Japan Association for Language Teaching

Articles
115 大学入試英語問題語彙の難易度と有用性の時代的変化 [A Chronological Study of the Level of Difficulty and the Usability of the English Vocabulary Used in University Entrance Examinations] — 長谷川修治 • 中條清美 • 西垣知佳子 (Hasegawa Shuuj, Chuujyou Kiyomi, Nishigaki Chikako)

135 Motivation for Learning English as a Foreign Language in Japanese Elementary Schools — Junko Matsuzaki Carreira

159 英語リスニング力における学習環境の差異：PC教室と普通教室の比較 [A Study of Different Learning Environments for English Tasks: Comparison of PC-Equipped Classrooms and Regular Classrooms] — 倉本充子 • 吉田信介 • 吉田晴世 (Kuramoto Atsuko, Yoshida Shinsuke, Yoshida Haruyo)

177 The Effect of EFL Students’ Self-monitoring on Class Achievement Test Scores — Akio Suzuki

197 Differences in Reading Strategies Employed by Students Constructing Graphic Organizers and Students Producing Summaries in EFL Reading — Terry Vanderveen

Reviews
207 Teaching English to the World: History, Curriculum, and Practice (George Braine, Ed.) — Reviewed by Michael Lessard-Clouston

210 Practical English Usage (3rd ed.) (Michael Swan) — Reviewed by Fiona MacGregor

212 Spoken English, TESOL and Applied Linguistics (Rebecca Hughes, Ed.) — Reviewed by Cheryl Kirchhoff

215 The Language of Language: Core Concepts in Linguistic Analysis (2nd ed.) (Madalena Cruz-Ferreira and Sunita Ann Abraham) — Reviewed by Byron O’Neill

217 Text, Context, Pretext: Critical Issues in Discourse Analysis (Henry G. Widdowson) — Reviewed by Gerry Lassche

ISSN 0287-2420
¥950
Contents

113 In this Issue
114 From the Editors

Articles
115 大学入試英語問題語彙の難易度と有用性の時代的変化 [A Chronological Study of the Level of Difficulty and the Usability of the English Vocabulary Used in University Entrance Examinations]
長谷川修治・中條清美・西垣知佳子 (Hasegawa Shuuji, Chuujyou Kiyomi, Nishigaki Chikako)
135 Motivation for Learning English as a Foreign Language in Japanese Elementary Schools
Junko Matsuzaki Carreira
159 英語リスニング力における学習環境の差異: PC教室と普通教室の比較 [A Study of Different Learning Environments for English Tasks: Comparison of PC-Equipped Classrooms and Regular Classrooms]
倉本充子・吉田信介・吉田晴世 (Kuramoto Atsuko, Yoshida Shinsuke, Yoshida Haruyo)
177 The Effect of EFL Students’ Self-monitoring on Class Achievement Test Scores
Akio Suzuki
197 Differences in Reading Strategies Employed by Students Constructing Graphic Organizers and Students Producing Summaries in EFL Reading
Terry Vanderveen

Reviews
207 Teaching English to the World: History, Curriculum, and Practice (George Braine, Ed.) — Reviewed by Michael Lessard-Clouston
210 Practical English Usage (3rd ed.) (Michael Swan)
Reviewed by Fiona MacGregor
212 Spoken English, TESOL and Applied Linguistics (Rebecca Hughes, Ed.)
Reviewed by Cheryl Kirchhoff
215 The Language of Language: Core Concepts in Linguistic Analysis (2nd ed.) (Madalena Cruz-Ferreira and Sunita Ann Abraham)
Reviewed by Byron O’Neill
217 Text, Context, Pretext: Critical Issues in Discourse Analysis (Henry G. Widdowson) — Reviewed by Gerry Lassche

JALT Journal Information
220 Information for Contributors (English and Japanese)
All materials in this publication are copyright ©2006 by JALT and their respective authors.
Japan Association for Language Teaching
A Nonprofit Organization

The Japan Association for Language Teaching (JALT) is a nonprofit professional organization dedicated to the improvement of language teaching and learning in Japan. It provides a forum for the exchange of new ideas and techniques and a means of keeping informed about developments in the rapidly changing field of second and foreign language education. Established in 1976, JALT serves an international membership of approximately 3,000 language teachers. There are 34 JALT chapters in Japan, along with 17 special interest groups (SIGs), and one forming chapter. JALT is one of the founders of the Pacific Asian Consortium (PAC), which is an association of language teacher organizations in Pacific Asia. PAC holds regional conferences and exchanges information among its member organizations. JALT is the Japan affiliate of International TESOL (Teachers of English to Speakers of Other Languages) and is a branch of IATEFL (International Association of Teachers of English as a Foreign Language).

JALT publishes JALT Journal, a research journal; The Language Teacher, a monthly periodical containing articles, teaching activities, reviews, and announcements about professional concerns; and JALT International Conference Proceedings.

The JALT International Conference on Language Teaching and Learning and Educational Materials Exposition attracts some 1,600 participants annually and offers over 300 papers, workshops, colloquia, and poster sessions. Each JALT chapter holds local meetings and JALT’s SIGs provide information on specific areas of interest. JALT also sponsors special events such as workshops and conferences on specific themes and awards annual grants for research projects related to language teaching and learning. Membership is open to those interested in language education and includes automatic assignment to the nearest chapter or the chapter you prefer to join, copies of JALT publications, and reduced admission to JALT-sponsored events. JALT members can join as many SIGs as they wish for an annual fee of ¥1,500 per SIG. For information, contact the JALT Central Office or visit the JALT website at <www.jalt.org>.

JALT National Officers, 2006

President: ................. Steve Brown
Vice President: ............. Steve Nishida
Auditor: ...................... Tadashi Ishida
Director of Treasury: ........ Peter Wanner
Acting Director of Records: ... Donna Tatsuki
Director of Programs: ......... Andrew Zitzmann
Director of Membership: ...... Hugh Nicoll
Director of Public Relations: ... Sayoko Yamashita

Chapters

Akita, Chiba, East Shikoku, Fuku, Fukuoka, Gifu, Gunma, Hamamatsu, Himeji, Hiroshima, Hokkaido, Ibaraki, Iwate, Kagoshima, Kitakyushu, Kobe, Kyoto, Matsuyama, Miyazaki, Nagasaki, Nagoya, Nara, Okayama, Okinawa, Omiya, Osaka, Sendai, Shinshu, Shizuoka, Tokyo, Toyohashi, Wakayama (forming), West Tokyo, Yamagata, Yokohama...

Special Interest Groups

Bilingualism; College and University Educators; Computer-Assisted Language Learning; Gender Awareness in Language Education; Global Issues in Language Education; Japanese as a Second Language; Junior and Senior High School Teaching; Learner Development; Material Writers; Other-Language Educators; Pragmatics; Professionalism, Administration, and Leadership in Education; Teacher Education; Teaching Children; Teaching Older Learners; Testing and Evaluation.

JALT Central Office
Urban Edge Building, 5F 1-37-9 Taito, Taito-ku, Tokyo 110-0016, Japan
Tel.: 03-3837-1630; Fax: 03-3837-1631; E-mail: jco@jalt.org;
Website: <www.jalt.org>
In this Issue

Articles
The main section of this issue contains five articles. First, Hasegawa Shu-uji, Chuujyou Kiyomi, and Nishigaki Chikako report on their study on the English vocabulary used in university entrance examinations focusing on the difficulty and usability of the words used. Next, Junko Matsuzaki Carreira examines elementary school students’ motivation for learning English and how it can change with age. Next, Kuramoto Atsuko, Yoshida Shinsuke, and Yoshida Haruyo discuss how different learning environments affect English learning tasks; their focus is on PC-equipped classrooms versus classrooms which do not have computers. Akio Suzuki compares the different reading strategies used by students making graphic organizers with the strategies used by students writing summaries. Finally, Terry Vanderveen discusses the effect that student self-monitoring has on test scores.

Reviews
In this issue we have five book reviews. In the first one, Michael Lessard-Clouston reviews an edited volume that looks at the experiences of 15 nonnative English teachers in learning and teaching English. Next, Fiona MacGregor reports on a new edition of what has become a classic reference to grammar and other aspects of language, Michael Swan’s Practical English Usage. Thirdly, Cheryl Kirchhoff reviews a book that will be of interest to anyone wanting to stay up to date on current research on spoken English and how it can be applied to the language classroom. Next, Byron O’Neill reviews an introductory text that provides an overview of the applied linguistics field geared for those with little or no prior knowledge of it. Finally, Gerry Lassche reports on a book that takes a critical look at the current state of discourse analysis.

From the Editors
By the time most of you read this, the 32nd international JALT conference in Kitakyushu will be history. I hope that those of you who were able to attend stopped by the publication table either to talk about ideas you have about potential articles or to find out how you can become more involved in this aspect of JALT. JALT Publications (the JALT Journal, The Language
Teacher, JALT Conference Proceedings) is always looking for volunteers to help with the reviewing, editing, and proofreading of articles. If you are interested, please contact me at <jj-editor@jalt-publications.org> and I will make sure you are put in touch with the appropriate party. Speaking of parties, at the conference there was a birthday celebration for The Language Teacher as it turned 30. Many readers may not realize it but The Language Teacher is the only monthly, blind-reviewed language teaching journal that we are aware of—Happy Birthday, TLT! In the last issue, Yuri Hosoda’s name was inadvertently left off the list of Editorial Advisory Board members. I have rectified that omission and wish to thank her for her work on the board. And I’d like to close by giving a special thanks to all the editorial board members, additional readers, proofreaders, and other volunteers who help make the JALT Journal what it is.

NonJALT Publishing Opportunity
Now entering its 40th year, Language Teaching is a quarterly, professional, peer-reviewed journal which is dedicated to providing a number of resources to its international readership of researchers and practitioners in L2 education at all levels of instruction.

One section of the journal publishes academic studies which use a replication approach. Language Teaching invites submissions of previously unpublished articles based on literal, approximate, or constructive replication of a previous study or studies. Language Teaching also encourages submissions of meta-analyses which attempt to combine or synthesise a series of comparable research replications. Papers can be based on a broad range of topics, including:

- language teaching
- teacher training
- curriculum design and materials development
- language learning
- language testing
- teacher education
- neurolinguistics
- bilingualism/bilingual education
- sociolinguistics
- psycholinguistics
- pragmatics
- applied linguistics

More information can be obtained from the editor, Graeme Porte at <gporate@gmail.com>.
Entrance exams at the university level in Japan contain a particular type of English vocabulary known as “juken-eigo,” which is specifically employed in entrance examinations and has—in some cases—no other practical application. It has been noted that such juken-eigo vocabulary is comprised of high-level words that are not even included in junior and senior high school (JSH) English textbooks. This study undertakes (a) to scrutinize the language used in the Examination of the National Center for University Entrance (Center exams) and the entrance exams administered by the individual universities (University exams) and (b) to discover how the language compares with the vocabulary to which students are introduced in JSH textbooks. Because Japan’s Ministry of Education, Science, and Culture revises its Course of Study Guidelines approximately every ten years,
exams from 1988, 1998, and 2004 were chosen to be used as samples for the study. The authors reviewed three Center exams and 48 university exams that had been administered during those years and they made quantitative/qualitative observations about the juken-eigo vocabulary.

To complete the study, the changes in the characteristics of juken-eigo vocabulary were observed, by year, from the following five aspects: (a) the total number of words (types and tokens) used in Center and in university examinations, respectively; (b) the coverage of JSH English textbook vocabulary for individual Center and University exams; (c) the vocabulary level of Center and university exams with respect to the level presented in JSH English texts, then compared to the high-frequency words of the British National Corpus (BNC); (d) the number of outstanding juken-eigo words in the University examinations, as identified from available usage statistics and compared with the high-frequency words of the BNC; and (e) the scope of the vocabulary used in the University exams compared with the coverage of the related JSH textbooks, with regard to seven practical applications in spoken English and in written English, respectively.

The study reveals that students are expected to be familiar with more words every year and that there are a higher percentage of juken-eigo words appearing in the individual University exams than in the Center exams. This study data shows that the juken-eigo vocabulary used in University exams is, in fact, above the level taught in JSH English textbooks, while the vocabulary used in the Center exams is very nearly appropriate for students who have just graduated from a senior high school in Japan. The authors discuss the implications of those findings and of the data collected when the vocabulary in respective entrance exams was compared with the high-frequency words of the BNC.

Although there is some controversy over the esoteric nature of some of the words used in the vocabulary of university exams, it is important to note that, as time passes, the words deemed to be a part of the juken-eigo vocabulary are more in line with current events and practical spoken applications than are the words presented in the standard JSH textbook vocabulary. The fact is that if the public’s increasing demand for practical communicative English language ability is taken into account, a student’s modest vocabulary can be supplemented with much of the juken-eigo vocabulary such as that used in the university exams. During the study, it was noted that the level of usability exhibited in the "exam vocabulary" has increased, slightly—in both spoken and written English. It is the authors’ contention, however, that students can be better prepared for the entrance examinations if the level of vocabulary used in the university exams was brought more into line with what students actually study when they are in school. This study provides valuable information and basic data, which can provide a solid foundation for the important discussion about what an entrance examination should be.
は、大学入試英語問題全体に使用される英語語彙の難易度と有用性を、1980年代以降を3年代に分けて定量的・定質的に調査し検証することを目的とした。調査分析は、（1）異音数と延べ音数、（2）中高英語教科書語彙によるカバー率、（3）British National Corpusを基準尺度にした語彙レベル、（4）特微音の抽出、（5）音声英語と文字英語の各々分野に対するカバー率の測定、という5項目で実施した。結果からは、共通一次・センター試験問題語彙の難易度はほぼ適切であるのに対し、個別大学入試英語問題語彙のそれは高すぎることが確認された。一方、個別大学入試英語問題語彙の有用性は高いことが判明した。

1. はじめに

日本の大学入試英語問題で扱われる英語は、入学試験のための特殊な英語とされ、いわゆる「受験英語」と称される。「受験英語」の起源は、明治36年（1903）に出版された南日恒太郎著『難問分類英文詳解』にあるとされるが（川澄, 1978; 若林, 1988）、以来100年以上経った現在、「さすがに珍問、奇問は少なくなくなったものの、中学から高校へどんな勉強をしてきた受験生なのかという認識が欠如している問題を平気で出す大学も目につく」（田崎, 2000: 19）という指摘がある。関連して、Watkins他（1997）は、受験英語についての特質と出題に関わる諸問題を具体的に分類し、「現在のような難解な、高校教育を全く無視したような入試問題がある限り受験生の負担は増えるだけである」（p. 50）と述べている。同様の意見は、新里（1990）、酒井（1996）、小林（2000）などにも見られる。

一方、別府（2003）によれば、昨今の大学入試の英語問題は「英文和訳問題や細かい文法問題が中心」という「イメージ」から脱却し、「長文問題や会話問題の増加」、「細かい知識を問う問題の減少」、「公立大学入試問題における英文和訳問題の比重のさささ」が「現実」であるという。さらに深沢（1999）は、1989〜1998年の10年間の国立大学約60の入試英語のライティング問題の推移を問題形式別に調査した結果、和文英訳の減少と自由英作文の増加を認め、ライティングにおいても入試問題は変化していることを報告している。

このような英文読解やライティングにおける問題形式の変化、特殊な構文・文法および礎末な知識を問う問題の減少を歓迎する声に対し、大学入試の英語問題に使用される「語彙」の難易度が依然として高すぎるのではないかという指摘がある。たとえば、小林（2003）は1982年度と2002年度の入試問題を各1000題無作為に選び比較したところ、長文で使用される頻出単語に有意な差は無かったという。したがって、学習指導要領が改訂され、「中高で学習する単語数が少なくても良いほど、入試の語彙レベルとの差が大きくなり、それだけ受験対策のための負担が大きくなる」（p. 8）と述べている。

関連調査として、長文読解問題に使用される語彙の難易度について、中條・長谷川（2004）は、中高英語教科書語彙のカバー率とリーダビリティの観点から、過去10年分（1993〜2002）のセンター試験と2002年実施の大学入試問題を分析した。その結果、センター試験では年度による差は少ないものの、語彙のカバー率とリーダビリティからはほぼ適正なレベルと判定された。一方、個別の大学入試問題では、中高英語教科書語彙で対応できる問題は4学部（10％）、高校卒業レベルとして適切なリーダビリティのものは12学部（30％）にとどまった。また、Matsuo（2000）は1991年から1997年までのセンター試験７年分、および国公立・私立各60大学の入試英語読
解問題と高校教科書7種類を比較し、その語彙の重なり量から高校教科書だけでは大学入試に対応できないという深刻な状況を報告している。

さらに、長文読解問題だけでなく、文法や語彙に関するテスト項目なども含めた試験問題全体で使用される語彙を対象にした調査に長谷川（2003）がある。長谷川はセンター試験を対象に、中高英語教科書語彙のカバー率を1993〜2002年の10年間にわたってを通時的調査をした。その結果から、センター試験においては、高校で難易度上級の教科書を使用した場合でも、カバー率がテスト理解の閾値とされる95%に達しない年度が多いことを報告している。しかしながら、個別の大学入試問題については、問題全体で使用される語彙の難易度を調査した報告はこれまでに無い。

近年、大学生の学力低下が問題となっているが、大学生の学力低下が指摘され始めたのは1990年代であると言われる（佐藤, 2001; 前谷, 2003）。また、大学生の学力は小学校から高等学校までの学習の積み重ねであるとすれば、英語は中・高の学習の積み重ねである。伊村（2003: 117）によれば、中・高で学習される英語の新語数の合計は、昭和26年（1951）以来、学習指導要領が約10年ごとに改訂されるたびに減少している。大学生の学力低下との関連で、1990年代とその前後を見れば、1980年代、1990年代、2000年代の順で、2,300〜2,950語→2,900語→2,700語となっている。したがって、実際に大学入試に出題される英語問題語彙の難易度を判定するためには、中高英語教科書語彙との関係で、このような時代的な変化を考慮にいれた調査が必要であると考えられる。

そこで、様々な議論がある日本の大学入試英語問題の現状に鑑み、センター試験および個別の大学入学試験の英語問題全体について、中高英語教科書語彙から見た語彙の難易度を、時代の推移にしたがって調査することにした。調査年代は、大学生の学力低下が指摘され始めた1990年代を中心に、1980年代、1990年代、2000年代という時代区分で行った。また同時に、昨今、急速に高まりつつある実践的コミュニケーション能力育成のための英語教育において、「大学入試英語問題語彙」がどの程度寄与できるかを、有用性の面から定量的・定質的に調査することにした。そして、大学での英語教育を効果的に行う上で、「中高英語教科書語彙」に「大学入試英語問題語彙」を加えた、大学入学時英語学習者の語彙力の上限を時代の推移とともに実際の使用場面を想定して推定しようと試みた。

2. 研究の目的

大学入試英語問題（共通一次・センター試験問題、および国公立・私立大学の個別試験問題）に使用される語彙の難易度が、中高英語教科書語彙から見て適切であるかを、1980年代、1990年代、2000年代の時代区分により定点観察をする。また同時に、大学生の学習における実践的コミュニケーション能力育成のための英語教育において、「大学入試英語問題語彙」がどの程度寄与できるかを、有用性の面から定量的・定質的に調査することにした。そして、大学での英語教育を効果的に行う上で、「中高英語教科書語彙」に「大学入試英語問題語彙」を加えた、大学入学時の英語学習者の語彙力の上限を時代の推移とともに実際の使用場面を想定して推定しようと試みた。

上記の研究目的を遂行するため、具体的な調査項目は下記の5点とする。大学入試問題語彙の「難易度」に関する調査が（1）〜（4）、「有用性」に関する調査が（5）である。
（1）各入試問題の延べ語数と異語数
（2）各入試問題に出現した語彙に対する中高英語教科書語彙の割合（カバー率）
（3）各入試問題および中高英語教科書語彙の語彙レベル
（4）年代別16大学合計入試問題語彙10に顕著に出現した語（特徴語）
（5）年代別「中高英語教科書＋16大学合計入試問題」語彙の有用性の計測

3．研究の方法

3.1 調査対象とする言語材料

1980年以降の高等学校学習指導要領の改訂施行年度は、1982年、1994年、2003年であり、各々3年後が当該学習指導要領で学習した生徒の最初の大学入試となる。したがって、2003年施行の学習指導要領による大学入試は2006年の開始である。さらに、中高学習指導要領の接続年度がずれている年があることを考慮に入れて、1988年、1998年、2004年の大学入試を調査対象とした。大学入試問題は、共通一次・センター試験問題と、一般に難関校と言われる学習者が目標とする国公立・私立の16大学（国公立8大学、私立8大学）の個別試験問題を選定した。比較対象とする中・高英語教科書は、各入試年度に対応した高校英語教科書を基準とした。詳細は下記のとおりである。

