の民間日本語学校

7月1日、2日

第2回 国際基督教大学

7月20日、21日、31日

実施会場を2か所としたのは、できるだけ様々な背景を持った被験者を得るためであり、日本語学校では主にアジア系、国際基督教大学では欧米系の被験者を期待していた。

2.1 被験者の属性

第1回の日本語学校での収録の被験者は18名で、内一名を除いて全員が韓国籍である。年齢は22歳が8名と最も多く、20～26歳までが18名中16名を占める。大半が就学生で、日本語学習歴、日本滞在期間、生活環境も似通っている極めて同質性の高いグループであった。

第2回目は国際基督教大学が夏に開いている夏期日本語講座の受講生を被験者とした。そのため被験者の国籍は7か国に渡り、母語や言語背景も多彩であった。母語が韓国語で米国籍の韓国系アメリカ人6名、日系アメリカ人1名、米国育ちの日本人1名もこの中に含まれる。大学などでの専攻や学習歴などもヴァリエーションに富んでおり(表3、表4)、多様性と言う観点からは、一ヶ所で得ることができる被験者のプロフィールとしては、かなり多様なデータを収集することができたのではない。この傾向はプリンストン大学についても同様で、欧

米系とアジア系がほぼ同数であった。年齢層は20歳代に偏っていたが、男女比、教育歴などは適当に分散していたと思われる。

2.2 収録に当たった確認事項

会話データの収録に際しては、事前にこの研究の趣旨とリスクに関して説明を行い、第1回目は口頭で同意を得、2回目は同意書に署名してもらった。確認事項は次の通りである。

1. この会話データ収録は15～30分の日本語によるインタビューである。
2. このインタビューは録音、録画される。
3. このデータは学術研究の目的以外には使用されない。
4. このデータはコンピュータネットワーク上で公開される。
5. 4に関しては現時点で予測不可能な犯罪行為を含むリスクについては、研究者およびインタビュアーは一切責任を負わない。

個人データの公開に伴うリスクにの可能性については、事前に被験者の了解を明確な形で得ておくことが不可欠であろう。実際にこの時点で被験者になることを拒否した応募者が一名あった。

2.3 会話データの形式と今回の問題点

会話データはすべてACTFL(全米外国語教師協議会)のOPIと呼ばれる会話能力試験の形式に則って収録した。これは会話モードとロールプレイモードの二つの要素によって構成されることが特徴である。もともと会話能力の評価を目的にした方法であるため、他の多くのスピーチデータの採集手法と異なり、インタビュアーはできるだけ被験者に話をさせるように配慮する。従ってデータ収録に当たっては、上記のような研究目的に関する説明に加えて、ACTFLのOPIの形式についての若干の説明を行う必要がある。OPIの一般的な注意事項としては次の2点が重要である。

1. この会話では、日本語を使ってできるだけたくさん話すようにする。
2. 個人的なことに関する質問や、答えたくない質問には答えたくないと言ってもよい。

今回の収録データには大きく分けて二つの問題点がある。一つは被験者のヴァリエーションについてである。既に述べたように、日本語学校で収録した会話データは同質性が際だっていた。日本語学校ではアジア系被験者を集めることが目的であったが、現在東京都内の多くの日本語学校では韓国の学習者が圧倒的に多い事実を考えれば、場所の選択に問題があったと思われる。

インドネシア、タイ、香港、フィリピン、中国などからの学習者を被験者に加えることができる場所を設定することが課題の一つであろう。

もう一つは会話データの形式である。スピーチサンプルがある一定の形式にそって行われたという点では今回の試みは画期的であるといえる。しかし実際には、15分程度から30分までインタビューの長さにかんがりの違いが見られる。評価という側面を考慮に入れる場合にはやむを得ない場合もあるが、