（1）共通一次・センター試験英語問題

1988年実施共通一次試験（本試験）、および1998年と2004年実施のセンター試験（本試験）とし、次のデータベースから必要部分を使用した。

『センターTen 英語』（ジェイシー教育研究所, 2003）
『センターTen Plus 英語』（ジェイシー教育研究所, 2004）

（2）大学個別入学試験英語問題

将来英語を使用する機会が比較的多いと考えられる学部・学科で、3年代を通じて同一大学での比較が可能となるように、1988年、1998年、2004年実施の国公立8大学、私立8大学、合計16大学の文学部（英文科）もしくは経済学部の問題とした。国公立8大学の内訳は、旧帝国大学4校（東京大学、東北大学、京都大学、九州大学）、首都圏新制大学4校（筑波大学、千葉大学、横浜国立大学、東京都立大学）、私立8大学の内訳は、東京都内4大学（早稲田、慶應義塾、上智、青山学院）、関西圏4大学（関西学院、関西、同志社、立命館）である。使用した資料とデータベースは、次のとおりである。

『昭和63年 全国大学入試問題正解 英語 国公立大編』（旺文社, 1988）
『昭和63年 全国大学入試問題正解 英語 私立大編』（旺文社, 1988）
『Xam ’98 全国大学入試問題データベース 英語』（ジェイシー教育研究所, 1998）
『Xam 2004 全国大学入試問題データベース 英語』（ジェイシー教育研究所, 2004）
（3）中学・高等学校英語教科書

中学校から高等学校までのひとりの生徒が英語学習のために使用する教科書は、中・高で各1シリーズの教科書であることが多いので、中・高ともに1980年代から2000年代まで採択数上位にあった教科書シリーズより選定した（cf.『内外教育』, 『教科書レポート』）。高等学校用は、大学進学者の多い普通高校で一般に使用されると考えられる「英語Ⅰ」「英語Ⅱ」「リーディング（1980年代は英語ⅡB）」とした。本調査で使用した教科書は以下のものであり、各教科書の「各課の本文」及び「Supplementary Reading」を対象とした。

中学校: New Horizon 1, 2, 3 (東京書籍, 1988, 2000)

3.2 調査項目の分析方法

本調査で使用した言語材料は、電子化されているものはデジタルデータを利用し、電子化されていないものはスキャナを使用して入力後、校正し、単語の変化形を基本形に集約した語彙リストを作成した。特定のテキストに多く出現して計測結果に影響を与えやすい固有名詞・数詞・略語・間投詞・記号は削除した。加えて、大学入試英語問題から注釈の付いている語を全て削除した。

言語材料のうち、1980年代から2000年代までの大学入試英語問題語彙の時代的変化を多角的に探るため、1988年、1998年、2004年の各々共通一次・センター試験問題語彙と個別大学試験問題語彙に対する比較調査5項目の分析方法は、下記(1)〜(5)のとおりである。

(1) 各入試問題の延べ語数と異語数

本研究では、読解問題のみならず文法や語彙に関する問題等も含めた各入試問題全体で使用される語彙の量的変化を探るため、各問題で使用される英単語の延べ語数と異語数を求めて、年代順に比較した。

(2) 各入試問題に出現した語彙に対する中高英語教科書語彙の割合(カバー率)

各入試英語問題の延べ語数に対し、その何％を中高英語教科書語彙でカバーできるかを、年度ごとに求めて比較した。現在、語彙研究の分野では英文の内容を理解するために、当該英文の95％以上にあたる語彙数が最低限必要であろうという考え方をする研究者が多いようである（Laufer, 1997; 梅内, 1997; Read, 2000; Nation, 2001; 林, 2002）。そこで、本研究でも難易度の判定にあたっては、「95％カバー率」をひとつの目安とした。

(3) 各入試問題および中高英語教科書語彙の語彙レベル

本研究ではBritish National Corpus (BNC)を基準尺度とし、その頻度上位何語で各入試問題および教科書の語彙を95％以上カバーできるかを算定する手法で、入試問題と中高英語教科書の語彙レベルを推定した。「95％」という基準は、上記Laufer (1997) 等による。基準尺度に用いたリストは、Chujo (2004)で作成したBNC頻度上位13,994語である。
（4）年代別16大学合計入試問題語彙に顕著に出現した語（特徴語）

「受験英語」という特定分野の英文の特徴を強く反映する語を抽出するには、対数尤度比という統計指標を利用することができる（Scott, 1999；中條他, 2005）。本研究では各年代別に16大学の入試問題を集めた入試語彙の総リストと、上述のBNC頻度上位13,994語を対数尤度比を用いて統計的に比較し、BNCのような汎用の英文資料の語彙出現状況に比べ、入試問題に顕著な出現状況を示す語を抽出した。

（5）年代別「中高英語教科書+16大学合計入試問題」語彙の有用性の計測

日本人英語学習者が高校卒業後に大学生となり、グローバル化社会の中で生きるために必要とされる英語という観点から、音声英語と文字英語、各5領域7分野の英文に対する「中高英語教科書語彙と16大学入試問題語彙の総和」によるカバー率を計測し、年代別に比較した。その結果は、長谷川・中條(2004)で調査した英語教科書語彙の有用性の計測結果とも比較できるように、表1に示した英文資料16を用いて有用性の指標とした。

特に、「英語コミュニケーション能力試験」については、文部科学省（2003）の「英語が使える日本人の育成のための行動計画」に示された英語の到達目標として使用される指標の中から、TOEICとTOEFLという2分野の言語材料を用意した。同様に「情報収集」についても、中條・長谷川(2003)を参考にして、日本人英語学習者が到達目標にすると考えられるもの（PBS、TIME）と、初心者向けの教育的配慮のあるもの（VOA、News for You）という観点から2分野ずつとした。

### 表1 大学入試問題の有用性の計測に用いた英文資料

<table>
<thead>
<tr>
<th>音声英語</th>
<th>文字英語</th>
</tr>
</thead>
<tbody>
<tr>
<td>英語コミュニケーション能力試験</td>
<td>TOEIC (リスニング・セクション)</td>
</tr>
<tr>
<td></td>
<td>TOEFL (リスニング・セクション)</td>
</tr>
<tr>
<td>大学留学</td>
<td>チュートリアル</td>
</tr>
<tr>
<td>情報収集</td>
<td>PBS (TVニュース)</td>
</tr>
<tr>
<td></td>
<td>VOA (ラジオ・レポート)</td>
</tr>
<tr>
<td>日常生活</td>
<td>サバイバル英語 (生活英語)</td>
</tr>
<tr>
<td>趣味・教養</td>
<td>映画 (Titanic)</td>
</tr>
</tbody>
</table>

カバー率の計測は、信頼性の高い計測結果を得るため、Chujo & Utiyama (2005a, 2005b)を参考にして、分野ごとに1,500語のサンプルを5個無作為に抽出し、その各々に対するカバー率を求め、その平均値を使用した。なお、
TOEICとTOEFLは試験問題という性格上、問題全体を対象とし、リスニング・セクションとリーディング・セクションについてそれぞれ2回分の試験問題のカバー率を計測し、その平均値を使用した。TOEICとTOEFLは、それぞれリスニング・セクションとリーディング・セクションの延べ語数が各々3,000語以上であることから、サンプル2個の平均でも安定した結果を得られると判断した（Chujo & Utiyama, 2005a; 2005b）。

4. 結果と考察

4.1 延べ語数と異語数

共通一次・センター試験、個別大学入試の各英語問題に使用された英語の語彙について、1988年、1998年、2004年でそれぞれ延べ語数と異語数を計測した結果を表2に示した。個別大学は国公立と私立に分類し、そのカテゴリーの中で2004年の異語数が降順になるようにした。最下段に示した平均値は、国公立・私立を合わせたものである。

| 表2 大学入試問題に用いられた語彙数の変化（単位：語） |
|-----------------|--------|--------|-----------------|--------|--------|
|                 | 1988年 | 1998年 | 2004年         |
| 共通一次・センター試験 |         |         |                 |
| 異語数          | 延べ語数| 異語数  | 延べ語数        | 異語数  | 延べ語数|
| 635             | 2541   | 657     | 3005            | 639     | 2943   |
| 国公立大学      |         |         |                 |
| 東京            | 584     | 1872    | 465             | 1795    | 682    | 2403   |
| 筑波            | 335     | 846     | 493             | 1370    | 597    | 1748   |
| 九州            | 475     | 1452    | 470             | 1411    | 556    | 1391   |
| 東北            | 512     | 1593    | 662             | 2126    | 482    | 1721   |
| 東京都立        | 344     | 869     | 405             | 1073    | 451    | 1380   |
| 横浜国立        | 522     | 1671    | 552             | 1553    | 432    | 1480   |
| 千葉            | 437     | 1682    | 666             | 1996    | 420    | 1199   |
| 京都            | 224     | 646     | 309             | 745     | 359    | 892    |
| 私立大学        |         |         |                 |
| 立命館          | 438     | 1603    | 781             | 3486    | 923    | 4241   |
| 上智            | 584     | 1998    | 650             | 2487    | 838    | 3207   |
| 関西            | 500     | 1493    | 620             | 2049    | 685    | 2691   |
| 同志社          | 457     | 1435    | 672             | 2535    | 650    | 2640   |
| 早稲田          | 470     | 1156    | 761             | 2553    | 623    | 2021   |
| 関西学院        | 596     | 2043    | 570             | 1815    | 595    | 2005   |
| 青山学院        | 399     | 1427    | 446             | 1429    | 419    | 1307   |
| 慶應義塾        | 426     | 1400    | 354             | 790     | 397    | 1085   |
| 国公立私立大学平均 | 456     | 1449    | 555             | 1826    | 569    | 1963   |

表2から、共通一次・センター試験では、延べ語数が年代別に、2541語→3005語→2943語となっており、80年代と比較して90年代と2000年代が相対的に多い。異語数では、635語→657語→639語と、ほぼ延べ語数の多寡に比例してい
個別大学入試問題の延べ語数と異語数を年代別に平均値で比較すると、延べ語数は、1449語→1826語→1963語となっており、試験で使用される英語の分量が年代ごとに増加している。また、異語数は、456語→555語→569語と、使用される語の種類も増加している。したがって、近年の入試英語の特徴のひとつである「読解問題の長文化」（安竹内, 1997）が、試験問題全体の語彙数の増加に反映しているのではないかと考えられる。

4.2 入試問題語彙に対する中高英語教科書語彙の割合（カバー率）

中高英語教科書の語彙で個々の入試問題の延べ語数の何％の語が既習となるかというカバー率を計測した結果を、年代別に表3に示した。国公立・私立ごとに2004年のカバー率で降順に示しており、最下段の平均値は国公立・私立を合わせたものである。なお、中高英語教科書の異語数は、80年代：2779語、90年代と2000年代：3098語であった。

表3 入試問題における中高英語教科書語彙のカバー率（単位：%）

<table>
<thead>
<tr>
<th></th>
<th>1988年</th>
<th>1998年</th>
<th>2004年</th>
</tr>
</thead>
<tbody>
<tr>
<td>共通一次・センター試験</td>
<td>94.8</td>
<td>96.4</td>
<td>96.4</td>
</tr>
<tr>
<td>国公立大学</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>東京</td>
<td>92.9</td>
<td>95.8</td>
<td>93.9</td>
</tr>
<tr>
<td>東京都立</td>
<td>88.5</td>
<td>90.5</td>
<td>92.9</td>
</tr>
<tr>
<td>横浜国立</td>
<td>87.7</td>
<td>85.6</td>
<td>92.6</td>
</tr>
<tr>
<td>千葉</td>
<td>93.5</td>
<td>88.7</td>
<td>92.2</td>
</tr>
<tr>
<td>東北</td>
<td>89.6</td>
<td>91.4</td>
<td>91.9</td>
</tr>
<tr>
<td>京都</td>
<td>87.0</td>
<td>88.6</td>
<td>91.0</td>
</tr>
<tr>
<td>筑波</td>
<td>92.7</td>
<td>87.1</td>
<td>89.6</td>
</tr>
<tr>
<td>九州</td>
<td>90.7</td>
<td>89.9</td>
<td>85.4</td>
</tr>
<tr>
<td>私立大学</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>関西</td>
<td>88.3</td>
<td>89.8</td>
<td>93.1</td>
</tr>
<tr>
<td>関西学院</td>
<td>88.4</td>
<td>90.7</td>
<td>91.6</td>
</tr>
<tr>
<td>早稲田</td>
<td>88.1</td>
<td>88.9</td>
<td>90.1</td>
</tr>
<tr>
<td>上智</td>
<td>89.5</td>
<td>93.4</td>
<td>89.7</td>
</tr>
<tr>
<td>同志社</td>
<td>87.7</td>
<td>88.8</td>
<td>89.7</td>
</tr>
<tr>
<td>慶應義塾</td>
<td>89.8</td>
<td>84.7</td>
<td>86.6</td>
</tr>
<tr>
<td>立命館</td>
<td>93.0</td>
<td>92.3</td>
<td>86.3</td>
</tr>
<tr>
<td>青山学院</td>
<td>89.6</td>
<td>87.9</td>
<td>83.5</td>
</tr>
<tr>
<td>国公立私立大学平均</td>
<td>89.8</td>
<td>89.6</td>
<td>90.0</td>
</tr>
</tbody>
</table>
表3において、共通一次・センター試験に対する中高教科書語彙のカバー率は、年代順に94.8％→96.4％→96.4％であり、いずれもほぼ95％前後である。したがって、共通一次・センター試験で使用された語彙の選定は、概ね適切であったと考えられる。また、90年代と2000年代のカバー率が、80年代に比較して上昇しているということは、近年、センター試験が易しくなったのではないかと言われる（cf. 片山他, 1997; 谷口, 1997; 武田, 2004）一因を裏付けるものと考えられる。

個別大学入試問題に対する中高英語教科書語彙のカバー率を年代ごとに平均で比較すると、89.8％→89.6％→90.0％となり、約90％でほとんど変化が無い。ただし、カバー率が90％以上の大学数は、80年代と90年代がそれぞれ5校、6校であるのに対し、2000年代は9校に増加している。したがって、近年、出題者側にも、大学入試問題語彙に対する配慮が現れ始めているのではないかと考えられる。中でも興味深いのは、受験生にとっては最難関の東京大学が、国公立・私立両大学の中で、80年代92.9％（3位）、90年代95.8％（1位）、2000年代93.9％（1位）、と一貫して90%を超え上位にあることである。

4.3 BNCを基準尺度とした入試問題の語彙レベル

各入試問題の語彙レベルは、対応する中高英語教科書語彙と比較すると相対的にどのような位置付けになるのかを、BNC頻度上位語の語数で表して比較した結果を図1に示した。図の数字は各試験・教科書語彙を95％カバーするのに必要なBNC頻度上位語の語数である。図には年代ごとに各試験における語彙レベルの平均値と、学習者の語彙レベルを表すと考えられる中高英語教科書の語彙レベルを示した。大学受験者は中学と高校の教科書を使用して英語の基本的部を学習してきているため、中高英語教科書語彙が彼らの語彙レベルの目安と考えられる。

図1から、共通一次・センター試験は年代順に、2185語レベル→2534語レベル→1903語レベルであり、いずれも対応する中高教科書の語彙レベル（80年代: 3295語レベル、90年代と2000年代: 3299語レベル）より低いので、高校修了時の英語の学力をみる試験として適切なレベルであると考えられる。

一方、個別大学入試問題の語彙レベルは年代別平均で、4125語レベル→4435語レベル→4224語レベルであり、対応する中高教科書の語彙レベルと比較して、明らかに難易度が高いと考えられる。図1には平均値のみを示したが、個々に観察した大学入試問題の語彙レベルからも、国公立・私立に関わりなく、各年代を通じて全ての大学入試問題が中高教科書の語彙レベルを超えていたことも確認されたことを付記する。

4.4 大学入試英語問題の特徴語

大学入試問題を特徴付ける語を、対数尤度比を用いて、一般的な英語使用を代表すると考えられるBNCと比較することによって抽出し、その特徴度の強い順に上位20語を表4に示した。なお、調査対象となった16大学入試英語問題の総和から求めた異語数は、1988年：2934語、1998年：3407語、2004年：3432語であった。
表4から、各年代における各入試問題の特徴語第1位にランクされた語は、1988年：fast-food、1998年：language、2004年：fishである。これらの語と上位20位以内に入った他の語を合わせて考えると、88年はrestaurant, fry, food, cornなど「食文化」について、98年はlisten, information, learn, communicationなど「言語」について、04年はoverfish, shrimp, aquaculture, fisher, preservation, mangroveなど「自然保護」についてのトピックが、大学入試問題の、特に長文読解問題の特徴をなしていたのではないかと推察される。さらに、98年はtsunami, deep-sea, dolphin, oceanなど「海洋」について、child, family, parent, humanなど「家族関係」についてのトピックも考えられる。

2004年度入試に出題された長文問題のトピックを分析した旺文社（2004a, 2004b）によれば、国公立・私立いずれも上位3位までは、「文化」「日常生活」「自然」の順で合計70％以上を占めている。表4からは、このようなトピックが1980年代より続いているのではないかということが推測できる。

大学入試問題に出題されるトピックは、その時代に関心を集めた事柄が取り上げられる可能性が高いのは、受験関係者には周知の事実である。時の話題に関するキーワードをその都度学習しておくことは、実践的コミュニケーション能力を高める上でも重要なことと考えられる。受験生にとっては、特定のトピックに関連語彙の知識が合否に影響を与えることも示唆されている（Brown & Yamashita, 1995a: 27）。また、大学入試の望ましくない波及効果が強調されるなかで、本来あるべき望ましい波及効果として、Mulvey (1999: 132)は、読解スキルの習得等のほかに語彙の補強を挙げている。このことから、視点を変えた入試問題の利用もありえるのではないかと考えられる。ただし、あくまでも受験生の学習負担が過剰にならないような配慮をすべきことは言うまでもない。
表4 大学入試問題の特徴語

<table>
<thead>
<tr>
<th>順位</th>
<th>1988年入試</th>
<th>1998年入試</th>
<th>2004年入試</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>fast-food</td>
<td>language</td>
<td>fish</td>
</tr>
<tr>
<td>2</td>
<td>snake</td>
<td>tsunami</td>
<td>tooth</td>
</tr>
<tr>
<td>3</td>
<td>blacksnake</td>
<td>child</td>
<td>print</td>
</tr>
<tr>
<td>4</td>
<td>restaurant</td>
<td>deepsea</td>
<td>overfish</td>
</tr>
<tr>
<td>5</td>
<td>fry</td>
<td>patina</td>
<td>shrimp</td>
</tr>
<tr>
<td>6</td>
<td>weighing-machine</td>
<td>family</td>
<td>downtown</td>
</tr>
<tr>
<td>7</td>
<td>food</td>
<td>dolphin</td>
<td>dentist</td>
</tr>
<tr>
<td>8</td>
<td>frontier</td>
<td>alcohol</td>
<td>tear</td>
</tr>
<tr>
<td>9</td>
<td>mother</td>
<td>ocean</td>
<td>baby</td>
</tr>
<tr>
<td>10</td>
<td>we</td>
<td>listen</td>
<td>advertiser</td>
</tr>
<tr>
<td>11</td>
<td>dead</td>
<td>dream</td>
<td>aquaculture</td>
</tr>
<tr>
<td>12</td>
<td>resurrection</td>
<td>information</td>
<td>fisher</td>
</tr>
<tr>
<td>13</td>
<td>lecture</td>
<td>facial</td>
<td>culture</td>
</tr>
<tr>
<td>14</td>
<td>author</td>
<td>parent</td>
<td>mathematics</td>
</tr>
<tr>
<td>15</td>
<td>gentility</td>
<td>we</td>
<td>preservation</td>
</tr>
<tr>
<td>16</td>
<td>philosophize</td>
<td>oxidize</td>
<td>recreational</td>
</tr>
<tr>
<td>17</td>
<td>doublet-and-hose</td>
<td>prejude</td>
<td>gender</td>
</tr>
<tr>
<td>18</td>
<td>corn</td>
<td>learn</td>
<td>people</td>
</tr>
<tr>
<td>19</td>
<td>I</td>
<td>communication</td>
<td>difference</td>
</tr>
<tr>
<td>20</td>
<td>civilization</td>
<td>human</td>
<td>mangrove</td>
</tr>
</tbody>
</table>

4.5 入試英語語彙の有用性

受験生は受験準備として過去に出題された入試問題を学校の内外で使用するため、本研究では、大学入学時の学生の語彙力の上限を、中高教科書語彙に加えて、今回調査に使用した16大学の入試問題語彙を全て習得した場合と仮定した。その際、どの程度の有用性が期待できるかを、音声英語と文字英語について各5領域7分野で計測した。結果はそれぞれ表5と表6に示した。表中の各年代において、左列がその年代に使用された中高英語教科書のみによるカバー率、右列（網掛け部分）が「中高英語教科書＋16大学入試英語問題」によるカバー率である。最下段には、各々7分野の平均カバー率と未知語に遭遇する割合を示した。なお、1990年代と2000年代では、受験生の学習した教科書語彙が共に90年代の学習指導要領に基づいているため、左列の数値が等しくなっている。調査対象とした「中高英語教科書＋16大学入試英語問題」の異語数は、1988年：4170語、1998年：4664語、2004年：4662語であった。
4.5.1 音声英語について