データの公開という観点からみれば、やはり統一したほうがよいだろう。

2.4 データ処理

データ記録媒体については、画像データを高画質8ミリビデオテープに、音声データを光磁気ディスク(MD, デジタル録音専用メディア)にそれぞれ収録した。収録時間は被験者1人につき20-30分程度である。現在までに、音声データからテキストデータへの書き起こし作業の一部(約10名分)が完成し、圧縮した音声データとともにインターネットのWWWサーバ(<http://corpus.fit.ac.jp>)およびFTPサーバ上で順次公開している。

特に、音声データの公開方法に関しては、最近インターネット上でのオン・デマンド型音声サーバ技術として注目されているReal Audio Serverをいち早く導入し、本年9月から実験的にインタビューの内容の一部(フリートーキングとロールプレイについて各々約20人分)を転写テキストと一体化した形で提供している。デジタル変換された画像データ(動画)は被験者1人当り200~300MB(圧縮ファイル)に達するので、現時点ではインターネット上では公開せず、追記型光ディスク(CD-R)に保存している。

3 日本語教育からみた本研究の意義

本研究プロジェクトが日本語教育の現場に与えるインパクトは限りない。その一つを挙げれば、本年度実施した2回の会話データをみるだけでも、第二言語の習得に関わる諸要因についての多くの示唆を読み取ることができるだろう。(転写テキストの具体例は本稿末尾のサンプル・データを参照。)さらに数多くのデータが得られれば、個々の要因についての研究も可能になることが期待される。さらに会話データの中のロールプレイモードのヴァリエーションは、社会的文化的に適切な言語使用を教師がどのように指導すべきなのかという問題に対する解決の糸口を与えられる。次年度以降に予定している、NSの会話データ収集が進めば、現在多くの日本語教師が持っている社会的・文化的に適切な言語使用のイメージの見直しを迫られることになるかもしれない。

4 今後の研究計画

1996年度は本年度に引き続いて、会話データの収集、テキスト転写作業を行う。また、情報処理、日本語教育関連の学会において、本研究の第2次中間報告を発表する。実験データの収録作業については、同年度内に次の2点を実行する。

- 1) 米国在住の日本人研究協力者(牧野成一プリンストン大学教授)の指導の下に、国内と同一の実験条件で現地在住のNS, NNSそれぞれの会話データを収録。
- 2) 研究分担者と国内2大学の協力の下に、主としてNSの会話データを収録。

音声データについては、オーディオテープ(DATを含む)やMDに録音した場合、データベース作成

時に膨大な量の変換作業を必要とするため、別途ノートブック型パソコンと16ビットサンプリング可能な音声入力・編集ソフトを用いて直接デジタル録音を試みる。このことにより、書き起こし(転写)作業から音声データベース部分の構築に至る作業工程を大幅に短縮することができる。

実験で収録したデータは、日本語教育関係の研究分担者が大学院生等に委託した形で転写作業を行った後、責任を持って校閲する。転写作業にあたっては、16ビットサンプリングによりパソコン上で直接デジタル録音したファイルから、デジタル録音・編集機を用いて適宜分割・再編集した音声データを用いる。

1997年度は研究班をコーパス作成班とコーパス分析班に分けて、適宜協力しながら研究計画を遂行する。まず、「作成班」は前年度内に収集した音声データと転写テキストデータについて、任意の文字列から当該個所の音声データを検索するプログラムの開発を試みる。特に、時系列データを扱う検索用言語としては、HyTime処理系等を参考にしながら、SGML/DTDの拡張を考える。同時に、デジタル動画データ(MPEG形式)の検索方法も検討する。なお、コーパス本体は上記のインターネット上で提供するネットワーク版と、CD-ROMで提供するスタンドアロン版の2種類を作成する予定である。また、CD-ROM版の検索ソフトウェアの更新及び追加情報の提供等はすべてインターネットを利用し、国内外の言語研究者および日本語教育関係者に公開する。