表5から、中高英語教科書のみのカバー率で、80年代教科書と2005年3月まで使用されていた90年代教科書を比較すると、90年代教科書の方が調査した7分野のほぼ全てにおいてカバー率が上昇し、平均値で見ても、91.4％から92.4％に上昇している。

次に、教科書語彙に入試語彙を加えた（網掛部分）、①「80年代教科書＋88年入試」、②「90年代教科書＋98年入試」、③「90年代教科書＋04年入試」を比較すると、指標とした7分野に対するカバー率の平均は、①から③へ向かって、94.1％→95.0％→95.5％と近年になるにつれて上昇している。結果的に入試問題語彙を習得すれば、実際の使用場面での有用性の向上が期待できる可能性が見とれる。

調査対象とした7分野の項目別に見た場合、カバー率が95％以上になっているのは、①では「TOEIC（リスニング・セクション）」と「サバイバル英語（生活英語）」だけであるが、②では「チュートリアル」が加わり、③ではさらに「TOEIC（リスニング・セクション）」と「VOA（ラジオ・レポート）」が加わって、7分野中5分野になる。したがって、大学入試問題語彙は、近年では実際的使用が可能な分野にも広がりを見せていると言える。

近年の時代的変化として見た場合、98年入試と04年入試では、受験生の学習した教科書語彙が共に90年代の学習指導要領に基づいているため、カバー率の向上は98年と04年の入試問題の変化を如実に反映している。したがって、明らかに98年より04年の入試問題語彙の方が有用性が高いと考えられる。特に04年では、TOEICのリスニング・セクションでカバー率が95％を超えたことは注目に値する。TOEICと並んで英語能力判定の資料に利用されることの多いTOEFLにおいては、1980年代までに95％を超えている。一方、TOEICでは2000年代になってようやく95％を超えた。昨今、大学英語教育の中でTOEIC対
策の授業を展開しているところも増えているが、試験問題作成者にもTOEICの問題が意識されているのではないかと推測できる。

4.5.2 文字英語について

表6から、まず中高英語教科書のみのカバー率を観察すると、調査した7分野に対するカバー率の平均（表中最下段）は、80年代教科書と90年代教科書とでは83.2%から85.3%に上昇している。指標とした7分野のそれぞれを見ても、全ての分野でカバー率が上昇している。しかしその90年代教科書でさえ、95%カバー率を目安にした場合、音声英語では平均2.6%（95%－92.4%）不足していたのに対し、文字英語では平均9.7%（95%－85.3%）不足しており、英語を聞いて理解する場合には困難を生じるが、読んで理解する場合には、それ以上の困難を伴うことが予想される。特に、「生活案内」「TIME（英文雑誌）」「TOEIC（リーディング・セクション）」の分野におけるカバー率が低い。

表6 年代別に見た教科書と入試語彙の文字英語の英文資料に対する有用性

<table>
<thead>
<tr>
<th></th>
<th>1980年代</th>
<th>1990年代</th>
<th>2000年代</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>80年代</td>
<td>+88年</td>
<td>90年代</td>
</tr>
<tr>
<td>教科書</td>
<td>入試</td>
<td>教科書</td>
<td>入試</td>
</tr>
<tr>
<td>文字英語</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOEIC（リーディング・セクション）</td>
<td>80.0</td>
<td>87.1</td>
<td>81.7</td>
</tr>
<tr>
<td>TOEFL（リーディング・セクション）</td>
<td>82.9</td>
<td>88.9</td>
<td>85.2</td>
</tr>
<tr>
<td>大学留学</td>
<td>大学入学案内</td>
<td>81.9</td>
<td>89.6</td>
</tr>
<tr>
<td>情報収集</td>
<td>TIME（英文雑誌）</td>
<td>79.7</td>
<td>85.0</td>
</tr>
<tr>
<td></td>
<td>News for You（ESL英字新聞）</td>
<td>88.5</td>
<td>92.3</td>
</tr>
<tr>
<td>日常生活</td>
<td>生活案内</td>
<td>78.2</td>
<td>85.3</td>
</tr>
<tr>
<td>趣味・教養</td>
<td>小説（Harry Potter）</td>
<td>91.1</td>
<td>93.1</td>
</tr>
<tr>
<td>平均カバー率（%）</td>
<td>83.2</td>
<td>88.8</td>
<td>85.3</td>
</tr>
<tr>
<td>未知語に遭遇する割合（語）</td>
<td>5.9</td>
<td>8.9</td>
<td>6.8</td>
</tr>
</tbody>
</table>

一方、教科書語彙に入試語彙を加えた（網掛部分）、①「80年代教科書＋88年入試」、②「90年代教科書＋98年入試」、③「90年代教科書＋04年入試」を比較すると、文字英語における実際の使用場面と想定した7分野に対するカバー率の平均は、①から③へ向けて、88.8%→90.3%→90.8%とわずかであるが増加している。調査した7分野の分野別に見ても、①→②→③と時代を追うごとに、ほぼ全ての項目でカバー率が上昇している。とはいえ、どの分野においてもカバー率が95%には到達していない。

③の「教科書語彙＋04年入試」においてカバー率が相対的に平均（90.8%）より高い分野は、高い順につき「小説（Harry Potter）」（94.2%）、「News for You（ESL英字新聞）」（93.9%）、「大学入学案内」（91.1%）である。一方において、英語上級レベルの学生や社会人の目標とする「TIME（英文雑誌）」（87.5%）や、学校英語教科書の弱点として指摘されてきた日常生活語彙（中条他、1993）の分野におけるカバー率が低い。
中條, 2004)に該当する「生活案内」（88.3%）が相対的に低い。大学においては、このような分野の語彙の補強が必要であると言える。とりわけ、中高英語教科書から大学入試英語問題に至るまで全ての学習段階を通じて不足する日常生活語彙については、その指導に関して特別な配慮が必要であると考えられる。

近年、英語教育の到達目標値を設定する際にしばしば利用される「TOEIC」と「TOEFL」において、③のリーディング・セクションのカバー率はそれぞれ89.6%と90.7%である。各リスニング・セクションがそれぞれ95.2%と95.9%であることを考慮に入れると、大学生のTOEIC、TOEFL両テストの試験対策のひとつとしては、リーディング・セクション用の語彙を補強する必要があると考えられる。ただし、中高教科書語彙に大学入試問題語彙を追加することによるカバー率の上昇を3年代で比較すると、90年代教科書語彙に04年入試語彙を追加した場合の「TOEIC」が最大（7.9ポイント：89.6% - 81.7%）である。したがって、音声英語と同様に文字英語においても、近年、コミュニケーション能力試験への関心の高まりが入試問題の作成に影響を与えているのではないかと考えることができる。

5. まとめ

日本の大学入試英語問題に出題される英語は「受験英語」と呼ばれ、大学に入れるための特別な英語であって、実際の使用場面ではあまり役立たないのではないかと考えられている。さらに、その語彙の難易度は、文部省の学習指導要領に基づく中高英語教科書語彙のそれを超えて、難しすぎるのではないかということがしばしば指摘されている。そこで、本研究の目的は、大学入試英語問題全体に使用される英語語彙の難易度と有用性を1980年代、1990年代、2000年代とに分けて実際に調査し検証することであった。

結果から判明した点は、（1）本稿で調査した個別大学入試英語問題語彙は、異語数と延べ語数が1980年代から2000年代に向けて近年増加していること、（2）中高英語教科書語彙によるカバー率が、共通一次・センター試験問題語彙では3年代を通じてほぼ95％前後であるのに対し、個別大学入試英語問題語彙は平均90％であること、（3）BNCを基準尺度とした相対的な語彙レベルが、共通一次・センター試験問題語彙では中高英語教科書語彙より低いのに対し、個別大学入試英語問題語彙では高いこと、（4）個別大学入試英語問題の特徴語は各時代の話題と関係がありそうなこと、（5）中高英語教科書語彙に16大学合計入試英語問題語彙を加えると、カバー率が音声英語では7分野中5分野で95％以上であるのに対し、文字英語ではどの分野でも95％には及ばなかったものの、音声英語・文字英語共に近年カバー率が上昇しているということであった。

ただし、大学入試問題語彙の難易度は、個々の問題の目標や内容等を考慮に入れた場合、その評価は多少異なる様相を示す可能性もある。また、有用性の査定では、本研究で調査した5領域7分野以外にも指標として利用可能なものもあると考えられる。

このような前提条件に立った上で、本研究の結果から判明したことより、個別大学の入試英語問題作成にあたっては、使用される語彙が中高英語教科書語彙の難易度を超えているという、従来からの指摘を充分考慮に入れる必要があると考えられる。一方で大学受験生にとって、受験英語は将来必要となる実際
の使用場面で役立つ可能性が高いということが確認できたことは、朗報と言えるだろう。また、視点を変えて、大学入試問題を語彙力補強のための牽引車となる教材として、積極的に利用していこうという方法も考えられる。このような現実を踏まえつつ、個別の大学入試英語問題語彙の選定においては、今後さらに改善がなされることにより、中・高校の教育現場への望ましい波及効果が高まることを期待したい。

注
1. 別府（2003）がこのように判断するのは、Beppu（2001）、宇都・柳瀬（2000）のデータを根拠としている。Beppuは1996～2000年の5年間にわたる国公立5大学、私立大学20学部、およびセンター試験5回分の延べ合計130の入試問題を、宇都・柳瀬は1998年の国立40大学・47学部、公立13大学・13学部、私立51大学・83学部の全問題から長文問題を対象として調査した。
2. カバー率（coverage）とは、ある語または語の集合が、テキスト全体（延べ語数）の何パーセントを占めるかという指標である（Schmitt & McCarthy, 1997: 328）。
3. リーダビリティ（readability）は、「文章を読みやすくする要因、すなわち単語の難易、単語の長さ、セッションの長さなどの要因を組み合わせ、公式に代入して計算し、その数字を読書学年レベルとするものである」（高梨・卯城, 2000: 31）と定義される。
5. 閾値（threshold level）とは、「外国語の機能的な能力を達成するために必要な最低限の言語熟達レベル」（リチャーズ他, 1988: 381）のことである。
6. たとえば、小池（2004: 124）は、「第二言語学習者が習得目標とするべき語彙とは、使われる頻度が高い、有用性のある語彙ということになる」と述べている。また、西澤（2003: 10）は、「日本人英語学習者という視点に立った語彙選定は次の6項目に配慮することが必要である」とし、「①頻度、②汎用性、③生活用語、④授業用語、⑤日本人にとっての有用性、⑥カタカナ英語」を挙げている。このような観点を参考にし、本調査では、現在の日本人英語学習者の置かれた状況を考慮し、学習した語彙が、ある特定の目的で使用される可能性を想定して、「有用性」に焦点を当てるということとした。
7. 高等学校用教科書には、上級、中級、基礎という学習者向けレベルがあるとともに、大学入試には偏差値による大学のランク付けがある。これらのことを考慮し、高等学校用教科書は上級レベルを使用した上で、大学入試は偏差値ランク上位大学を受験した場合を想定し、その際に必要とされる語彙力を、「語彙力の上限」と定義した。
8. 「共通一次試験」が始まったのは1979年1月である。その後、1990年からは「大学入試センター試験」となり現在に至っている（清水, 1997: 16）。本研
究で調査対象とする入試問題は、1980年代から2000年代までの範囲であるため、「共通一次・センター試験問題」と表記することにする。

9. 国立大学は2004年4月より「国立大学法人」となり、公立大学は2005年4月時点では随時「公立大学法人」に移行しているが、受験関係誌では国公立大学という旧来の名称が使用されているため、本研究においてもこの呼称に従う。

10. 「16大学」とした理由は、本研究に先立って行った試算では、16大学を合計した入試英語問題語彙の異語数は約3,000語となり、中高英語教科書語彙の異語数とほぼ等しくなるため、両者を比較する上での量的目安として適当であると考えたからである。


14. 近年、日本において急速に受験者数の増加しているTOEICの大学専攻別受験者数の内訳（TOEIC運営委員会, 2005: 7）によれば、1位が語学・文学系（英語専攻）、2位が社会学系（商・経・法）である。このような英語運用能力に対する関心の高さという観点から、実際に入手できた大学入試問題という現実的理由から、文学部（英文科）と経済学部の問題を調査対象とした。

15. 2005年から首都大学東京と大学名を変更したが、調査対象とした入試問題は2004年までのものなので旧来の大学名を用いる。

16. 出典の詳細は、長谷川・中條（2004: 153）より次のとおりである。

17. 旺文社（2004a, 2004b）によれば、2004年度入試の英語問題全体に占める読解問題の割合は、国公立大で73.5％、私立大で52.5％であった。また、安竹内（1997）でも指摘されているように、現在の入試英語を支えている特徴のひとつは、読解問題の長文化である。これらのことから、本研究では英語問題全体で使用される語彙を対象にしているが、その過半数は読解問題で使用された語彙で、特徴語の多くは長文読解問題に起因しているのではないかと推定できる。

18. 羽鳥他（1979: 110）によれば、未知語に遭遇する割合が20語に1語であれば、なんとか英語のテキスト理解が可能であるという。したがって、20語よりも少なければ理解が困難であると考えられる。20語に1語というのは、95％カバー率に該当する。
参考文献
安竹内ひろし（1997）。「大学入試英語問題―悪い問題・良い問題」。『現代英語教育』、第34巻、第8号、4-9。
別府有紀（2003）。 「大学入試が高校英語教育に対して与える影響」。 『関東甲信越英語教育学会研究紀要』、第17号、77-88。
中條清美・長谷川修治・竹蓋幸生（1993）。 「日米英語教科書の比較研究から」。 『現代英語教育』、第29巻、第12号、14-16。
中條清美・長谷川修治（2003）。 「時事英語の授業で用いられる英文素材の語彙レベル調査―BNC (British National Corpus) を基準にして―」。 『時事英語学研究』、No. XLII. 51-62。
中條清美・長谷川修治（2004）。 「語彙のカバー率とリーダビリティから見た大学英語入試問題の難易度」。 『日本大學生産工学部研究報告B』、第37巻、45-55。
中條清美・内山将夫・長谷川修治（2005）。 「統計的指標を利用した時事英語資料の特徴語選定に関する研究」。 『英語コーパス研究』、第12号、19-35。
長谷川修治（2003）。 「英検2級とセンター試験に対する英語教科書語彙の効果—過去10年間の通時的調査」、『STEP BULLETIN』、Vol. 15、日本英検定期協会、152-158。
長谷川修治・中條清美（2004）。 「学習指導要領の改訂に伴う学校英語教科書語彙の時代的変化—1980年代から現在まで」。 『Language Education & Technology』、第41号、外国語教育メディア学会（LET）、141-155。
羽鳥博愛他（1979）。 『英語指導法ハンドブック4＜評価編＞』。 東京：大修館書店。
林洋和（2002）。 『英語の語彙指導―理論と実践の統合をめざして』。 広島：渓水社。
深沢清治（1999）。 「大学入試英語問題を考える」。 『現代英語教育』、第35巻、第12号、59。
伊村元道（2003）。 「日本の英語教育200年」。 東京：大修館書店。
菅谷剛彦（2003）。 『なぜ教育論争は不毛なのか―学力論争を超えて』。 東京：中央公論新社。
片山七三雄・大谷道子・金谷憲（1997）。 「変わりゆく入試英語と高校での対応」。 『英語展望』、No. 104、Summer、19-24、14。
川澄哲夫（編）(1978)。 『資料日本英学史2・英語教育論争史』。 東京：大修館書店。
小林功（2000）。 「2000年度の大学入試問題の傾向と今後の課題―改善が見られない大学入試問題」。 『英語教育』、第49巻、第3号、12-15。
小林功（2003）。 「入試の英単語と学習法を巡って」。 『英語通信』、第33号、8-9。
小池生夫（編集主幹）（2004）。 『第二言語習得研究の現在―これからの外国語教育への視点』。 東京：大修館書店。
文部科学省（2003）。 「『英語が使える日本人』の育成のための行動計画」。
新里真男（1990）。 「高校から大学入試を見ること」。 『英語教育』、第39巻、第9号、22-23。
西澤正幸（2003）。 「語彙数はどれだけ必要か？」。 『英語教育』、第52巻、第7号、8-10。
旺文社（2004a）。 『2005年受験用全国大学入試問題正解 英語（国公立大編）』。
旺文社（2004b）。 『2005年受験用全国大学入試問題正解 英語（私立大編）』。
リチャーズ,J.・プラット, J.・ウェーバー, H.（山崎真稔 他訳）（1988）。 『ロングマン応用言語学用語辞典』。 東京：南雲堂。
酒井邦秀（1996）。 『どうして英語が使えない？』。 東京：筑摩書房。
HASEGAWA, CHUJIYOU, & NISHIGAKI

佐藤学（2001）．『学力を問い直す―学びのカリキュラムへ』，岩波ブックレット，No. 548．東京：岩波書店。

清水裕子（1997）．「大学入試センター試験」の変遷—問題構成と配点をもとに」．『英語教育』，9月増刊号，Vol. 46，No. 7，16-18。

出版労連（1987）．『教科書レポート』，No. 31。

出版労連（2002）．『教科書レポート』，No. 46。

高梨庸夫・卯城祐司（2000）．『英語リーディング辞典』．東京：研究社出版。

武田一（2004）．「2004年度入試を斬る！—高校入試／センター試験／大学入試」．『英語教育』，10月増刊号，Vol. 53，No. 8，73-87。

谷口賢一郎（1997）．「新学習指導要領でセンター試験はどう変わったか」．『現代英語教育』，第34巻，第3号，8-12。

田崎清忠（2000）．「日本人にとっての「英語力」を見極めよ」．『英語教育』，第49巻，第2号，8-19。

TOEIC運営委員会（2005）．『TOEICテスト DATA & ANALYSIS 2004』．東京：財団法人国際ビジネスコミュニケーション協会。

投野由紀夫（編著）（1997）．『英語語彙習得論—ボキャブラリー学習を科学する』．東京：河源社。

宇都裕・柳瀬実佳（2000）．『大学入試における「和訳」の比重と内容—1998年度大学入試英語問題の徹底的研究 国・公立大学編、私立大学編の分析から—』．東京：学校法人 文京学園 文京語学教育研究所。

若林俊輔（1988）．「入試英語」．『英語教育』，9月増刊号，Vol. 37，No. 7，65-67。

Watkins, G.・河上道生・小林功（1997）．『これでいいのか大学入試英語』（上下巻）．東京：大修館書店。


Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language acquisition* (pp. 20-34). Cambridge: Cambridge University Press.


Motivation for Learning English as a Foreign Language in Japanese Elementary Schools

Junko Matsuzaki Carreira
Tsuda College

This study investigated how 345 Japanese elementary school pupils’ intrinsic and extrinsic motivation for learning English changed with age (174 third and 171 sixth graders). Factor analysis identified five underlying factors: interest in foreign countries, intrinsic motivation, caregivers’ encouragement, instrumental motivation, and anxiety. The results of an ANOVA showed significant differences in intrinsic motivation, interest in foreign countries, and instrumental motivation between the third and sixth graders. The third graders’ mean scores were higher than those of the sixth graders. This study revealed a rather steady developmental decline in intrinsic and extrinsic motivation for learning English, which might be attributed to general development trends in contemporary Japanese elementary school pupils. Within a consensus that there is considerable room for improvement in primary–school English education in Japan, the results of the present study suggest that the area of motivation can shed light on how the teaching methods for elementary school students in the higher grades can be improved.

本研究は、日本の小学生英語学習者の英語学習に対する内発的・外発的動機が年齢によりどのように変化していくかを調べたものである。調査参加者は、3年生174名、6年生171名、合計345名である。質問紙で得たデータを因子分析した結果、外国に対する興味、内発的動機、親の励まし、道具的動機、不安の5つの因子が抽出された。さらに分散分析の結果、外国に対する興味、内発的動機、道具的動機に有意差がみられ、どれも3年生の平均値が6年生の平均値よりも高かった。この結果から、学年があがるほど英語学習に対する内発的動機と外発的動機が低下していることがわかる。これは今の小学生の一般的な発達的傾向に原因があるという推測も成り立つが、特に、本論文では日本の初等英語教育においても改善すべき点があるとの立場から、高学年の指導法の改善の必要性を示唆した。
Many people consider motivation, which refers to “the process whereby goal-directed activity is instigated and sustained” (Pintrich & Schunk, 2002, p. 5), to be one of the most important factors involved in learning English. Gardner and Lambert (1972) suggest that motivation can be divided into two types: integrative motivation, referring to positive attitudes and feelings toward the target language group, and instrumental motivation, referring to the potential utilitarian gains of second language (L2) proficiency, such as getting a better job or higher salary. By the 1990s Gardner’s motivation theory had overwhelming dominance in L2 motivation research (Dörnyei, 2001). Some researchers of second language acquisition (SLA) have however started to pay attention to mainstream motivation psychology, incorporating psychological and “education-friendly” motivation research into their own (Dörnyei, 2001).