次に、「分析班」は言語学と日本語教育の立場から、NSとNNSの発話内容を比較検討し、段落形成、連接名詞、代名詞類、繋ぎ語等の各トピックについて会話分析を試みる。データの蓄積が一定レベルに達した後、当初の研究目的であった「言い誤り」の類型に関する分析を開始する。年度内に本研究プロジェクトの最終報告を行うが、論文とデータは通常の印刷物に加えて、電子化テキストの形式で作成し、インターネット上のFTPサーバからも利用可能にする。

5 繋ぎ語の分析

5.1 分析の目的

繋ぎ語 (connected words) は言語を問わず談話の節々に確実に現れ、会話の流れを円滑にしたり制御するなどの様々な機能を持つため、日常会話においては必要不可欠でありまた多用される。繋ぎ語自体は個別に意味を持たない場合が多いが、談話の各単位を確実にマークするため、discourse markerと呼ばれたりもする。会話全体の構成をみる場合、談話のユニットがどのような種類の繋ぎ語で結合されているかをみることで談話構成を知る手がかりとなるので、談話研究者にとって重要な意味を持つ。また日本語教育の観点からも、日本語学習者は日本語を母語とする人たちの会話の中の繋ぎ語を意識し、習得することにより更に上級へと進むことができる。

本研究では、OPIにより得られたデータを基に、インタビュー形式の中での繋ぎ語の機能的分析を試みた。この場合、会話の管理者たる interviewer の使

う繋ぎ語の種類とその用法、頻度、それに対する会話の受け手の立場に立つintervieweeの使う繋ぎ語の種類、用法、頻度などに研究の焦点は絞られた。更に言うならば、ピッチャー役を演じなければならないinterviewerが、発話を起こすために、どのような繋ぎ語を、どのように使い、またintervieweeの話を受けるために、どのような繋ぎ語をどのように使ったかを観察することが、本研究の主題となろう。

5.2 分析方法

1994年までに取得したデータは、その大半が日本人のtesterと日本人testeeによるものであり、その中から特にnativeどうしによる会話を選び、その中から主要な繋ぎ語の使用回数を調べ、その結果を表1にまとめた。表1では正味約3時間の会話から74000文字(全角)のテキストファイルを利用した。

表2については、1995年の夏と秋に東京で取得した会話データを利用したもので、日本人のtesterと外国人のtesteeによるものである。ここでの会話データは正味約3時間半であり、文字数は88000文字(全角)である。よって表1と表2を合わせると約6時間半の会話から162000文字(全角)のテキストファイルを利用したことになる。

また繋ぎ語の分析にあたっては上記のテキストファイルばかりではなく、それに照応するビデオ及びオーディオテープを言語資料として使用した。繋ぎ語の抽出についてはそれが使用されているコンテキストがきわめて重要になってくるのでサンプリングする際には周辺部も含めて切り取っていくよう心がけた。またテキストファイルだけではわからない微妙な音声の長短や高低については、オーディオテープ、ビデオテープに頻繁に戻らざるを得なかった。

下記の表2及び表3に示している様に、6種類の繋ぎ語を選び、その各々についての使用回数と、testerとtesteeのインタビューとロールプレイでの各々の使用回数を調査した。繋ぎ語については、あくまでも談話の中で何らかの機能的役割を持っているものだけに限定した。また、シリアルに起こる会話の背景音として差し込まれ、発話とoverlapした「うん」「そう」「ええ」「ほんと」等の俗に言う「相づち的発話」は、談話に於ける話者交替(turn-taking)という観点からも実質的な場所を占めているとは考えにくいので、今回の分析からは除外した。

表2 繋ぎ語使用回数(日本人tester・日本人testee)

繋ぎ語	度数	Tester		Testee	
		FT	RP	FT	RP
そうですか	183	176	0	4	3
じゃ(あ)	153	95	32	6	20
なるほど	58	58	0	0	0
そうですね	142	21	5	100	16
やはり/やっぱり	79	5	0	74	0
でも	63	23	2	2	10