The difference between intrinsic motivation, referring to “motivation to engage in an activity for its own sake” and extrinsic motivation, referring to “motivation to engage in an activity as a means to an end” (Pintrich & Schunk, 2002, p. 245), is another well-known distinction in psychological motivational theory (Dörnyei, 2001). Brown (2000) suggests that intrinsic and extrinsic factors can be easily identified in foreign–language classrooms regardless of differences between cultural beliefs and the attitudes of learners and teachers. Although research into intrinsic and extrinsic motivation has been systematically introduced into SLA by Noels, Pelletier, Clément, and Vallerand (2000), to date there has been little research into children’s intrinsic and extrinsic motivation for learning foreign languages.

At elementary schools across Japan, English classes are being conducted to promote international understanding under the aegis of the Ministry of Education, Culture, Sports, Science, & Technology’s [MEXT] “Period for Integrated Study” programme (Ministry of Education, 2001). English activities were conducted in about 92.1% of all Japanese public elementary schools in 2004 (Ministry of Education, 2005). Although several researchers (e.g., Higuchi, Kunikata, Miura, Kitamura, Nakamoto, & Moriya, 1994; Takada, 2003) have conducted a few motivational studies of English as a foreign language (EFL) in Japanese elementary schools, most studies have focused only on the effect of childhood English learning. There have been very few attempts to reveal the motivational factors underlying Japanese elementary school pupils’ EFL learning.

Harter (1981), Lepper, Sethi, Dialdin, and Drake (1997), and Sakurai and Takano (1985) found that pupils’ intrinsic motivation for learning in
general decreases with age. It is often said that students in lower grades seem to participate in English lessons actively through playing games and singing songs, but students in upper grades tend less to enjoy such activities. However, there has been little research into developmental trends in motivation for learning EFL. Thus, the purpose of the present study is to clarify several underlying factors behind Japanese elementary school pupils’ motivation for learning EFL and their developmental trends.

**Background**

**Intrinsic and Extrinsic Motivation**

Defining intrinsic and extrinsic motivation has always been very controversial with several definitions being discussed to date (Sansone & Harackiewicz, 2000). First, Heider (1958) introduced perceived locus of causality, referring to actions or outcomes which can be perceived either as being personally caused or the result of impersonal causes. Personal causality refers to “instances in which P causes x intentionally” (p. 100). For example, people who have personal causality act on their own with their goal being to get an object. Impersonal causality refers to instances in which “P may cause x unintentionally merely because his or her physical or social being exerts some influence on the environment” (p. 100). For example, an outburst of anger may cause unintended displeasure. Thus, impersonal causality is often enacted unconsciously.

deCharms (1968/1983) expanded Heider’s concept and proposed the concepts of origin and pawn. “An Origin is a person who perceives his behavior as being determined by his own choosing; a Pawn is a person who perceives his behavior as being determined by external forces beyond his control” (p. 273). An origin is intrinsically motivated, whereas a pawn is extrinsically motivated (deCharms, 1968/1983).

Building upon these concepts, Deci and Ryan (1985) and Ryan, Connell, and Deci (1985) developed their self-determination theory, which suggests that under certain conditions there can be a shift from extrinsic control to self-regulation: internalization, referring to “the process by which children assimilate the socializing environment and accommodate to its demands and affordances” (Ryan et al., 1985, p. 33).

Secondly, Kruglanski (1975) introduced endogenous-exogenous attribution, where endogenous action refers to “an end in itself” (p. 390). For example, people who have endogenous attribution learn English for the sheer enjoyment of studying and knowing English; they have no other goal. On the other hand, exogenous action refers to “a means that medi-
ates a further goal, one exogenous to it” (Kruglanski, 1975, p. 390). For example, persons who have exogenous attribution study English for external reasons, such as for careers and entrance examinations. Endogenous action is thereby linked with intrinsic motivation and exogenous action with extrinsic motivation.

In SLA, Noels et al. (2000) developed the Language Learning Orientations Scale-Intrinsic Motivation, Extrinsic Motivation, Amotivation Subscale (LLOS-IEA), on the basis of the self-determination theory introduced by Deci and Ryan (1985). Noels et al. investigated the relations between the items of LLOS-IEA and the four orientations discussed by Clément and Kruidenier (1983), including instrumental orientation, friendship orientation, travel orientation, and knowledge orientation. Noels et al. found that instrumental orientation was associated with external regulation, whereas travel, friendship, and knowledge orientations were correlated more with self-determined and intrinsic motivation. That is, Noels et al. considered the relationships between intrinsic/extrinsic and integrative/instrumental motivation from the viewpoint of self-determination theory.

On the other hand, Schmidt, Boraie, and Kassabgy (1996) developed a questionnaire including concepts from cognitive and educational psychology. They defined extrinsic motivation as motivation to obtain an external reward and intrinsic motivation as motivation to get sufficient rewards from the activity itself. They went on to say that both instrumental and integrative motivation can be seen as subtypes of extrinsic motivation because both are related to goals and outcomes. Thus, it might be said that Schmidt et al. devised their questionnaire based on endogenous-exogenous attribution, although they did not specifically mention this.

Several lines of research which show developmental changes in intrinsic and extrinsic motivation have been conducted. Harter (1981) examined intrinsic and extrinsic motivation for studying in general in Connecticut, New York, Colorado, and California. Over 3,000 pupils (third through ninth graders) participated. She found a decline in intrinsic motivation in third through ninth graders. Harter investigated five subscales: challenge, curiosity, mastery, judgment, and criteria, which she defined as a “preference for challenge versus preference for easy work, curiosity/interest vs. teacher approval, independent mastery attempts vs. dependence on the teacher, independent judgment vs. reliance on the teacher’s judgment, and internal vs. external criteria for success/failure” (p. 300). Harter found that students’ responses on the challenge, curiosity, and mastery subscales changed from intrinsic to extrinsic motivation.
with age. In contrast, there was a shift from extrinsic to intrinsic motivation on the judgment and criteria subscales.

Harter’s questionnaire forced children to decide which of the options was more true for them. There were two sentences in each item: one based on intrinsic motivation and the other on extrinsic motivation. An example, one of the items had the following sentence for its intrinsic choice: “Some kids know whether or not they’re doing well in school without grades” (p. 305). For its extrinsic choice the item had: “Other kids need to have grades to know how well they are doing in school” (p. 305). The children were asked to decide which kind of child they were like and were then asked whether this description was only partially true or completely true for them. Each item was scaled ranging from 1 indicating maximum extrinsic motivation to 4 indicating maximum intrinsic motivation.

Harter (1981) assumed that intrinsic and extrinsic motivation were negatively correlated. Some researchers (Lepper et al., 1997; Lepper & Henderlong, 2000; Pintrich & Schunk, 2002) have started to cast doubt on Harter’s results, however. For example, Lepper et al. (1997) tried to replicate Harter’s research. Harter’s scale forced children to make a choice, but Lepper et al. modified the scale. The modified scale allowed children to answer intrinsic and extrinsic items independently. That is, the Lepper team assigned each of Harter’s items to separate scales for intrinsic and extrinsic motivation, each with a five-point Likert scale. Using these modified scales, Lepper et al. found a decrease in intrinsic motivation with age, but no evidence of an increase in extrinsic motivation. That is, intrinsic and extrinsic motivation proved far from negatively correlated. These results seem to show that intrinsic and extrinsic motivation can operate independently (Pintrich & Schunk, 2002). Intrinsic and extrinsic motivation are not two ends of a continuum but separate continua, each ranging from high to low (Pintrich & Schunk, 2002).

Following Harter (1981), Sakurai and Takano (1985) administered a questionnaire to 486 students in Japan: second–graders through seventh–graders. Sakurai and Takano changed several items, including those dealing with perceived locus of causality, endogenous-exogenous attributions, and enjoyment, as well as items addressing curiosity, challenge, and mastery. Using factor analysis (30 items), they identified six factors: curiosity, causality, enjoyment, mastery, challenge, and attribution. Sakurai and Takano found three types of developmental trends. First, the curiosity, causality, and enjoyment subscales declined gradually from second through fifth grades but increased in sixth grade and decreased again in seventh grade. Second, with age, the mastery and
challenge subscales decreased. Third, the attribution subscale increased with age. Although there are slight differences among their results and ideas, Harter (1981), Lepper et al. (1997), and Sakurai and Takano (1985) all found that intrinsic motivation decreases with age.

Motivation of Elementary School Pupils for Learning Foreign Languages

It is crucial to reveal what motivational differences exist between motivation of adults and children. Although for learning foreign languages, there has been a lot of research into the motivation of adults and adolescents, research into the motivation of elementary school pupils for learning foreign languages has been hard to find. Children are different from adults in terms of various factors such as cognitive skills and physical maturity. We cannot directly apply to children the results of motivational research on adults.

Masgore, Bernaus, and Gardner (2001) examined the attitudes and motivation of 499 Spanish children ages 10 through 15 who were studying English in Spain. In their study, they developed a children’s version of the mini-Attitude/Motivation Test Battery (ATMB) because the items of the AMTB and mini-AMTB were difficult for young children and the administration time was too long. The Attitude/Motivation Test Battery (AMTB) and a shorter version (the mini-AMTB) were originally developed for older school-aged students and adults. A factor analysis was conducted and after varimax rotation, a five-factor solution was chosen. The factors were as follows: attitudes toward English, orientation to learn English, English proficiency, age, and parental encouragement. According to Masgore et al. (2001), there was “no clear indication of independent clusters reflecting Integrativeness, Attitudes Toward the Learning Situation, Motivation, or Language Anxiety” (p. 291) and the factors that underlie motivation of children for learning EFL were significantly different from those of adults.

Nikolov (1999) reported on EFL motivation of 6- to 14-year-old children in Hungary in what was called the Pécs project. Three studies were conducted in 1977, 1985, and 1987 respectively and followed three groups of children over their first eight years of schooling. A questionnaire consisting of the same six open-ended questions was used during the eight-year period. The results of the first question (Why do you learn English?) were grouped into four types: classroom experience, teacher, external reasons, and utilitarian reasons. Nikolov found that the 8- to 11-
year olds had more external reasons to learn English than the 6- to 8-year
olds and tended to look ahead more into the future. The 11- to 14-year
olds stated more utilitarian reasons than the younger children. Nikolov
said that integrative motivation was not found in the answers. Nikolov’s
study showed some developmental trends of motivation for learning
EFL, which is very informative. However, the number of participants
was small and the research was done in Hungary, where, like Japan, the
first language is not related to any spoken in neighboring countries but
where the necessity, the choice, and the opportunity to learn foreign lan-
guages are all quite different from Japan. Thus, the research conducted by
Nikolov might not be directly applicable to a Japanese context.

In Japan, there have been a few studies (e.g. Higuchi et al. 1994; Takada,
2003) related to elementary school pupils’ motivation for learning EFL.
Most of them investigated the effect of early childhood English learning
and were conducted on junior high, high school, and college students,
ot elementary school pupils. Harter (1981), Lepper et al. (1997), and
Sakurai and Takano (1985) found that intrinsic motivation for studying in
general declines gradually with age, but there have been no such studies
on EFL pupils in Japan. Therefore, the present study explores the factors
underlying Japanese elementary school pupils’ motivation for learning
EFL and their developmental trends. It is hoped that the results of this
study will shed new light on the concept of motivation for learning EFL
among Japanese elementary school pupils and will expand its scope and
implications.

**Research Hypotheses**

Generally speaking, extrinsic motivation means external causality
based on the perceived locus of causality, which was suggested by Heider
(1958), deCharms (1968/1983), and Deci and Ryan (1985). However, at
the moment, most Japanese elementary schools do not have homework
or regular tests. Because pupils seem to feel little external pressure, it
may not be appropriate to discuss external causality. Thus, this study
investigates only the endogenous and exogenous attributions suggested
by Kruglanski (1975). Referring to Schmidt et al. (1996), the present study
defines intrinsic motivation as motivation to get sufficient rewards from
the activity itself and extrinsic motivation as motivation to obtain an ex-
ternal reward.
**Hypothesis 1**

As has been mentioned above, Harter (1981) found a decline in intrinsic motivation related to challenge, curiosity, and mastery with age. Lepper et al. (1997) also found a decrease in intrinsic motivation with age. In Japan, Sakurai and Takano (1985) found that although curiosity, causality, and enjoyment increased temporarily in sixth grade, these gradually declined with age. Hence,

Hypothesis 1: Third graders have more intrinsic motivation for learning EFL than sixth graders.

**Hypothesis 2**

As has been discussed above, Nikolov (1999) revealed that instrumental motivation emerged at age 11 or 12, and that the 11- to-14-year olds suggested more utilitarian reasons for learning English than the younger children. Therefore,

Hypothesis 2: Sixth-graders have more extrinsic motivation for learning EFL than third graders.

**Method**

**Participants**

Third (8 to 9 years old) and sixth (11 to 12 years old) graders in two public elementary schools (N and K) and one private school (S) were selected, for a total number of 354. Individuals with missing values were excluded and full analyses were performed on 345 participants: 174 third and 171 sixth graders. There were 193 boys and 161 girls (one respondent did not answer the question on gender). Teachers in each school signed an informed consent statement assuring each participant of anonymity and confidentiality.

**K Elementary School**

There are 425 pupils in K Elementary School. One hundred and thirty-four pupils, 73 third- and 61 sixth- graders, participated in the study. English lessons are provided for pupils in the first through sixth grades. They have an English lesson once a week. English lessons are given by an assistant language teacher (ALT) or a Japanese teacher of English (JTE)
together with their homeroom teacher (HT). Their main English activities are playing games and singing songs. In addition, English instruction was adopted during music lessons once a week.

**N Elementary School**

There are 307 pupils in N Elementary School. One hundred and two pupils, 51 third and 51 sixth graders, participated in the study. English lessons are provided for pupils in first through sixth grade. They have a 20-minute English lesson four times a week. English lessons are team-taught by the HT and ALT. Students learn English through games, songs, and stories. In addition, English instruction has been adopted during music and physical education lessons.

**S Elementary School**

There are 360 pupils in S Elementary School. One hundred and nine pupils, 50 third- and 59 sixth-graders, took part in the study. English lessons started 14 years ago. English lessons are provided for pupils in third through sixth grade and are taught by a JTE twice a week. In addition, fifth- and sixth-graders have English lessons taught by an ALT once a week. Although they learn to read and write in English, they primarily learn English through games, songs, and stories. They have regular English tests and get a grade at the end of each term.

**Pilot Study**

A pilot study was conducted in July 2003 with 41 elementary school pupils who attended a private English school in Iruma City, Saitama. Through observing the respondents and listening to their questions, the wording of the items on the questionnaire was modified and revised for the present study.

**Instruments**

This study employed a survey research design using a questionnaire. Questionnaires have been developed for adult and adolescent EFL learners (e.g., Dörnyei, 1990; Schmidt et al., 1996) and in Japan, questionnaires have been developed to study the motivation of Japanese EFL learners (e.g., Kimura, Nakata, & Okumura, 2001). The studies conducted in Japan are informative and helpful in creating new motivational questionnaires for EFL in Japanese situations. However, most of these questionnaires are
for adults and adolescent students and are not suitable for elementary school pupils. Several researchers have investigated the motivation of elementary school pupils for learning languages. There has, however, been no research in Japan and little elsewhere focusing on children’s intrinsic and extrinsic motivation for learning foreign languages.

Thus, the author created the Motivation and Attitudes toward Learning English Scale for Children (MALESC) based on the body of literature regarding motivation in SLA and intrinsic and extrinsic motivation in psychology, referring in particular to Schmidt et al. (1996), Sakurai and Takano (1985), and Gardner (1985). There are 19 items on the questionnaire, including items on motivation to learn English, interest in foreign countries, caregivers’ encouragement, and anxiety (see Appendix). Some researchers suggest using an even number of response options because some respondents might use the middle category, (i.e., “neither agree nor disagree”, “not sure”, or “neutral”) too often (Dörnyei, 2003). Thus, this study used a four-point Likert scale. Each response option was assigned a number for scoring purposes: strongly agree = 4, agree = 3, disagree = 2, strongly disagree = 1.

**Procedure**

The data was collected in September 2003. In order to reduce bias that could be created across the different schools, a tape recorded by the researcher was used to give instructions during data collection sessions. The pupils listened to the tape after the teachers distributed the questionnaires. The total administration of the questionnaire lasted between 10 and 15 minutes.

**Data Analysis**

The collected data was analyzed using the Statistical Package for the Social Sciences (SPSS) computer program. The mean and standard deviations (SD) of the third, sixth, and both the third- and sixth-graders together were reported for each question. In the present study, exploratory factor analysis was used for extracting underlying factors behind the motivation of elementary school pupils. Based on the results of this analysis, subscales were created by adding the scores for the items within each factor. In order to test hypotheses 1 and 2, a one-way ANOVA was used with SPSS.
General Description

The total number of subjects in the present study was 345, 174 third and 171 sixth graders. The mean and SD of individual items in the questionnaire of the third, sixth, and third-and-sixth graders are indicated in Table 1.

Factor analysis was used for extracting underlying factors behind motivation of elementary school pupils. After varimax rotation, five factors with eigenvalues over 1.00 emerged. The five factors accounted for 66.27% of the total variance. The items loading highest on each factor are listed in Table 2.
Table 2. Factor Loadings for 19 Questionnaire Items

<table>
<thead>
<tr>
<th>Item Description</th>
<th>Factor loading</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor 1: Interest in Foreign Countries (α = .86)</strong></td>
<td></td>
</tr>
<tr>
<td>2. I would like to go to various foreign countries.</td>
<td>.82</td>
</tr>
<tr>
<td>4. I would like to make a lot of foreign friends.</td>
<td>.77</td>
</tr>
<tr>
<td>12. I would like to try and talk to foreigners when my English becomes proficient.</td>
<td>.73</td>
</tr>
<tr>
<td>15. I would like to live abroad.</td>
<td>.77</td>
</tr>
<tr>
<td>17. I would like to know more about foreign countries.</td>
<td>.67</td>
</tr>
<tr>
<td><strong>Factor 2: Intrinsic Motivation (α = .84)</strong></td>
<td></td>
</tr>
<tr>
<td>1. English lessons are great fun. (I really enjoy learning English.)</td>
<td>.85</td>
</tr>
<tr>
<td>3. I always look forward to the day when we have English class.</td>
<td>.87</td>
</tr>
<tr>
<td>6. I would like to try to use the English which I have learned.</td>
<td>.57</td>
</tr>
<tr>
<td>8. I hope that we have more English lessons.</td>
<td>.81</td>
</tr>
<tr>
<td><strong>Factor 3: Caregivers’ Encouragement (α = .78)</strong></td>
<td></td>
</tr>
<tr>
<td>9. In my family, we all feel that it is very important to learn English.</td>
<td>.86</td>
</tr>
<tr>
<td>13. My parents hope that my English will be proficient.</td>
<td>.87</td>
</tr>
<tr>
<td>18. My parents tell me to study English hard.</td>
<td>.71</td>
</tr>
<tr>
<td><strong>Factor 4: Instrumental Motivation (α = .72)</strong></td>
<td></td>
</tr>
<tr>
<td>7. I study English in order to make English easier for me in junior high school.</td>
<td>.81</td>
</tr>
<tr>
<td>11. I study English because I think English will be necessary for me when I am an adult.</td>
<td>.63</td>
</tr>
<tr>
<td>14. I am studying English for a future job.</td>
<td>.45</td>
</tr>
<tr>
<td>19. I am studying English in order to enter a high school or a university.</td>
<td>.82</td>
</tr>
<tr>
<td><strong>Factor 5: Anxiety (α = .64)</strong></td>
<td></td>
</tr>
<tr>
<td>5. I get worried when I am doing worse than my classmates in English class.</td>
<td>.79</td>
</tr>
<tr>
<td>10. I am somehow always anxious in the English class.</td>
<td>.75</td>
</tr>
<tr>
<td>16. I get nervous when I answer or give a presentation in English class.</td>
<td>.75</td>
</tr>
</tbody>
</table>

*Note.* Item descriptions can be found in Appendix.
Factor 1 seems to address feelings towards foreign countries or foreigners and can be labeled Interest in Foreign Countries. The items loading on Factor 2 represent intrinsic motivation and can be labeled Intrinsic Motivation. The items loading on Factor 3 primarily ask how the caregivers feel about their children learning English and thus can be labeled Caregivers’ Encouragement. All of the items that load on Factor 4 ask about the practical reasons for studying English. Factor 4 can be called Instrumental Motivation. Factor 5 clearly represents anxiety and can be called Anxiety.

Based on the results of factor analysis, the scores for the items within each factor were added up to create the five subscales. Cronbach alpha for each subscale ranged from .65 to .86. Descriptive statistics for the five subscales are presented in Table 3.

Table 3. Descriptive Statistics for Five Subscales

<table>
<thead>
<tr>
<th>Factor</th>
<th>Min</th>
<th>Max</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest in Foreign countries</td>
<td>5.00</td>
<td>20.00</td>
<td>15.60</td>
<td>4.07</td>
</tr>
<tr>
<td>Intrinsic motivation</td>
<td>4.00</td>
<td>16.00</td>
<td>12.87</td>
<td>2.90</td>
</tr>
<tr>
<td>Caregivers’ encouragement</td>
<td>3.00</td>
<td>12.00</td>
<td>8.08</td>
<td>2.64</td>
</tr>
<tr>
<td>Instrumental motivation</td>
<td>4.00</td>
<td>16.00</td>
<td>12.24</td>
<td>2.99</td>
</tr>
<tr>
<td>Anxiety</td>
<td>3.00</td>
<td>12.00</td>
<td>7.16</td>
<td>2.63</td>
</tr>
</tbody>
</table>

Table 4 shows the means and SD for the five subscales of the third and sixth graders.

Table 4. Means and SD for Subscales of Third- and Sixth-Graders

<table>
<thead>
<tr>
<th>Factor</th>
<th>Third-graders</th>
<th>Sixth-Graders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Interest in foreign countries</td>
<td>16.46</td>
<td>3.75</td>
</tr>
<tr>
<td>Intrinsic motivation</td>
<td>13.82</td>
<td>2.48</td>
</tr>
<tr>
<td>Caregivers’ encouragement</td>
<td>8.19</td>
<td>2.74</td>
</tr>
<tr>
<td>Instrumental motivation</td>
<td>12.98</td>
<td>2.77</td>
</tr>
<tr>
<td>Anxiety</td>
<td>7.30</td>
<td>2.79</td>
</tr>
</tbody>
</table>
An ANOVA (see Table 5) was run in order to reveal differences between the third and sixth graders on each subscale. The results of the ANOVA showed significant differences in interest in foreign counties, intrinsic motivation, and instrumental motivation between the third- and sixth-graders. There were no statistically significant differences between the third- and sixth-graders in caregivers’ encouragement or anxiety.

**Table 5. One-Way Analyses of Variance for Effects of Grades on Five Dependent Variables**

<table>
<thead>
<tr>
<th>Variable and source</th>
<th>SS</th>
<th>MS</th>
<th>$F$ (1, 343)</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest in foreign countries</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>261.25</td>
<td>261.25</td>
<td>16.47***</td>
<td>.046</td>
</tr>
<tr>
<td>Within groups</td>
<td>5441.75</td>
<td>15.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intrinsic motivation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>316.41</td>
<td>316.41</td>
<td>42.00***</td>
<td>.109</td>
</tr>
<tr>
<td>Within groups</td>
<td>2583.98</td>
<td>7.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caregivers’ encouragement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>4.13</td>
<td>4.13</td>
<td>.59</td>
<td>.002</td>
</tr>
<tr>
<td>Within groups</td>
<td>2395.60</td>
<td>6.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instrumental motivation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>194.88</td>
<td>194.88</td>
<td>23.21***</td>
<td>.063</td>
</tr>
<tr>
<td>Within groups</td>
<td>2879.63</td>
<td>8.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>6.54</td>
<td>6.54</td>
<td>.95</td>
<td>.003</td>
</tr>
<tr>
<td>Within groups</td>
<td>2376.37</td>
<td>6.93</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. $\eta^2$ = effect size.*

***$p < .001$
Test of Hypothesis 1

The results of intrinsic motivation using ANOVA (see Table 5) showed that there was a statistically significant difference between the third and sixth graders, \( F(1, 343) = 42.00, p < .001 \), although the effect size was small (partial \( \eta^2 = .109 \)). Table 4 shows that the mean of the scores of the third graders in intrinsic motivation (13.82) was higher than that of the sixth graders (11.91). Therefore, Hypothesis 1 has been supported (third graders have more intrinsic motivation for learning EFL than sixth graders).

Test of Hypothesis 2

As Schmidt et al. (1996) stated, interest in foreign countries and instrumental motivation can be considered as subtypes of extrinsic motivation. The results of interest in foreign countries using ANOVA (see Table 5) showed that there was a statistically significant difference between the third and sixth graders, \( F(1, 343) = 16.47, p < .001 \), although the effect size was small (partial \( \eta^2 = .046 \)). Table 4 shows that the mean of the scores of the third graders in interest in foreign countries (16.46) was higher than that of the sixth graders (14.72). The results of instrumental motivation using ANOVA (see Table 5) showed that there was a statistically significant difference between the third and sixth graders, \( F(1, 343) = 23.21, p < .001 \), although the effect size was small (partial \( \eta^2 = .063 \)). Table 4 shows that the mean of the scores of the third graders in instrumental motivation was higher (12.98) than that of the sixth graders (11.48). The third graders indicated more interest in foreign countries and greater instrumental motivation than the sixth graders. Contrary to Hypothesis 2, third graders have more extrinsic motivation for learning EFL than sixth graders.

Discussion

Hypothesis 1

Hypothesis 1, which stated that third graders have more intrinsic motivation for learning EFL than sixth graders, was supported. Just as Harter (1981), Lepper et al. (1997), and Sakurai and Takano (1985) found that intrinsic motivation for studying in general declines gradually with age, the present study on learning EFL also shows Japanese elementary school pupils’ intrinsic motivation for learning English as well as for learning in general seems to decrease with age.

Why does pupils’ intrinsic motivation decrease with age? When young children are given some interesting and exciting tasks by teachers and
parents, they often become absorbed in them. As pupils become more logical and have their own ideas as they get older, they may not enjoy activities set by others as much. Developmental decreases in intrinsic motivation may be inevitable to some degree.

**Hypothesis 2**

Hypothesis 2, which stated that sixth graders have more extrinsic motivation for learning EFL than third graders was not supported: instead, the opposite was revealed. In the present study, while there were significant differences between the third and sixth graders in extrinsic motivation (interest in foreign counties and instrumental motivation), it was the third graders who possessed higher extrinsic motivation. Based on these results, it could be said that not only intrinsic but also extrinsic motivation decreases with age. However, the effect sizes were small. There might be other factors which influence students. It is necessary to explore other potential factors in the future.

Nikolov (1999) revealed that instrumental motivation appeared in children 11 or 12 years old, and that the 11 to 14 year old group gave more utilitarian reasons than the younger groups. Here the younger groups may have had some latent instrumental motivation, but because they had much stronger classroom- and teacher-related motivation than instrumental motivation, instrumental motivation might not have been revealed in responses to the open-ended questionnaire. Moreover, as has already been suggested, there may well have been cultural and even political factors at work which do not directly pertain to the situation in Japan, so it is not unusual that the results were different from those of the present study.

Why does pupils’ instrumental motivation decrease with age? This may be happening not only in English learning, but also in their general learning. It has often been said that Japanese pupils tend to lose their goals and become unmotivated with age. This might be a sort of general trend in contemporary Japanese elementary school pupils.

Also, why do pupils appear to lose interest in foreign countries with age? Separating the learning of academic skills from real-world contexts can entail motivational costs (Lepper & Henderlong, 2000). In order to enhance communicative skills, listening and speaking have been emphasized, but students have few opportunities to use English with foreigners outside the classroom. When they are in lower grades, they can be satisfied with exciting and fun tasks. As they grow, their motivation might
decrease when they cannot connect the English they learn in class to real communication.

**General Discussion**

The present study identified five affective factors that were labeled: interest in foreign countries, intrinsic motivation, caregivers’ encouragement, instrumental motivation, and anxiety. The questionnaire appears to be a reliable and valid measure sensitive to individual differences in both intrinsic and extrinsic motivation. Masgoret et al. (2001) also found five factors: attitudes toward English, orientation to learning English, English proficiency, age, and parental encouragement. Some similar and different factors can be found in both studies, although it is necessary to be conservative when drawing comparisons because the questionnaires used and the ages of the participants were different. In both studies, parental or other caregiver encouragement emerged as one factor. This factor may be necessary for children.

In the present study, the factors of intrinsic motivation, instrumental motivation, interest in foreign countries, and anxiety were clustered into separate factors. On the other hand, Masgoret et al. (2001) did not find any clear clusters even though they have frequently been found in other studies using the AMTB. The AMTB and mini-AMTB were developed in Canada, where many need an L2 for their social life. As some researchers (e.g., Dörnyei, 1990; Schmidt et al., 1996) have argued, the different contexts of SLA and EFL is significant, and thus the mini-AMTB might not be suitable for EFL contexts such as Spain (Masgoret et al., 2001), or even Japan.

Harter (1981) assumes that intrinsic and extrinsic motivation may be negatively correlated. Lepper et al. (1997) found a decrease in intrinsic motivation with age, but no evidence of an increase in extrinsic motivation. The present study shows that there are some developmental declines in both intrinsic and extrinsic motivation. This suggests that extrinsic motivation is more susceptible to various factors than intrinsic motivation. Further, intrinsic and extrinsic motivation may not be two ends of a continuum but separate continua (Pintrich and Schunk, 2002).

The results of the present study have demonstrated a developmental decrease in motivation for learning EFL, which may be influenced by such external factors as education, teachers, parents, peers, and the classroom. These factors can enhance or lower pupils’ motivation. In other words, for students in higher grades, motivation for learning EFL can
be enhanced by providing an appropriate environment and improving teaching methods and materials. How, then, can teachers create an intrinsically motivating climate in English lessons for the upper grades? Lepper and Henderlong (2000) suggest several potential responses to developmental decline in intrinsic motivation (e.g., “promote autonomy and self-determination” and “promote children’s sense of curiosity by placing learning in meaningful and exciting contexts” [p. 289]).

How can we apply these suggestions to English classes for children in Japan? According to Paul (2003), in order to enhance English ability, pupils need to repeat patterns a number of times, but this should be done in a meaningful way, not as a dry classroom drill. Paul suggests, “the children first come across new targets inside an activity, and then later they may do some less game-like activities, such as writing sentences in their notebooks” (p. 51) and that inserting the new targets into games elicits a genuine and meaningful desire and need to learn, giving some space for pupils to think. He goes on say that pupils are likely to feel a sense of ownership of those new words and patterns. Ideas such as this one suggested by Paul might be part of the solution to combat developmental decline in motivation for learning EFL.

This study revealed a rather steady developmental decline in intrinsic and extrinsic motivation for learning EFL, which might be attributed to general development trends in contemporary Japanese elementary school pupils. There is, however, still considerable room for improvement in primary English education in Japan. Before English is formally established in elementary school education, more research about what sorts of activities and materials are appropriate for the upper grades should be conducted.

**Conclusion**

This study has revealed developmental declines in intrinsic and extrinsic motivation of pupils for learning EFL, which could be happening in other subjects, too. Factor analysis identified five underlying factors behind the motivation of pupils: interest in foreign countries, intrinsic motivation, caregivers’ encouragement, instrumental motivation, and anxiety.

There are several limitations to the present study. First, motivation may be influenced by various context-specific factors. While the motivation of grade-schoolers was fairly accurately revealed, the present results may be applicable only to the pupils investigated here. Second, the data
was collected only by means of a questionnaire. Qualitative research, including interviews with students and teachers, assessments of actual performance, and classroom observations, may significantly provide further layers of analysis. Third, data collection was cross sectional. Cross-sectional studies are less effective in identifying individual variations in growth or establishing causal relationships (Cohen, Manion, & Morrison, 2000). Longitudinal studies are particularly appropriate in research on human growth and development (Cohen et al., 2000). In order to investigate more precisely the change in motivation with age, longitudinal data collection should be conducted in the future.

Notwithstanding its limitations, this study does make a contribution towards revealing and understanding the motivation of Japanese elementary pupils for learning EFL. The present study implies that the teaching method for higher grades should be improved. This study narrowed its focus to motivational studies, which will certainly be an area that will form the foundation of elementary school English teaching in Japan. Other areas such as cognitive and linguistic development, strategies, and sociocultural factors will have to be investigated in the future.

Acknowledgements

My sincere thanks go to Dr. Hiroko Tajika for her generous and valuable guidance. I would like to thank Dr. Makoto Mori, Dr. Roger Barnard, Dr. Jonathan Piken, and Ms. Masae Konishi, for their valuable advice. I would also like to thank the editor and the two anonymous referees of JALT Journal for their advice on the revision of this paper.

Junko Matsuzaki Carreira is a Ph.D candidate at Tsuda College. She also teaches part time at Keiai Universiy, Tokyo Keizai University, and Takushoku Universiy. Her research interests include motivation and teaching children.

References


Appendix

Motivation and Attitudes toward Learning English Scale for Children (MALESC)

次の質間はあなたの英語学習にたいする今の気持ちをたずねるものでです。例のように最もあてはまる番号に○をつけてください。例：本を読むことは好きです。

はい（  ）まあまあ（ ○  ）あまり（  ）いいえ（  ）

1. 英語の授業はとてもたのしいです。
2. いろいろな外国にいってみたいのです。
3. 英語の授業のある日は楽しみです。
4. 外国のお友達をたくさん作りたいです。
5. 英語の授業でみんなよりできないと心配です。
6. 習った英語をもっと使ってみたいのです。
7. 中学校に入って英語の勉強で困らないように勉強します。
8. もっと英語の授業があったほうがいいです。
9. おうちの人は英語がとても大切だと思うています。
10. 英語の授業中はなんだかいつも心配です。
11. 大人がになったら自分にとって必要になると思うので、英語を勉強します。
12. 英語が上手になって外国の人と話してみたいのです。
13. おうちの人は私が英語ができるようになることを望んでいます。
14. 将来なりたい仕事のために英語を勉強します。
15. 外国に住んでみたいのです。
16. 英語の授業で答えたり、発表するときどきどきします。
17. 外国のことをもっと知りたいです。
18. おうちの人は英語を一生懸命勉強しなさいといいます。
19. 高校や大学に入るために必要なので英語を勉強します。

English Translation of MALESC

Q1. English lessons are great fun. (I really enjoy learning English.)
Q2. I would like to go to various foreign countries.
Q3. I always look forward to the day when we have English class.
Q4. I would like to make a lot of foreign friends.
Q5. I get worried when I am doing worse than my classmates in English class.
Q6. I would like to try to use the English which I have learned.
Q7. I study English in order to make English easier for me in junior high school.

Q8. I hope that we have more English lessons.

Q9. In my family, we all feel that it is very important to learn English.

Q10. I am somehow always anxious in the English class.

Q11. I study English because I think English will be necessary for me when I am an adult.

Q12. I would like to try and talk to foreigners when my English becomes proficient.

Q13. My parents hope that my English will be proficient.


Q15. I would like to live abroad.

Q16. I get nervous when I answer or give a presentation in English class.

Q17. I would like to know more about foreign countries.

Q18. My parents tell me to study English hard.

Q19. I am studying English in order to enter a high school or a university.
In this paper, we attempted to investigate the following issues: (a) how different learning environments may cause different interactions in the classroom; (b) how such differences may affect learners’ attitudes towards listening tasks and test results. We compared the results of tests and questionnaires in two types of learning environments: PC-equipped classrooms and regular classrooms.

Participants were all Japanese undergraduate university students majoring in English. The experimental group (PC-group) consisted of 116 participants (PCL: lower level group, 51; PCH: higher level group 65), while the control group (regular classroom-group) consisted of 117 participants (RGL: lower level group, 56; RGH: higher level group 61). Materials used in the present study were several series of VOA Special English news. The participants in the PC-group were allowed to listen to the materials recurrently using an application on the individual PC. The applica-
tation enabled real-time feedback of their performances and their individual ranks in the classroom for their tasks. By contrast, the participants in the regular classroom-group had recurrent but teacher-controlled chances to listen to the same materials and were given overall feedback for their performances. We conducted two types of the same paper tests and questionnaires in both the pre and post phases of this research project. ANOVA, multiple comparison Scheffe and Wilcoxon signed rank test with Bonferroni adjustment were used for analysis.

As for the fill-in-the-blank tests, the ANOVA repeated Measure detected a significant difference between the pre and post test scores \((F (1,231) = 184.77, p = .01)\), and a significant interaction between the two groups \((F (1,231) = 45.00, p = .01)\). The average score of the PC-group was significantly higher than that of the regular classroom-group. As for the TOEIC listening tests, the ANOVA repeated Measure detected a significant difference between pre and post test scores \((F (1,231) = 52.84, p = .01)\), but no significant interaction between the two groups. The significant interaction between two groups was detected only for the fill-in-the-blank tests. This means that learning environment with PCs may have impact on learners’ attitudes and improved results of listening tasks at perception level. In addition, the ANOVA repeated Measure among four groups (PCL, PCH, RGL, RGH) detected a significant difference between the pre and post test scores: the fill-in-the-blank tests \((F(3,229)= 15.231, p=.01)\), the TOEIC listening tests \((F(3,229)= 18.10, p=.01)\). Multiple comparison Scheffe revealed significant interactions among the following tests: the fill-in-the-blank tests: PCH and RGL, RGL and PC/RGH; the TOEIC listening tests: PCL and RG/PC H, PCH and RG/PC L, RGL and RG/PC H, RGH and RG/PC L. This means that learning environment with PCs may have influence on listening performance at perception level (rather shallow cognitive processing level) but not at comprehension level (rather deep cognitive processing level).

Wilcoxon signed rank test with Bonferroni adjustment detected significant differences on the pre and post questionnaires as follows: two items for the PC-group and eight items for the regular classroom-group. This means that learners’ awareness of their listening skills and learning strategies did not affect the results of the tests. In other words, it was different learning environments that caused differences in interactions in the classroom. In conclusion, learning environment may have affected learners’ learning attitudes. Focusing on fewer listening skills and learning strategies in the PC-equipped classroom had a positive impact on the scores of the listening perception test. The learning environment with PCs might have facilitated learner autonomy.