合計 678 378 39 212 49

表3 繋ぎ語使用回数(日本人tester・外国人testee)

繋ぎ語	度数	Tester		Testee	
		FT	RP	FT	RP
そうですか	318	257	25	24	12
じゃ(あ)	178	136	16	17	9
なるほど	67	61	0	6	0
そうですね	110	25	76	5	4
やはり/やっぱり	132	33	1	94	4
でも	102	15	7	69	11
合計	907	527	125	215	40

*FT: Free Talking, RP: Role Play

次に各々の繋ぎ語について、機能的な側面から個別に観察してみよう。繋ぎ語の分析についてはそれが使用される状況が極めて重要になってくるので、繋ぎ語が実際使用されている前後の文脈も含めた文の切り出しから始めなければならなかった。ここでは、紙面の都合上、繋ぎ語の使用回数で1位を占めtester側の使用回数でも最高回数を記録した「そうですか」を観察してみよう。

5.3 繋ぎ語「そうですか」の機能分析

表2と表3から、「そうですか」はここで扱った6個の繋ぎ語の中で最も使用回数が多く、しかも会話の中では管理的な役割を果たすtesterが95%以上の圧倒的な比率で使用していることがわかる。

「そうですか」は形式的にならざるをえないインタビュー形式では、testerがtesteeを管理するために多用せざるを得ない繋ぎ語であり、非形式的な状況を想定したロールプレイの中の会話の中ではほとんど使われていないことがわかる。また「そうですか」は「あ」「ああ」などの間投詞と一緒に「あ(あ)、そうですか」という様に使われる場合が大半である。英語では、'Oh, really' ぐらいが、この日本語の繋ぎ語の意味に最も近いだろう。また、この「そうですか」には話し手が相手から意外な情報、もしくは際立った情報をもたらして驚いたという含蓄も持つ。

(1) Continuer

相手(特にinterviewee)の発話が完結途中で差し込まれるもので、当人は話者交替(turn-taking)の意図はなく、相手の話を継続させる機能を持つ。これは実質的な価値を持たない「相づち」だと考えられがちではあるが、話し手に何らかの働きかけをする機能的価値を持っているのでここで扱った。またこの場合のcontinuerはMaynard(1986)と同義である。特にinterviewerによって使用され、相手の話を確認したり、促したりする機能を持つ。

(2) Answer to (Questions)

上昇もしくは水平イントネーションで発せられ、話し相手（特にinterviewee）の発話を確認する目的で使用され、それに対してintervieweeは「はい」「いいえ」で返答する。

(3) Acknowledgement (確認)

(a) 下降イントネーションで発せられ、相手の直前の発話を確認した後、それに関連した質問をする。

(b) interviewerが相手の一連の談話 (a unit of talk) を再確認した後、新たな話題を提供する。

(以下のdialogue sample中の1はtester, 2はtesteeを指す。また、一方の話者の談話中に他方の話者のあいづちなどのbackground speechが存在する場合はカッコ付きで挿入している。)

(1) Continuer の例

1: で、あの、花岡さんはなぜまた、アメリカ史なんかを専攻なさったんですか？

2: そうですねえ...えーと、日本文化をやりたいかったですけど。

1: うん、あ、そうですか。

2: たまたまこの学校は、日本...文化、専門にせ、専攻する、(1: うん。)こ、ゼミがなくて、その一、わたしがやりたい日本文化(1: うん。)っていうのを、考えるときに、やはり、アメリカの存在... (1: うーん。)という、戦後の日本ですね、(1: うん。)日本文化っていうことだったので、アメリカという、の影響のもとに、という方面で、入っていいのかな、と思っていて、(1: うーん。)それがアメリカ史...