本研究では、学習活動の場である教室内に備えられた人工物（PC）の有無により、学習者が経験する学習行為に差が生じるのか、また、学習者のリスニング力伸長や学習意識に差異をもたらすのかを調査する目的で、同じ教材で同じ教育目標によるPC設置有無の二条件（PC教室と普通教室）での授業を実施し、その影響を検証した。教室環境の差により「課題遂行」と「結果表示」の方法などにおいて教師－学習者間の役割遂行配分が異なり、PC教室では学習者が中心型の授業を展開し、普通教室ではより教員主導型の授業を展開した。この授業展開の差異がリスニング学習の意識変化に影響を及ぼし、結果としてリスニングの知覚レベル（深い意味判断を要求しない）でのテスト得点増加に影響が
KURAMOTO, YOSHIDA, & YOSHIDA

あり、PC教室群が有意に高かった。教室環境固有の条件－PC教室では自己ベース学習が可能である状況、普通教室では教師主導が生み出す状況－がそれぞれの学習者の着目するリスニング技術や学習方略に影響を与えたものと考えられる。

研究の背景

本研究では、言語学習における学習効果を考察する上で、大きく二つの視点からの調査分析を行った。一つは、学習者要因の視点から、自律学習に影響を与えるとされる学習方略に対する学習者の意識を調査することであった。他方は、学習環境の差異がもたらす影響を調査した。教育における大きな二つの枠組みを概観することで、本研究で述べる学習環境の差異が学習過程の何に影響を与えるかについて述べた。

Oxford（1990,7-10）は学習方略と学習者の意識との関連について次のように述べている。学習方略は、学習者自らによる情報の習得、記憶、検索、利用などを行い易くする作業である。言語学習における言語の処理、保持、伝達能力などの伸張には欠かすことのできない要素で、広義には、学習をより簡単に、履く、楽し、自主的に、効率良く、場面に応じて応用できるようになるために取る学習者の行為である。ところが、学習者がさまざまな方略を認識できず意識的な活用ができないため、自身が言語学習への適正を欠いているという判断にとどまり、学習を詰めより良い学習結果に結びつかない場合がみられる。Oxfordによる以上の指摘から、学習者の自己の学習における学習方略に対する意識のあり方が、学習の成果や自律学習の成り立ちに影響をもたらすものと考えられる。


一方、授業あるいは教育における大きな枠組みとしては、次の二つの形態がある。ひとつは、産業革命以降の近代学校成立の背景ともなり、また「Back to Basics」という1970年代にアメリカを中心に起こった教育運動にも見られる、
伝統的な知識を習得させることを基本とする教師主導のリテラシー教育である。他方は、Rousseau（1911）、Pestalozzi（1907）、Dewey（2001）などの児童中心主義と呼ばれる教育思想に根ざした学習者中心型の授業である。市川（1995）によれば、実際の教育場面においては、この両者の考え方が明確に質的に二分されるのではなく、「（1）教育目標の設定（2）具体的な学習課題の設定（3）課題遂行方法の選択と実施（4）結果の評価」という四つの過程の教師－学習者間の遂行配分により、かなりのバリエーションが見られると指摘している。これらのことから、同じ学習内容について、2つの異なる学習環境で異なる授業を行なった場合、教師－学習者間の役割遂行配分が異なるため、結果として学習効果に何らかの差が現れるのではないかと本研究筆者らは考えた。

また、1980年代半ばから認知理論の一つとして注目が寄せられている状況論的学習論の立場では、知識は個人の頭脳の中に留まるものではなく、実際に生活をするコミュニティにおけるアーチファクト（人工物）や同じ時間空間に生活する家族・学校・会社など、つまりコミュニティの人々との相互作用において集積されるという（Wenger, 1998）。「相互作用」とは、コミュニティでの実生活を通して、周りの人工物や人々と学習者の関係に変化が起こり、学習者が自身がどのようなかが学習の効率につながり、学習環境の差異が学習効果に影響を与えるとされている。学校での授業においては、一般社会におけるようなあらゆる面での「相互作用」の選択幅ではなく、一定の学習環境に教えられる。その一定の学習環境においての特有の相互作用が起き、知識や技能の変化が始まる。

本研究で述べる「学習環境」とは、状況論で述べる「相互作用」が起こる場としての教室のコミュニティを指す。人工物とは、机やいす、黒板、CDプレーヤー、コンピュータ（以下PC）などさまざまな道具を含むが、本研究では特に現代の社会的要因を反映した教室内のPCの有無に着目した。人工物の一例であるPC利用の有無により条件の異なる教室コミュニティにおいて、これらの学習環境の差異が学習者の意識のあり方などに何らかの影響を与え、学習効果に差異をもたらすのではないかと考え、以下の研究を行った。

研究の目的と調査事項

本研究では、教室コミュニティという学習活動の場で学習者が経験する他者（教授者）や人工物（PC）との相互作用の差異、つまり学習環境の差異が、学習成果に影響するのかを調査することを目的とした。PC教室と普通教室という二条件下での授業展開で、リスニング課題遂行時に学習者の利用するリスニングスキルや学習方略に対する意識に差異をもたらすのか、また差異があるとすればどのような影響があるのかについて調査した。同じ教師が、同じ教材で同じ教育目標により、学習環境の異なるPC教室と普通教室という二条件下的授業を実施し、その影響を検証するため、量的データによる分析を行った。
研究の方法

英語専攻の日本大系生1年生2群（PC教室群：実験群116名（内下位群51名、上位群65名）、普通教室群：対照群117名（内下位群56名、上位群61名）、各2クラス編成、内男子124名、女子109名、2クラスのクラス内の男
女比がほぼ等しい、年齢18〜19歳）を対象とした。

教材・テスト・質問紙

VOA（Voice of America）Special Englishのリスニング素材から、同一のものを選んで教材とし、PC教室群では文字と音声ともにPCソフトを利用して、普通教室群では印刷物とCDに録音された音声を使って、取組み課題を実施した。紙ベ
ースによる事前・事後テスト二種類（リスニング空所補充テスト（付録A）、
TOEICリスニングテスト（TOEIC公式ガイド＆問題集pp.23-29、Part 1 〜 4）お
よびリスニング課題遂行時に学習者の利用するリスニングスキルと学習方略に
する意識の質問紙調査（表1）を4月第一週授業日と7月最終週授業日に実
施した。

質問紙調査は、24項目、6段階のリカートスケールで行った。質問紙24項目の信頼度係数はCronbach’s coefficient α=.92と高く、これらの項目は本調査に適
切であると判断した。6段階の評定尺度は下記の通りである。

1 全く当てはまらない。  2 ほとんど当てはまらない。
3 あまり当てはまらない。  4 少し当てはまる。
5 かなり当てはまる。      6 そのとおり当てはまる。

SILLで用いられている学習方略の分類は、記憶、認知、補償、メタ認知、感情、社会の6方略であるが、本研究では教室コミュニティでの教師やPCの有無
による人工物との相互作用の差異が学習意識などへの変化をもたらすかどうかを
調査するため、SILL、リスニングに関わる質問紙(Vogely，1995)、リスニン
グ力を構成する技術の分類（Rost, 1991, Omaggio, 1986）、
Vygotskyの概念発達論(Vygotsky, 1987, 1997a, 1997b)の4資料を参照し、新たに質問紙を作成した。
新たな項目の分類は以下の通りである。リスニングスキルについては、知覚技
術、分析技術、統合技術についての3分類とし、学習方略については、メタ認
知的方略、認知的方略、補償方略お3分類とし、本研究での学習環境の差異による
学習者の意識変化を読み取れる内容を質問項目として用意した。質問項目は表
1の通りである。

実施の手順

両群とも、「オンラインで学ぶVOAイングリッシュ・ライフ系総合英語」(吉田、他、2002)を教材として使用し、半期間（約4ヶ月間）教員1名により下
記の要領1〜3に従って授業を行った。両群とも聞き取りの後、当該単語を入力
するリスニング課題全体の制限時間および解答に関する質疑応答を含む解説時
間（提示方法は下記の通り異なる）の配分は、各ユニットの長さに応じて同じ
時間配分（30〜40分、30〜40分）とした。
表1：質問項目

<table>
<thead>
<tr>
<th>質問番号</th>
<th>項目</th>
<th>認知的方略</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>一語一義しっかり聞き取ろうとする。</td>
<td>認知的方略</td>
</tr>
<tr>
<td>2</td>
<td>話の展開を頭に描きながら聞くとする。</td>
<td>認知的方略</td>
</tr>
<tr>
<td>3</td>
<td>話の展開を予想して聞くとする。</td>
<td>認知的方略</td>
</tr>
<tr>
<td>4</td>
<td>トピックになじめがなければ聞き取りやすい。</td>
<td>認知的方略</td>
</tr>
<tr>
<td>5</td>
<td>視覚情報（絵、写真、図、漫画など）があれば、参考にして聞く。</td>
<td>認知的方略</td>
</tr>
<tr>
<td>6</td>
<td>英字母を母語とする人が自分の言語でゆっくり話して聞き取れる。</td>
<td>英聡的方略</td>
</tr>
<tr>
<td>7</td>
<td>教材のテープは速度がゆっくりのものであれば聞き取れる。</td>
<td>英聡的方略</td>
</tr>
<tr>
<td>8</td>
<td>多少長い話を最後まで聞き取ろうとする。</td>
<td>英聡的方略</td>
</tr>
<tr>
<td>9</td>
<td>多少解らないところがあっても、全体を最後まで聞くとする。</td>
<td>認知的方略</td>
</tr>
<tr>
<td>10</td>
<td>英語のまま聞き取り理解しようとする。</td>
<td>認知的方略</td>
</tr>
<tr>
<td>11</td>
<td>教材のメニューは何度も聞き返せば聞き取れる。</td>
<td>認知的方略</td>
</tr>
</tbody>
</table>

PC教室群（実験群）の要領：
1. 当該ユニット関連のインターネットサイト情報（背景知識の提供）を受け
   る。画面のテキストを見ずに一度聞き取る。
2. 当該ユニットを各自のペースで繰り返し聞き取り、オンラインで画面上に表
   示されたテキストの順に当該単語を入力する。入力後リアルタイムで解
   答と正答率が個別にフィードバックされる。その他に、正誤問題、語彙問
   題も同様にオンラインで行う。
3. サーバーに送られた各個人のデータに基づいて、その場で提示されるクラス
   内の順位、全員の解答状況およびニュース特有の表現などの解説を受け
   る。

普通教室群（対照群）の要領：
1. 当該ユニット関連の印刷された情報（背景知識の提供）を受ける。テキスト
   見ずに一度聞き取る。
2. 一斉に流れてくる音声を複数回聞きながら、テキストの順に適語を書き込
   んでいく。その他に、正誤問題、語彙問題も同様にテキストを見て行う。
3. 教師の口頭でのテキストの空所部分の答えあわせを受けながら、引き続きニュース特有の表現などの解説を受ける。
PC教室群では、自己ペースでの繰返し聞き・入力する学習を可能にするため、コンピュータ上で自由に聞きたい部分を簡単に再生できる音声再生ソフトと、インターネットブラウザの画面上でリアルタイムに解答が添削されフィードバックされるソフト（吉田、他、2002）が用意されていた。
一方、普通教室群では、教卓からの一括操作でCD音声が再生され、文単位ごとに停止して、複数回流される聞き取りで、印刷されたテキストの空所へ書き込んだ。教師の判断で学習者の多くが聞き取りに困難を感じているとみなした場合は、同じ空所部分を再度繰返すという、学習者個別への対応ではなく教室参加者一斉のベースであった。
紙ベースによる事前・事後テスト二種類（リスニング空所補充テスト（fill-in-the blank test以下FBT）、TOEICリスニングテスト（TOEIC listening test以下TLT））では、リスニング力の伸長度を測定した。実験群、対照群とも、同じ条件で教室に一斉に音声を流した。FBTは、学習者が初めて聞くニュース原稿に7語ごとの空所が50箇所あり、一斉に流される音声を追って、それらの空所に単語を書き取っていく課題であった。聞き取る素材がニュースであるため、背景知識などから補充すべき語を推測できる場合もあるが、ニュース全体を2度繰返し、見直しの時間を設けていない、その場での詳細な意味把握課題を課さない、という条件での音声聞き取り課題であり、機械的に7語ごとに内容語も機能語も入り混じった形での空所が設けられているため、純粋に音声に対する知覚レベルを測る性格の強いテストとなっていた。詳細な意味把握課題を課さないとということは、その場での深い意味理解を要求しないこと、つまり課題の処理負荷を下げているといえよう。本研究における二種類のリスニング課題のうち知覚レベルの素材としては妥当であると判断した理由は、以下の二点である。クローズ（Cloze）テストの一種であり、ESLにおいてCloze Dictationが通常のリスニングテストと.83の相関があると有効性が実証されている（Johansson, 1972）。VOAのSpecial English（付録A）は1500語レベルであり、文章全体のリーダビリティは、比較的平易（Flesch Reading Ease 59.3、Flesch-Kincaid Grade Level 7.5）である。TLTは、英語運用力を測るテストとして広く認められているTOEICテストの準備用として自己診断により現在の力を測り、本試験での予測点を算出するためのミニTOEICテストのリスニングセクション（国際ビジネスコミュニケーション協会TOEIC運営委員会, 2000）から28問を使用した。
結果と考察

テストの結果と考察

PC教室群と普通教室群のFBTおよびTLTの研究授業開始期の各テスト結果に有意な差はなかった。FBT (F(1,231)=.23, n.s.) TLT (F(1,231)=.46, n.s.)。

PC教室群と普通教室群のFBTおよびTLTの事前・事後テストの平均（素点）、SD、参加者数を表2に示す。

表2：群別の前後テストの平均値 (n=参加者数)

<table>
<thead>
<tr>
<th>テスト種類</th>
<th>PC教室群(n=116)</th>
<th>普通教室群(n=117)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
</tr>
<tr>
<td>リスニング空所補充(FBT)</td>
<td>29.71</td>
<td>36.94</td>
</tr>
<tr>
<td></td>
<td>(7.67)</td>
<td>(6.43)</td>
</tr>
<tr>
<td>TOEICリスニング(TLT)</td>
<td>11.97</td>
<td>13.65</td>
</tr>
<tr>
<td></td>
<td>(2.73)</td>
<td>(2.76)</td>
</tr>
</tbody>
</table>

括弧内は標準偏差、FBT 50点満点、TLT 28点満点

2群の反復測定による分散分析の結果、各事前・事後テスト間において、FBT、TLT共に有意差 (F(1,231)=184.77, p<.01)、(F(1,231)=52.84, p<.01) が認められた。2群間の教室条件と各事前・事後テストの得点変化は、FBTにおいてはPC教室群の方が多く、PC教室群と普通教室群の間に有意差 (F(1,231)=45.00, p<.01) (図1) が検出され、2群間の教室条件（PCの有無）と得点変化との間に交互作用が認められた。一方、TLTでは、2群間の教室条件（PCの有無）と得点変化との間には有意差がなく、交互作用は認められなかった (F(1,231)=.24, n.s.)。

図1. リスニング空所補充前後テスト（FBT）の推移 50点満点

さらに、本研究におけるリスニング課題が、成績上位群と下位群に同様の効
果をもたらすか否かを検討するため、以下の通り各群を上下2群に分割した。英文理解度の程度を上下2群に分割するため、開始期に行ったリスニングテストのうち、内容理解度を測るTLTを利用した。TLTの平均得点の11.83点を基準として、各群を上位群（12点以上）、下位群（11点以下）に分割したところ、4群の度数配分にも比較的偏りがなく、比較検討するのに適当であると判断した。PC教室群（下位群・上位群）と普通教室群（下位群・上位群）のFBTおよびTLTの事前・事後テストの平均（素点）、SD、参加者数を表3、表4に示す。また、一元配置分散分析（ANOVA）により群間に有意な差が認められた（F(3,232)=113.17, p< .01）。

表3: PC教室・普通教室群別（下位・上位）
リスニング空所補充前後テスト（FBT）の平均 (n=参加者数)

<table>
<thead>
<tr>
<th>グループ</th>
<th>n</th>
<th>テスト平均値</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
</tr>
<tr>
<td>PC-L</td>
<td>51</td>
<td>27.29</td>
</tr>
<tr>
<td></td>
<td>(7.80)</td>
<td>(6.36)</td>
</tr>
<tr>
<td>PC-H</td>
<td>65</td>
<td>31.60</td>
</tr>
<tr>
<td></td>
<td>(7.06)</td>
<td>(6.07)</td>
</tr>
<tr>
<td>RG-L</td>
<td>56</td>
<td>26.52</td>
</tr>
<tr>
<td></td>
<td>(11.89)</td>
<td>(10.60)</td>
</tr>
<tr>
<td>RG-H</td>
<td>61</td>
<td>31.53</td>
</tr>
<tr>
<td></td>
<td>(8.22)</td>
<td>(7.56)</td>
</tr>
</tbody>
</table>

括弧内は標準偏差
PC-L: PC教室下位群、PC-H: PC教室上位群
RG-L: 普通教室下位群、RG-H: 普通教室上位群

表4: PC教室・普通教室群別（下位・上位）
TOEICリスニング前後テストの平均（TLT）(n=参加者数)

<table>
<thead>
<tr>
<th>グループ</th>
<th>n</th>
<th>テスト平均値</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
</tr>
<tr>
<td>PC-L</td>
<td>51</td>
<td>9.53</td>
</tr>
<tr>
<td></td>
<td>(1.76)</td>
<td>(3.07)</td>
</tr>
<tr>
<td>PC-H</td>
<td>65</td>
<td>13.88</td>
</tr>
<tr>
<td></td>
<td>(1.58)</td>
<td>(2.39)</td>
</tr>
<tr>
<td>RG-L</td>
<td>56</td>
<td>9.04</td>
</tr>
<tr>
<td></td>
<td>(1.99)</td>
<td>(2.75)</td>
</tr>
<tr>
<td>RG-H</td>
<td>61</td>
<td>14.13</td>
</tr>
<tr>
<td></td>
<td>(2.41)</td>
<td>(3.14)</td>
</tr>
</tbody>
</table>

括弧内は標準偏差
PC-L: PC教室下位群、PC-H: PC教室上位群
RG-L: 普通教室下位群、RG-H: 普通教室上位群

4群（PC教室群（下位群・上位群）と普通教室群（下位群・上位群））の反復
測定による分散分析の結果、各事前・事後テストの得点はFBT、TLT共に有意差（F(1,229)=183.34, p<.01）、（F(1,229)=74.68, p<.01）が認められた。4群間における教室条件と各事前・事後テストの得点変化は、FBT、TLT共に有意差（F(3,229)=15.23, p<.01）（図2）、（F(3,229)=18.10, p<.01）（図3）が検出され、4群間の教室条件と得点変化との間に交互作用が認められた。多重比較シェフェの検定より、群間に有意差が検出され、交互作用が認められたものは、FBTでは、普通教室下位群と両上位群（表5）、TLTでは、教室環境に関係なく、両下位群と両上位群（表6）であった。

図2. PC教室・普通教室群別リスニング空所補充前後テスト（FBT）の推移
PC-L：PC教室下位群、PC-H：PC教室上位群
RG-L：普通教室下位群、RG-H：普通教室上位群

図3. PC教室・普通教室群別TOEICリスニング前後テスト（TLT）の推移
PC-L：PC教室下位群、PC-H：PC教室上位群
RG-L：普通教室下位群、RG-H：普通教室上位群 28点満点
表5：FBTのシェフェ検定における群間の差

<table>
<thead>
<tr>
<th>(I) 群別</th>
<th>(J)群別</th>
<th>平均値の差 (I-J)</th>
<th>標準誤差</th>
<th>有意確率</th>
<th>95%信頼区間</th>
</tr>
</thead>
<tbody>
<tr>
<td>RG-L</td>
<td>PC-L</td>
<td>-3.54</td>
<td>1.53</td>
<td>0.15</td>
<td>-7.85, 0.78</td>
</tr>
<tr>
<td>PC-H</td>
<td>-7.51*</td>
<td>1.44</td>
<td>0.00</td>
<td>-11.57, -3.45</td>
<td></td>
</tr>
<tr>
<td>RG-H</td>
<td>-5.36*</td>
<td>1.46</td>
<td>0.00</td>
<td>-9.48, -1.23</td>
<td></td>
</tr>
</tbody>
</table>

PC-L：PC教室下位群、PC-H：PC教室上位群
RG-L：普通教室下位群、RG-H：普通教室上位群
*平均値の差は.05水準で有意。

表6：TLTのシェフェ検定における群間の差

<table>
<thead>
<tr>
<th>(I) 群別</th>
<th>(J)群別</th>
<th>平均値の差 (I-J)</th>
<th>標準誤差</th>
<th>有意確率</th>
<th>95%信頼区間</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC-L</td>
<td>PC-H</td>
<td>-2.75*</td>
<td>0.36</td>
<td>0.00</td>
<td>-3.77, -1.74</td>
</tr>
<tr>
<td>RG-L</td>
<td>0.83</td>
<td>0.37</td>
<td>0.18</td>
<td>0.00</td>
<td>-0.22, 1.88</td>
</tr>
<tr>
<td>RG-H</td>
<td>-2.99*</td>
<td>0.37</td>
<td>0.00</td>
<td>-4.02, -1.96</td>
<td></td>
</tr>
<tr>
<td>RG-L</td>
<td>PC-L</td>
<td>-0.83</td>
<td>0.37</td>
<td>0.18</td>
<td>-1.88, 0.22</td>
</tr>
<tr>
<td>PC-H</td>
<td>-3.58*</td>
<td>0.35</td>
<td>0.00</td>
<td>-4.57, -2.59</td>
<td></td>
</tr>
<tr>
<td>RG-H</td>
<td>-3.82*</td>
<td>0.36</td>
<td>0.00</td>
<td>-4.82, -2.81</td>
<td></td>
</tr>
</tbody>
</table>

PC-L：PC教室下位群、PC-H：PC教室上位群
RG-L：普通教室下位群、RG-H：普通教室上位群
*平均値の差は.05水準で有意。

2群における二種類のテスト得点を比較した結果、両テストとも後テストでの得点が前テストに比べて有意に増加したが、FBTの得点の伸びにのみPC教室群と普通教室群で有意な差が認められた。二種類のテスト結果から、同じテキストを利用し、異なる学習環境条件で指導した場合、その場での深い意味理解を要求しない英語音声の知覚作業を主とするリスニング力の伸長に環境(人工物：PC)の影響が有意に認められた。

リスニング課題全体の制限時間および解答に関する質疑応答を含む解説時間は同じ時間配分であったにも関わらず、FBTで群間の差が有意であったということは、学習環境の差が両群での学習行為に差をもたらし、結果として学習者のリスニングスキルや学習方略に対する意識に影響を与えたことが原因ではないかと思われた。普通教室では、リスニング課題取り組み中にどの部分を繰り返すのかは学習者自身の意思のみで変更できない環境であった。一方、PC利で群の学習環境では、個人の意思により自己のベースでどの部分を繰り返すのかをPCという人工物を活用して変更するという学習行為が加わり、学習行為に自ずと起こる差異が学校プロセスおよび学習効果に意味ある影響を与え、PC教室群のリスニング知覚レベルが有意に伸びたと考えられた。半期間という短い期間の、音声の知覚レベルというその場での深い意味理解を要求されない、比較的認知作業量の軽い課題に早晩成果が現れたのではないかと推定された。なお、PCから流れ出す音声はデジタル化されたものであり、ヘッドセット経由で雑音が入らず、音量調整も各個人で可能だったため、教室全体で聞くCD再生音声に比べて、外部音や雑音の影響は無いものと思われた。
一方、TLTは伝える意味内容に関する理解を問うものが多く、理解力を測定する性格が強い課題である。上述したように、事前事後のテスト間の得点の伸びが有意ではあったものの、2群間に統計的に有意な差は認められなかった。つまり、PCの利用は理解力の向上にも役立つものと考えていたが、本研究での半期という短い研究授業期間では、理解力に差異をもたらすほど十分な時間がなかったのではないかと考えられた。この点については、吉田、吉田、倉本（2004）において、本研究と同様の条件での指導、事前・事後テストを通年で実施し調査分析した結果、半期では表れなかった統計的な有意差が認められたとの報告がある。この結果は、英文音声を繰り返し聞くことで「繰り返し学習：リハーサル」の効果を得ることができ、結果的に「長期記憶に結びつく」ものであると予測（御領、他：1995, pp.105-108）されるという指摘に符合するものと思われる。

2群間の事前・事後テストの得点変化を比較し、統計的に有意な結果についての考察を述べてきたが、さらに2群を成績上位群と下位群に分割し、4群間での事前・事後テストの得点差への影響を測る多重比較シェフェの検定を行った。上述の通り、FBTでは、普通教室下位群と両上位群に有意差があり、PC教室上位群とPC教室下位群では有意な差がなかった。一方、TLT（TOEIC listening test）では、両教室の下位群同士、上位群同士に有意な差はなく、PC教室下位群と両教室の上位群および普通教室下位群と両教室の上位群で、共有有意な差が認められた。これは2群間での検定結果「PC教室環境下では普通教室に比べて、FBT結果に有意な伸びがある」を追認するもので、知覚作業を主とするリスニング練習には、下位群上位群共に、PC環境が相応しいことを示唆している。しかし、TOEICリスニングのように、伝える意味内容に関する理解をともなう作業においては、研究授業遂行が群間に有意差を検出したのは下位群のみで、教室環境の差の影響を受けないのではないかと思われた。

上述の吉田、吉田、倉本（2004）の中期的な研究結果とは異なり、短期間ににおいての同じ内容による異なる学習環境下の授業では、PC環境は知覚レベルのリスニング練習に影響があり、意味理解を伴うリスニング練習では、教室環境差の影響は明確ではなかった。また、意味理解を伴うリスニング練習においては、成績下位群は、どちらの環境下でも、上位群に比べて有意なレベルで成績が伸ばしたことから、研究授業遂行は、共有下位群により相応しい可能性があるといえる。そこで、これらの原因を質問紙調査から考察した。

質問紙調査の結果と考察

授業開始期に行った質問項目で、学習者が利用していると感じるリスニングスキルや学習方略に対する意識傾向を調査し、終了期との比較を行なった。質問紙による前後結果に関して、1項目ごとにウィルコクスンの符号付順位和検定を行ったところ、PC教室と普通教室で有意な変化のあった項目は、下記の通りである（表7、表8）。
表7: PC教室前後質問紙の変化のウィルコクスン符号付順位和検定

<table>
<thead>
<tr>
<th>質問番号</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q10</th>
<th>Q16</th>
</tr>
</thead>
<tbody>
<tr>
<td>同順位補正Z値</td>
<td>-2.05</td>
<td>-2.29</td>
<td>-2.95</td>
<td>-4.27</td>
<td>-4.22</td>
<td>-2.18</td>
<td>-2.51</td>
</tr>
<tr>
<td>同順位補正P値（両側検定）</td>
<td>0.04</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.01</td>
</tr>
</tbody>
</table>

表8: 普通教室前後質問紙の変化のウィルコクスン符号付順位和検定

<table>
<thead>
<tr>
<th>質問番号</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q7</th>
<th>Q18</th>
<th>Q19</th>
<th>Q21</th>
<th>Q24</th>
</tr>
</thead>
<tbody>
<tr>
<td>同順位補正Z値</td>
<td>-4.65</td>
<td>-5.58</td>
<td>-7.04</td>
<td>-3.64</td>
<td>-2.60</td>
<td>-6.01</td>
<td>-3.84</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>同順位補正P値（両側検定）</td>
<td>-4.27</td>
<td>-4.81</td>
<td>-2.69</td>
<td>-3.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

上記検定の結果、PC教室群では、質問項目2, 3, 4, 5, 6, 10, 16の同順位補正z値は、帰無仮説の棄却域に入ると、またp値についても同順位補正p値が0.05以下となり、帰無仮説は棄却された。更に、24個の質問項目に関して同検定にボンフェローニの修正（両側検定での危険率0.2パーセントの上側境界値z以上）を加えた多重比較の結果、上記7項目の内下記の2項目のみが事前質問紙と事後質問紙の結果に有意な差が認められた。質問項目5：英語を母語とする人の話す音は全体がつながっていて聞き取りにくい（z=-4.27, 3.08<|-4.27|, p<0.00）、質問項目6：英語独特の音声のつながり方が有ると知っている（z=-4.22, 3.08<|-4.22|, p<0.00）。

普通教室では、質問項目1, 2, 3, 4, 5, 6, 7, 18, 19, 21, 24の同順位補正z値は、帰無仮説の棄却域に入ると、またp値についても同順位補正p値が0.05以下となり、帰無仮説は棄却された。更に、ボンフェローニの修正を加えた多重比較の結果、上記11項目の内下記の8項目のみが事前質問紙と事後質問紙の結果に有意な差が認められた。質問項目1：英語を母語とする人が自己に対してゆっくり話せば聞き取れる、質問項目2：教材のテープは速度がゆっくりのものであれば聞き取れる、質問項目3：教材のテープは何度も聞き返せば聞き取れる、質問項目4：英語は話すリズムが日本語とは異なるので聞き取りにくい、質問項目6：英語独特の音声のつながり方が有ると知っている、質問項目7：言葉と言葉のつながりに消える音があることを知っている、質問項目18：トピックになじむ音があれば聞き取りやすい、質問項目19：英語を聞くとき文法に注意する。それそれぞれ両側検定での危険率0.2パーセントの上側境界値z以上（質問項目1（z=-4.65, 3.08<|-4.65|, p<0.00）、質問項目2（z=-5.58, 3.08<|-5.58|, p<0.00）、質問項目3（z=-7.04, 3.08<|-7.04|, p<0.00）、質問項目4（z=-3.64, 3.08<|-3.64|, p<0.00）、質問項目6（z=-6.01, 3.08<|-6.01|, p<0.00）、質問項目7（z=-3.84, 3.08<|-3.84|, p<0.00）、質問項目18（z=-4.27, 3.08<|-4.27|, p<0.00）、質問項目19（z=-4.81, 3.08<|-4.81|, p<0.00））。次に、PC教室群、普通教室群を上位群と下位群に分けた4群間での比較を行ったが、2群間での比較との特徴ある差を認めなかった。

上記のとおり、PC教室群ではリスニングの知覚技術に関わる質問項目の5や6が有意に変化しているが、その他の項目に有意な変化は見られなかった。一方、普通教室では、リスニングスキルや学習方略のさまざまな項目に有意な変化が認められた。
つまり、PC教室群では、限られた時間であるが、自身の意思でPCを操作し聞くという自由に行える学習環境にあるため、学習者は多岐に渡る学習方略を用いず、意識して音声を聞き取る目的に絞ってリスニングスキルの学習を遂行したのではないかと考えられる。一方、普通教室では、教師主導による一斎聞き取り作業が主であるため、教員の助言に沿った意識変化が起こり、PC教室群には見られないさまざまな学習方略を用いてリスニングスキルの学習を遂行したものと考えられる。普通教室では、CDからの英語情報（教材）を聞き取る場合に、他者（教員）からの情報や他の学習者による発表など、周辺情報にさらされる機会がPC教室に比べて多い。普通教室では、より頻繁な周囲との相互作用から、より多くの学習方略やリスニングスキルに目を向けるようになったものと思われる。しかし、テスト結果との比較において、より多くの学習方略やリスニングスキルに関する意識の変化が、直接テスト得点の伸びに結びついていないことがわかった。また、PC教室群、普通教室群とも上位群と下位群に分けた4群間での比較において、学習方略やリスニングスキルに関する意識の変化に特に有意な変化がみられなかったことからも、本研究において上位群と下位群のテスト結果の差と意識変化の差を結びつけることはできなかった。

以上の比較分析から、PC教室群では、普通教室群より少ない意識の変化だったが、自己ベースで繰り返し聞き取ること作業を通じて、リスニング知覚の力を有意に伸ばした結果に影響があったことが示唆された。用いるリスニングスキルや学習方略の種類の多さが、直接知覚レベルの得点結果に現れるのではなく、自己ベースで変更のできる学習環境が、知覚レベルの伸張に影響があったものと考えられる。一方、理解力の得点については、両群間に有意な差がなく、PC教室でも普通教室と同じ学習効果を得ることができるわけだった。4ヶ月間という短期間ではあったが、学習環境に差のある指導を行った場合、PC教室環境が、深い意味理解を要求しない知覚レベルでのリスニングテストの得点増加に有意な差をもたらす一方、リスニングスキルや学習方略に関する意識の変化は普通教室群に比べて、より少ないことがわかった。ただし、両群の教室室外でのリスニング学習に関してはコントロールできないため、データ分析においては、教室外学習の影響を無視している点が本研究デザインの限界であることを付記する。

最後に、冒頭「研究の背景」で述べた「教師－学習者間の遂行配分の四つの過程」（1）教育目標の設定、（2）具体的な学習課題の設定、（3）課題遂行方法の選択と実施、（4）結果の評価、（巻川、1995）に本研究を照らし合わせてみた。本研究では、学習開始期の事前テストでリスニング力に差がない二つの学習者群に、同じ内容、同じ速度のリスニング素材を用いた。つまり、上記（1）と（2）の設定は、両群で同じであった。しかしながら、（3）課題遂行方法の選択と実施においては、学習者中心型（PC教室群）の指導法をとるか、教師主導の指導法（普通教室群）をとるかという差異があり、（4）結果の評価については、教師が同じ基準を用意するとしても、その表示法に大きな差異が認められた。すなわち、PC教室群ではその場で即時に個人ペースで解答が得られ、グループ内での順位や個人の誤答パターンなどもフィードバックされるが、普通教室群では、解答はグループ全体での同時の答え合わせであり、グループ内順位や個人の解答パターンなどはフィードバックされなかった。
提供した学習環境の違いによる（3）と（4）の条件の差異が、学習者の選択取るリスニングスキルや学習方略の選択に影響し、自己ベースでの学習成就の差となった。つまり、PC教室群では個人のPC画面にフィードバックされる正誤表示を基に個人が意識的に繰り返し作業することで、普通教室群の学習者よりもより自主的に特定のリスニングスキルに着目して、知覚レベルでの音の確認を行う習慣が養われたものと考えられた。しかし、「音の確認」作業では、PCに入力する自分の単語のつく里が正解か不正解かに集中し、ニュース内容のより深い意味理解を意識しなかった可能性があり、そのことがTOEIC得点で普通教室群との有意な差が現れなかった原因ではないかと考えられる。一方、普通教室群では、教師が正解をまとめて提示するため、学習者は「音の確認」作業においても、内容の意味理解に関しても受動的な範囲を超えず、より多くの学習方略に注意を向けたため、自主的なスキル選択が不十分となり、テストの得点が伸びなかったのではないかと考えられる。この点は、冒頭の「研究の背景」で引用したVogely(1995)やNyikos & Oxford (1993)が指摘しているように、学習方略の多さではなく、学習者自身が選んだ学習方略をどのように意識し、適切に利用したかの差異が学習内容の改善の差に影響したものと考えられた。

意味理解を含むTOEICリスニングテストで群間の差が有意ではなかった点は、半期間の授業においてどちらのグループの指導方針でも、音声の知覚段階を超えて、英文が伝える内容についてのより正確なより深い意味の理解まで導けなかったことが原因であると考えられる。利用するリスニングスキルや学習方略の種類や数の影響については、4か月の本研究では調査項目が不十分であったと思われる。また同時に、PC教室と普通教室という学習環境の差異がTLT（TOEIC listening test）に必要な英語力を高める要因として差を生じさせるには不十分である可能性もあるが、上述の長期的研究において有意な差が出た結果も視野に入れ、さらに調査研究をせねばならない。

「音の確認」の学習課題において、PC教室群では「自主的」、普通教室群では「受動的」という異差が生じた。この異差が作用して、学習者がどのようなリスニングスキルや学習方略をどのように意識するかに影響が現れ、個人ペースの学習が成り立ったPC教室群では自主的な作業を通じて自律学習傾向が高まり、結果としてリスニング知覚レベルの伸張に影響を与えたと考えられる。リスニング知覚レベルの伸張には、「自己ベースによる集中」という学習環境提供が重要な影響力を持ち、この学習環境下ではより少ない学習方略で学習を遂行することが学習効果を高めると示唆された。

おわりに

普通教室群では、ユニットに関する言語や内容に関わる解説などが学習者の反応に応じて付加され、PC教室での授業展開に比べると、教師主導型の授業が展開された。グループでの相互作用としての学習が進む環境では、周囲の影響も受け、より多くのリスニングスキルや学習方略を選んだものと考えられる。他方PC教室群では教師の解説や指導があるものの、個人ペースによる単独での作業を通じて、学習者中心型の授業が展開された。即ち、PC操作とPC情報の助けを得て、学習者は自身の学習に対する意識をより強め、特定のリスニングスキルの向上を選んだのではないか。学習の促進は、学習者と教授者の関係
や教室条件が異なる学習環境においても図れるものであるが、課題遂行方法の選択と実施や結果の評価の提示方法を変更すると、学習者の意識や学習行為に異なる影響を及ぼすことが示唆された。

しかし、本研究では、量的データによる分析でのPC教室と普通教室という条件の学習環境で、リスニング力の伸張と学習意識に差異が起こるか否かについてを調査の目的としたため、学習者がリスニングスキルや学習方略を具体的にどのように利用したのかは調査しなかった。このため、リスニングスキルや学習方略の種類や数が、テスト結果にどのように影響したのかを明らかにすることができなかった。今後は、竹内（2003）が指摘しているように「方略の使用は内的な活動」であるから、インタビューなどの質的アプローチでのデータ収集による精緻な分析とあわせて、自律学習の進展度やテスト結果との関係を明らかにしていきたいと考えている。今回調査で問うたリスニングスキルなどについて、参加者から具体的にどのような場面、どのような機会にスキルの選択や変更を行ったか、その理由、また、学習環境に対する学習者の評価などを調査していきたい。


参考文献


国際ビジネスコミュニケーション協会TOEIC運営委員会編 （2000）. 「TOEIC公式ガイド＆問題集：日本語版」 東京: 国際ビジネスコミュニケーション協会

TOEIC運営委員会

竹内理 (2003). 「より良い外国語学習法を求めて」 東京: 松柏社

吉田信介、吉田晴世、池田真生子 （2002）. 「オンラインで学ぶVOAライフ系総合英語」 東京: 松柏社

付録A
空所補充テスト(サンプル8問目まで)

**TOEFL**

Our VOA listener question this week comes from many places around the world. 1) want to know about the TOEFL 2) required before a foreign 3) can attend an American college. That 4) the subject of our report today. 5) information can be found on the 6) English web page at www.voa.gov/special. If 7) do not have a computer, get 8) to write down some addresses.

解答：1) Students, 2) test, 3) student, 4) is, 5) This, 6) Special, 7) you, 8) ready
Differences in Reading Strategies Employed by Students Constructing Graphic Organizers and Students Producing Summaries in EFL Reading

Akio Suzuki
Josai University

The big difference between summaries and graphic organizers (GOs) as adjunct aids for EFL reading is that GOs can reduce the cognitive burden on readers because of their two-dimensional spatial displays while summaries cannot. In this study, five Japanese high school students were required to construct GOs and another five were required to produce summaries while reading a passage written in English. They were required to report what they were thinking while reading the passage and producing adjunct aids so that their use of reading strategies could be examined. The results gained from think-aloud protocol analysis indicated that the GO group reported more general comprehension strategies than the summary group. The findings are discussed from the perspective that the task of constructing the GOs using visual argument allowed the students to employ more general comprehension strategies.
Graphic organizers and summaries as adjunct aids in EFL Reading

With regard to instructional techniques to teach English as a foreign language for teaching reading, there are currently a variety of adjunct aids, such as outlines (see Silberstein, 1994), summaries (see Irwin, 1986), typographical cueing (see Hershberger & Terry, 1965), phrase reading (see Hatch, 1979), pictures (see Omaggio, 1979), and graphic organizers (see Grabe & Stoller, 2002). Adjunct aids are assumed to facilitate students’ comprehension of a text, either by attaching further reference material to a sentential text or by requiring students to independently produce summaries or graphic organizers.

Among adjunct aids, the writing of summaries has recently become relatively popular as a technique for facilitating students’ comprehension in EFL reading classes in Japan. A number of reading texts authorized by the Ministry of Education and Science, such as New Cosmos Reading (Ohura, Tennuma, Ando, Toda, & Yanagisawa, 2002) and Sunshine Reading (Hashimoto, Ibe, Furukawa, & Takanashi, 2002) among others include summary production tasks. Considering that a number of reading texts include summary tasks and that many Japanese teachers use this technique for the comprehension of texts, it is safe to say that the writing of summaries is likely to be one generally accepted adjunct display used for EFL reading classes in Japan. Several studies have shown that the writing of summaries enhances both comprehension and memory (Rinehart, Stahl, & Erickson, 1986; Winograd & Bridge, 1986).

Like summary production, the use of graphic organizers (GOs, hereafter) as visual adjunct displays is also becoming popular; in fact, foreign language educators have long taken it for granted that visual aids are generally useful for the learning of a foreign language (Omaggio, 1979). GOs are understood to be a type of adjunct display developed so as to facilitate learners’ understanding of important interconceptual relations through the use of spatially displayed information (Robinson, 1998). A number of empirical studies have investigated GOs’ facilitative effectiveness in text comprehension in a native language (e.g., Levie & Lentz, 1982; Mayer, 1997; Robinson & Kiewra, 1995; Robinson, Robinson, & Katayama, 1999; Sims & Hegarty, 1997; Waller & Whalley, 1987; Winn, 1987; Winn, Li, & Schill, 1991).

There has been a variety of research that has reported positive outcomes concerning GO’s facilitative effectiveness in text comprehension when students are asked to construct their own GOs (e.g., Alesandrini, 1981; Alvermann & Boothby, 1986; Barron, 1980; Dean & Kulhavy, 1981;
Suzuki

Geva, 1983; Hawk, 1986; Holley & Dansereau, 1984; 1981; Novak, 1991; Snowman & Cunningham, 1975). In one study on student-constructed GOs, Bean, Singer, Sorter, and Frazee (1986) investigated the facilitative effectiveness of three written-product format treatments: (a) GOs-plus-summary-training, (b) GOs only, and (c) outlining only, with a group of 10th-grade students using history passages. Although their results revealed no significant differences between the three groups on short-term probes, the performance of the GOs plus summarization training group was significantly better than both the graphic-organizer and outlining groups on a delayed measure. In addition, on a written transfer passage, participants in both the graphic-organizer groups constructed significantly better summary paragraphs.

The act of constructing GOs and that of producing a sentential summary have several features in common (see Bean, et al., 1986), and, in fact, Moore, Chan, and Au (1993) have referred to graphic organizers as diagrammatic summaries (p. 59). Both summaries and GOs are expected to be effective in helping students comprehend texts better, but the crucial difference between the two is that GOs take the form of two-dimensional spatial displays whereas summaries take the form of one-dimensional sentential displays.

With regard to this difference, Larkin and Simon (1987) argued that one-dimensional sentential displays are comprised of sequential propositions such as sentences found in a text, while two-dimensional spatial displays, indexed by location in a plane, explicitly represent information that is only implicit in sentential displays. To be more specific, when information is presented as text, a linear search is required. When the first relevant element concerning a particular concept is found and processed, it must be stored in the memory before the search can continue for the next relevant element. This routine of find and process must continue in this fashion in the working memory of the student until the last element is found and processed. This will inevitably place a great deal of strain on the student’s cognitive resources and, as a result, it is likely that a reading error may occur. In the case of spatial displays, however, once a relevant element is found, the next element appears next to or near it. Due to this advantage, the burden upon the working memory of a student is reduced since both elements concerning the concept can be viewed simultaneously and thus efficiently. Larkin and Simon (1987) specifically referred to the efficiency of two-dimensional spatial displays as computational efficacy.

Between reading in a native language (L1) and in a foreign or second language (L2), a large difference can be expected in the cognitive burden.
In fact, researchers have posited a number of linguistic and processing differences between reading texts in L1 and L2 from a variety of perspectives, including lexical access (e.g., Grabe, 1999), knowledge of grammar (e.g., Urquhart & Weir, 1998), orthographic depth (e.g., Koda, 1999), and language threshold (e.g., Alderson, 2000). The general consensus is that reading texts in a foreign language puts readers under considerable strain, which can prevent them from gaining a completely effective understanding of the texts. However, it is my impression from the literature that if students construct GOs while reading a text written in English, their cognitive burden might reasonably be lowered, not only in the process but also once they have the GO product to refer to and to display. This can accordingly affect the students’ use of reading strategies.