上記の「そうですか」のように相手は話し続けている最中に、背景音として差し込まれるのでserialに続く談話の流れとはあまり関係を持たないが、言外の意味は大いに持つ。ここでの「あ、そうですか」は「アメリカ史を専攻した理由」についての質問に対して発話ではなく、その直前の「本当は日本文化をやりたいかった」という相手の発言に対して、意外性を示す目的で相手の発言が完了する前に割り込んでいったのである。

(2) Answer to (Questions) の例

2: あのー、書類の上からいいますとー、就職率は、かなり高いと思います。

1: ああ、そうですか。

2: はーい、多分、都内の大学ではー、女の子の就職率としたら、あの、群を抜いているんじゃないですか。

相手の発言に対しての確認と質問という二つの機能を持つので、受けた側はそれに対して何らかの形で返答する場合が多い。

(3a) Acknowledgement + Question の例

2: ま、近いという、いうので、あのー、やはり、あのー、近い大学ですから、見えますよね。(1: ええー、そうですね。)雰囲気やなんか。(1: ええ、ええ。)ですから、え、こう、入りやすかったっていうのもありますし、(2: なるほどね。)まあ、希望の大学に、入れなかったっていうのもあります

けどね。

1: ああ、希望の大学に入れなかったから。ああ、そうですか。浪人...はなさらない...

2: しませんでした。え、(1: ふん、ふん、ふん。)はい。

上記の会話の場合、1は2のtesteeの発言「希望の大学に入れなかった」を繰り返し「ああ、そうですか」で相手の直前の意外な新情報を再確認した後、さらにそれに関連した次の質問へと移る。もちろんここでも、「ああ、そうですか」は相手の発言に対して、意外性や驚きの意を含んでいる。

(3b) Acknowledgement + New Topic の例

1: あー、それは、あの、ことばに、でいうと、どういう事なんですか？どういうのが、あの、求め、企業が求めている、女の子像なんですか？

2: ええっとですね、(1: はい。)柔軟性があるって、素直なんです。

1: あ、柔軟性があるって、素直。

2: うーん。

1: ああ、言うことないですね。企業としては。

2: うーん、企業としては。

1: ああ、そうですか。

2: そうですね。あの、あんまり、こう、染まっていけないし、自分を出さない。

1: ああ、なるほどね。(2: うん、なんか。)あの、この学校の、卒業生の人は、就職するときは、大体、こう、ずっと働くつもりで、就職なさるんですか、それとも、あの、あの、結婚までの、ある一定時期っていう感じで。

基本的には(3a)の繋ぎ語の機能とは違いはないが、(3a)がローカルな機能しか持っていないのに対して(3b)はもっと大きな談話を管理していることがわかる。即ち、1のtesterは「そうですか」で相手の直前の発言を確認しているのではなく、そのもっと前にでてくる「どうして成蹊大学の女子学生が企業に人気があるか」と話題を取りまどめて次の話題へ移ろうとする。しかし、ここでは2のtesterが前の話題を続けて話そうとするので、1のtesterは次の「ああ、なるほどね」で強引にこの話題を終わらせ、次の話題への展開をはかる。

前が(3a) 後ろが(3b) の例

1: あ、結婚して(2: はい。)らっしゃるんですか。それは失礼...

2: それまでは、わたしも、これ、初めてですね。

1: ああ、そうですか。花岡さんに、お会いになるまでは。(2: ええー、はい。)

ああ、そうですか。で、ナオミさんっていうお名前、片仮名のお名前も、珍しいんですけど、(2: はい。)まあ、珍しいっていうか、芸名には、あると思うんですけども、どういう、由来なんですか？

前の「ああ、そうですか」は直前の「私も初めてだ」に対しての返答であるのに対して、後の「ああ、そうですか」はさらに大きな談話を再確認した後、新たな質問を相手に与えているのである。

表4 1994年までの収録データの内訳

インタビュー形式

A. 日本人のTesterと日本人のTestee (5組)		
	録音時間 (分)	文字数 (全角)
M(F)/H(F)	17.0	10790
H(F)/A(M)	27.0	11581
M(F)/M(F)	25.0	12150
H(F)/F(F)	23.0	11370
F(F)/T(F)	11.4	6054
B. 日本人のTesterと外国人のTestee (1組)		
F(F)/R(F)	26.5	3910

ロールプレイ形式

A. 日本人のTesterと日本人のTestee (4組5話)		
	録音時間 (分)	文字数
M(F)/H(F)	2.3	1225
H(F)/A(M)	4.5	1501
M(F)/M(F)	2.3	1126
	2.2	1174
H(F)/F(F)	2.5	1013
B. 日本人のtesterと外国人のtestee (4組5話)		
F(F)/R(F)	3.3	959
	3.5	1572
P(F)/U(F)	3.2	1000
	2.3	724
	2.3	762

日常会話形式

日本人学生の会話 (4組5話)

23.0 7266

*H(F)/A(M): /の左側が*testerの頭文字で右側が*testeeの頭文字。また () 内のF,Mはそれぞれ女性と男性を指す。

総録音時間: 181.3 (分) 総文字数: 74177 (全角)

表5 1995年度収録データの内訳

インタビュー形式

日本人のTesterと外国人のTestee (10組)		
	録音時間 (分)	文字数 (全角)
(1)F/S	13:50	5487
(2)F/I	17:10	6291
(3)H/C	16:40	6956
(4)H/I	14:20	5125
(5)F/K	12:00	5377
(6)H/N	17:20	7224
(7)Ma/K	27:40	14434
(8)Mu/Ma	21:00	8353
(9)Ma/B	26:30	12515
(10)Mu/Mi	20:15	8474
小計	186:45	80236

ロールプレイ形式

日本人のTesterと日本人のTestee (8組10話)

	録音時間 (分)	文字数
(1)F/S	2:32	1071
(2)F/I	1:20	493
(3)H/C	1:16	511
	1:25	501
(4)H/I	2:37	811
	0:45	285
(5)F/K	2:10	1020
(6)Ma/K	1:08	454
(7)Ma/B	4:05	1941
(8)Mu/Mi	2:24	1058
小計	19:42	8145

総録音時間: 206分27秒 総文字数: 88381 (全角)

†注) 本稿は公開シンポジウム「人文科学とデータベース」において研究代表者が行った講演原稿に大幅な加筆修正を加えたものである。

参考文献

- Armstrong, S.(Ed.): *Using Large Corpora*. MIT Press. (1994).
- Ball, W. J. (1986), *Dictionary of Link Words in English Discourse*, London: Macmillan.
- Brown, G. and G. Yule. *Discourse Analysis*, Cambridge: Cambridge UP. (1983),
- Buck, K. (Ed.) *The ACTFL Oral Proficiency Interview Tester Training Manual*. The American Council on the Teaching of Foreign Languages. (1989).
- Maynard, S. "On back-channel behavior in Japanese and English casual conversation." *Linguistics* 24:1079-1108. (1986),
- Schiffrin, D. (1987), *Discourse Markers*, Cambridge: Cambridge UP.
- Schourup, L. C. and T. Waida. *English Connectives*, Tokyo: Kuroshio Shuppan. (1988),
- Tabuki, M., E. Kizaki, and H. Shimatani. (1990), "Discourse Markers 'Oh' and 'So' in Simplified Speech." *KASELE Bulletin*, 18, 50-58.
- Uemura, R. : "Hypermedia Corpus Project of Japanese Conversation-Interim Report," *Language and Information Processing* 6, pp. 1-5. Fukuoka Institute of Technology. (1995).
- 上村隆一「コーパスによる日本語会話分析—指示詞の使用について」小泉保教授古稀記念論文集『言語学の展望』pp.93-105. 大学書林.(印刷中).
- ホーラー・サトウスキー『日本語の談話の構造分析』くろしお出版. 1993.
- マイナード・K・泉子『会話分析』くろしお出版. 1993.