Barnett (1988) summarized previous research on reading strategies and categorized them broadly into two types: one for comprehension at a text level and the other for comprehension at a word or grammar level. Similarly, Block (1986) categorized reading strategies broadly into two types: general comprehension strategies and local linguistic strategies. A general comprehension strategy involves readers’ recognizing text structure or integrating information at the text level, while local linguistic strategy deals with the attempt to understand specific linguistic units.

Because the differences in the ways of displaying information—two-dimensional spatial displays (GOs) and one-dimensional sentential displays (text summaries)—can influence the cognitive burden placed on EFL readers, the expected proportion of strategy use should vary depending on whether students construct GOs or produce summaries while reading a text written in English, as the type of strategy chosen reveals readers’ cognitive resources in comprehension (Langer, 1982).

**Purpose and Hypothesis**

The purpose of this study is to examine to what extent, if any, the production of different adjunct display types—graphic organizer (spatial) vs. summary (prose sentence) displays will cause EFL students to rely on different proportions of strategy use (categories according to Block, 1986) when reading a text written in English. My hypothesis is that in the process of constructing GOs, students will inevitably use greater cognitive resources for general reading comprehension strategies, such as recognizing text structure and integrating information, than those who instead produce summaries. This is because GOs possess the advantage of computational efficacy (Larkin & Simon, 1987), which is a particular
feature of two-dimensional spatial-display representations, such as maps, charts, graphs, and GOs. Students using such displays can see explicit, graphically-presented relationships, making it easier to formalize connections among concepts while reading. Spatial arrangement allows for data connections to be made more easily in GOs than in prose summaries, in which students must (a) search through many sentences to find one fact, (b) keep that fact in the working memory, (c) search for other facts and, finally, (d) draw inferences between and among the facts. Displaying, rather than describing, the organization or structure of concepts reduces the effort required to comprehend the intended message (Winn, 1990). In this paper I shall try to demonstrate that GOs possess the dual virtue of being (a) user-friendly (as product) and (b) cognitively challenging and therefore rewarding to construct (as process). Students will inevitably use greater cognitive sources for general comprehension reading strategies, such as recognizing text structure or integrating information, than those in the summary group, who are asked to produce summaries, which have, however, no relationship to computational efficacy, and thus need not to be aware of text structure or integrate information.

Method

Participants and Design

The participants in this study were initially 15 female third-year senior high school students (ages 17 to 18; all native speakers of Japanese). However, five were ultimately excluded because they failed to complete the assigned tasks of constructing GOs or writing summaries. Thus, the data of 10 participants was used in this study. All were students in my class, a one-year compulsory English course, and all participated in this study voluntarily. As I had previously introduced and used GO techniques, all participants had had the experience of producing GOs.

The participants were randomly divided into two groups. One group was required to produce GOs during or after reading a text in order to enhance comprehension, while the other group was required to write summaries for the same purpose and under the same conditions. In order to ensure the homogeneity of English reading levels between the two groups, an English reading skills test was conducted. English reading levels were determined using the reading portion of a TOEFL sample (Educational Testing Service, 1998). The average score of the ten participants was 20.00 out of 50. According to the $t$ test, the average TOEFL reading skills of the two groups did not differ significantly, $t(8) = 0.18, p > .05$ indicating that
the GO group \((M = 20.60, SD = 6.92)\) had almost the same English reading skills as the summary group \((M = 21.40, SD = 5.64)\).

**Material**

*Text.* An English reading passage (245 words, Flesch Reading Ease = 72.5, Flesch-Kincaid Grade Level = 6.3) was used (see Appendix). This text was taken from a textbook designed to teach paragraph reading (Fukuzaki & Yoneyama, 1999). This particular passage concerned a psychological experiment in which researchers predicted the future character of children according to their behavior in early childhood. The rhetorical structure of this text appears in Table 1. This passage was printed on one B4 (250 x 353 mm) page.

*Adjunct display.* Students in the GO group were asked to construct GOs as an adjunct display to the text, while those in the summary group were required to produce summaries for the same purpose. The construction of both GOs and summaries is intended to assist with the individual comprehension of textual content. Students in both groups were instructed to produce their adjunct displays on one A4 (210 x 297 mm) page. It was emphasized that they were expected to make the most of these individually created adjunct aids for better understanding of the text.

**Procedure**

Several sessions were conducted during spring vacation using a typical senior high school classroom with one to three participants each time. Invited into the room individually, they were told that the purpose of the study was to establish the way in which they read a text and construct an accompanying adjunct display—not to test them personally. The data collection methods used in this study was think-aloud protocol analysis. Although this think-aloud protocol technique is still controversial in that (a) it is difficult to measure what participants do automatically and thus are not aware of (Okouchi, 2001), and (b) this technique itself can influence the students’ natural cognitive processes (Kaiho & Harada, 1993), I decided to employ it, since it has been used in a number of studies investigating reading strategies (e.g., Block, 1986; 1992; Hartman, 1995; Kletzien, 1991; Narvaez, van den Broek, & Ruiz, 1999; Olshavsky, 1977; Philips, 1988; Stromso, Helge, & Braten, 2002; Yoshida, 1997) as well as in some studies examining students’ use of GOs (e.g., Guthrie, Weber, & Kimmerly, 1993; Moore & Scevak, 1997; Schnottz, Picard, & Hron, 1993).
Table 1. The Rhetorical Structure of the English Passage Used

<table>
<thead>
<tr>
<th>[the 1st paragraph]</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>This paragraph explained an experiment in which small children were instructed not to eat a marshmallow, although they could get it.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>[the 2nd paragraph]</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>This paragraph described the different behaviors of those children; some could not resist the temptation of eating the marshmallow while others could endure the temptation.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>[the 3rd paragraph]</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>This paragraph reported the findings of the experiment. High school students, who had shown patience in their early childhood outperformed those who had not, with respect to both their personalities and their academic achievement.</td>
<td></td>
</tr>
</tbody>
</table>

Participants first underwent a practice session (of approximately 30 minutes’ duration) to familiarize themselves with the required think-aloud procedure and the audio recording process. The practice session was conducted according to the procedure proposed by Ericsson and Simon (1984). It consisted of participants reading four short sample English passages and solving mathematical problems while simultaneously reporting what they were thinking. They were cautioned not to explain or analyze their thoughts while thinking aloud, and if they were silent for a particularly long period of time, they were prompted to report what they were thinking. Participants were asked to report in the same language in which they were thinking—either Japanese or English; all participants used only Japanese. The practice session continued until the author determined that the participants were able to perform the think-aloud process with ease. In order to avoid bias or imitation, no particular model or example was given.

After the practice session participants were provided with the text, instructions, and a blank sheet of paper upon which to produce their GOs or summaries. They were allowed to read and study the material without time constraint (their average study time was about 44 minutes); they were also allowed to consult a dictionary at any time throughout the session. All verbal reports were audiotaped. Following this stage and a 10-minute break, the participants completed the English reading skills test
(50 minutes) and were finally dismissed after being asked to not discuss the content of the material with other participants. One day after the session, the participant(s) were reconvened in the classroom and were given a delayed-recall test where they were asked to write down everything they remembered about the text on one A4 sheet of paper. The entire session was conducted by the author alone and all instructions were issued in Japanese. Even when there were three participants (the maximum mentioned earlier) in the room where the session was conducted, the entire session—including the practice session—was conducted individually. The room was rather large, so even when one participant was being instructed or was doing the practice session, the others were unable to observe what was happening. Since this procedure was repeated several times with one to three students at a time, all were strictly instructed not to talk with other participants about the TOEFL English reading skills test, the English passage, or about what happened during the sessions.

**Think-Aloud Protocols And Scoring**

Think-aloud protocol analysis was used to explore the relative quantity of the types of reading strategies employed by participants. The verbal protocol data was categorized into either of the following two strategies: local linguistic strategies (such as questioning the meaning of a word), and general comprehension strategy (such as recognizing textual structure). In order to investigate which strategies were used more, the taxonomy of reading strategies developed by Block (1986) was employed. This taxonomy provides a useful classification system for general comprehension and local linguistic strategies and was developed using think-aloud protocols conducted with expository texts. Moreover, Block used beginner level ESL college students as participants in her study. Considering that the participants in this study were third-year senior high school students, this classification, developed using beginning level students, was deemed appropriate. Alternative reading strategy classifications, for example, those of Anderson (1991) or Sarig (1993), were not deemed appropriate to this study. Almost 40% (18 out of 47) of Anderson’s classifications were specific to test-taking strategies, and although Sarig did investigate the process of producing a summary, the classification system was developed by examining one proficient EFL student, and as such does not provide the beginner-level measure applicable to this study.

When participants reported anticipating the text, recognizing text structure, integrating information, questioning the text, interpreting the
text, associating information in the text with other information outside the
text, commenting on the text, monitoring their own reading, correcting
their reading, or reacting to the text, then these reports were placed in the
general comprehensive strategy category. If students reported paraphrasing
English words, rereading the English text, questioning the meaning of an
English clause or sentence, questioning the meaning of English words, or
solving vocabulary problems with English words, then the reports were
placed in the local linguistic strategy category. The frequency of each type
of strategy was counted and the results of the GO group and the sum-
mary group were compared.

All of the participants’ recorded verbalizations during the think-aloud
protocols were transcribed by the author. The transcripts of the verbal
protocols were then matched with the corresponding sentences in the
text. Each protocol was analyzed by the author and another coder to
identify which type of strategy the participants used more, local linguis-
tic or general comprehension. Intercoder reliability was 91%.

When scoring the delayed-recall test, each sentence of the source text
was considered to contain one specific and independent fact which could
be potentially referred to by participants in their tests. For instance, the
sentence "Some children grab for the treat as soon as they are out the
door," has two specific and independent facts ("grab for the treat" and
"out the door."). Thus, theoretically, if a participant in the recall test wrote
only "children took the treat," only one point would be given. However,
if they wrote "children took the treat after going out," two points would
be given. The source text contained 16 sentences in all, and a score of one
point was assigned to each. The author examined the completed tests and
divided the information in each into separate passages which referred to
the relevant sentence of the source text. Then, the author and another re-
searcher analyzed independently whether there was agreement between
the students’ passage and the corresponding source text sentence with
one point being awarded for each correct reference (16 points being the
maximum possible score). Intercoder reliability turned out to be 96.88%.
Any discrepancies remaining in coding were resolved through discus-
sion.

Results

All statistical tests were conducted at the level of $\alpha = .05$. 
Adjunct Display Tasks

As mentioned earlier, five students failed to successfully complete their GO or summary tasks. Three students failed to complete their GOs; they tried to construct GOs, but instead wrote sentential memorandums rather than organized spatial displays. Similarly, 2 students failed to successfully complete their summaries; they simply translated the entire text into Japanese. The data of those who failed to complete the task was excluded and not analyzed. Figure 1 shows examples of successful and unsuccessful GOs. Figure 2 shows an example of a successful summary. Interestingly, all successful GOs appear to be very similar in format; children who ate the marshmallow were represented on the top right of the page, children who didn’t were represented on the top left. Accordingly, the future characters of the children were recorded at the bottom of the page in corresponding positions and linked by arrows. All the successfully completed GOs reveal an advantage of two-dimensional spatial display: the GO allows immediate comparison between the behavior of the children in childhood and their future characteristics (discovered by vertical flow), as well as immediate comparison between the two different characteristics (discovered by using horizontal comparison).

Figure 1. Examples of successful and unsuccessful GOs. (Translated here from the original Japanese by the author)
Scientists can predict the future by watching four-year-olds interact with a marshmallow. Those who can resist the temptation to eat the marshmallow grow up to be well-adjusted, popular, adventurous, and confident, while those who cannot grow up to be lonely and stubborn.

Table 2. Means as a Percentage and Standard Deviations of Reading Strategies for Each Group

<table>
<thead>
<tr>
<th>Display type</th>
<th>GO</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>General comprehension</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$ (%)</td>
<td>51.64</td>
<td>23.24</td>
</tr>
<tr>
<td>$SD$</td>
<td>18.28</td>
<td>11.59</td>
</tr>
<tr>
<td>Local linguistic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$ (%)</td>
<td>48.48</td>
<td>77.56</td>
</tr>
<tr>
<td>$SD$</td>
<td>18.44</td>
<td>11.20</td>
</tr>
</tbody>
</table>
Table 3. Examples of General Comprehension and Local Linguistic Strategies According to Block’s (1986) Taxonomy (Sentences in brackets were translated by the author)

<table>
<thead>
<tr>
<th>General comprehension strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>The text: A survey of the children’s parents and teachers found that those who as four-year-olds had the patience to wait for the second marshmallow generally grew up to be better adjusted, more popular, adventurous, confident and dependable teenagers.</td>
</tr>
<tr>
<td>Protocol: “へぇ・・・冒険好きなの？でも待てるんでしょう？・・・あたし・待てないほうが冒険好きな気がするよ・・・”</td>
</tr>
<tr>
<td>[ Hmm… He likes adventure? But he can wait, right? I think the ones who can’t wait are more adventurous.]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Local linguistic strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>The text: the same sentence as above</td>
</tr>
<tr>
<td>Protocol: “adjust・適応？・・・adjustって言う言葉・・・んーかっこいい・・・ ・・・adjust・これ遅いなぁ・・・適用・・・順応・洗脳する・・・順応・適用・順応しやすい・・・順応性か”</td>
</tr>
<tr>
<td>[ Adjust, adapt? What’s this word, “adjust”? Let’s see, it sounds cool; adjust; I’m being slow; apply? Accustom, brainwashing? Accustom, adapt, accustom–this means accustom.]</td>
</tr>
</tbody>
</table>

A 2 x 2 (Adjunct Displays x Reading Strategies) factorial analysis of variance (ANOVA) was conducted in order to determine the statistical significance of the difference between the general comprehension and local linguistic strategies used by the two groups. Statistical analysis revealed that the effect of strategy (general comprehension strategy vs. local linguistic strategy) was not significant \(F(1, 8) = 1.68, p>.05\). However, there was a significant main effect of group (GO vs. Summary), \(F(1, 8) = 5.61, p>.05\] with the GO group employing more general comprehension strategies and fewer local linguistic strategies than the summary group (see Table 2). In addition, the Strategy x Group interaction was significant \(F(1, 8) = 7.08, p>.05\]. A post hoc Ryan’s procedure indicated that display type had a significant effect on the strategy. Using a sentential adjunct display (summary) caused students to employ more local linguistic strategies than general comprehension strategies, whereas using a spatial adjunct display (GO) caused equal strategy use.
**Task Completion Time**

Table 4 shows the means and standard deviations of participants’ task completion time. The average time was 44.40 minutes. According to the *t* test, the average task completion times of the two groups did not differ significantly, *t* (8) = 0.36, *p* > .05.

**One-Day Delayed Free Recall Test**

Table 4 also shows the means and standard deviations of the delayed free recall. The GO group’s average score was 6.73 out of 16, and the summary group’s was 5.97 out of 16. According to the *t*-test, the average scores of the two groups did not differ significantly, *t* (8) = 0.44, *p* > .05.

<table>
<thead>
<tr>
<th>Table 4. Means and Standard Deviations for Task Completion Time and Delayed Free-recall Tests for Each Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test</strong></td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>Task Completion Time</td>
</tr>
<tr>
<td><em>M</em> (minutes)</td>
</tr>
<tr>
<td>Delayed free-recall tests</td>
</tr>
<tr>
<td><em>M</em> (out of 16)</td>
</tr>
<tr>
<td><em>SD</em></td>
</tr>
</tbody>
</table>

**Discussion**

The research question addressed in this study is: do different adjunct display types (GO vs. summary) cause students to use different proportions of strategy use (general comprehension strategy vs. local linguistic strategy)?

Regarding strategy variation between the two groups, the results of the ANOVA indicated that students in the GO group used significantly fewer local linguistic strategies than those in the summary group, and that the GO group employed significantly more general comprehension strategies than the summary group.

It was confirmed that English reading skills did not differ between the groups and that students in both groups demonstrated an equally
successful comprehension of the text material and spent an almost equal amount of time on the task. The only difference in the treatment of the two groups was the type of adjunct display used as an aid to study. It can be concluded that the different qualities of the display types (spatial vs. sentential) are responsible for the variations evident in the proportion of strategy use (general vs. local linguistic comprehension).

The summaries and GOs produced bear out the difference in computational efficacy (Larkin & Simon, 1987). The following example summary was produced by a student:

*Scientists can predict the future by watching 4-year-olds interact with a marshmallow. Those who can resist catching it grow up to be well-adjusted, popular, adventurous, and confident, while those who can’t grow up to be lonely and stubborn.*

In order for the student creating this particular prose summary to contrast the contrasting characteristics of children, she had to a) search through this summary to discover that children who didn’t eat the marshmallow grew up to become good, b) retain this information in her working memory, c) search for the relevant information, and d) comprehend and compare the significance of the contrasting behavior and its subsequent outcomes. On the other hand, for a student using a GO to obtain the same contrasting information, she just has to a) scan vertically the data on children who didn’t eat the marshmallow, b) do likewise with the data on children who did eat it, and c) compare the data horizontally (see Figure 3).

The completed GOs produced by the GO group all contain this advantage of computational efficacy. It can be argued that in order to produce GOs equipped with this computational efficacy, they had to expend greater cognitive resources on general comprehension reading strategy, such as recognizing text structure or integrating information, than on local linguistic reading strategy such as paraphrasing English words or solving grammatical problems. As a result, the proportion of general comprehension reading strategy use of students in the GO group was significantly higher than that of those in the summary group.

With respect to the depth of text understanding, Kintsch (1994) distinguished learning text from learning from text. Learning text indicates being able to reproduce it in some form. For instance, when students are required to recite a poem or a passage of literature, this activity can be seen as learning text. Learning from text means using the information gained
from the text in other ways, not just for reproduction. For example, when reading a scientific explanatory text, students can connect text information with their prior knowledge and thus infer new ideas. For this reason, *learning from text* is viewed as a deeper understanding of text (Kintsch, 1988; van Dijk & Kintsch, 1983). Cognitive activities such as integration or inference are particular features of *learning from text* (Kintsch, 1994).

The findings presented in this study are of significance in that constructing GOs has the potential to urge EFL readers to use greater resources for general comprehension reading strategy, which can lead to students’ *learning from text* rather than *learning text* (Kintsch, 1994). Students in the GO group used more general comprehension strategies such as inference or information integration, which are both important cognitive activities in order for EFL readers to implement *learning from text*.

In this study, the number of participants (10) was rather small, so it is not possible to generalize that the different proportion of strategies employed by the two groups (GOs vs. summaries) can lead to either *learning text* or *learning from text* (Kintsch, 1994) solely from the findings presented. In addition, the task to determine how well students comprehended the text was a delayed free recall test, and thus only adequate to examine

![Diagram](image_url)
the students’ comprehension at the level of a propositional textbase for learning text, and not of a situation model for learning from text (Kintsch, 1988; van Dijk & Kintsch, 1983), so it is reasonable that there were not any significant differences regarding the delayed free-recall tests.

As this study is an exploratory study, further research is necessary with a larger number of participants, as well as a variety of text types, to consider and examine the application of students’ constructing GOs as aids to EFL reading. In order to promote learning from text (Kintsch, 1994), future research is needed that examines how students comprehend a text, that requires readers to connect what is explicitly stated in the text with their background knowledge, and that requires them to derive inferences from this linkage.

Akio Suzuki is a full-time lecturer at Josai University, Saitama. His research interests include TEFL, cognitive psychology, and educational psychology.

References


Appendix

(English passage used as material)

It turns out that a scientist can see the future by watching four-year-olds interact with a marshmallow. The researcher invites children, one by one, into a plain room and begins the gentle torment. You can have this marshmallow right now, he says. But if you wait while I run an errand, you can have two marshmallows when I get back. And then he leaves.

Some children grab for the treat, as soon as he’s out the door. Some last a few minutes before they give in. But others are determined to wait. They cover their eyes; they put their heads down; they sing to themselves; they try to play games. When the researcher returns, he gives these children their two marshmallows. And then science waits for them to grow up.

By the time the children reach high school, something remarkable has happened. A survey of the children’s parents and teachers found that those who as four-year-olds had the patience to wait for the second marshmallow generally grew up to be better adjusted, more popular, adventurous, confident and dependable teenagers. The children who gave in to temptation early on were more likely to be lonely, easily frustrated and stubborn. They couldn’t put up with stress and shied away from challenges. And when some of the students in the two groups took the Scholastic Aptitude Test, which U.S. students take when applying to university, the kids who had waited longer scored an average of 210 points higher.
The Effect of EFL Students’ Self-Monitoring on Class Achievement Test Scores

Terry Vanderveen
Kanagawa University

The effect of self-monitoring on an achievement test in an EFL classroom setting was investigated. The results of an experimental study that measured the difference in class achievement test scores between a self-monitoring and an external-monitoring group are reported. Participants were 114 first-year students in a Japanese university and college. Target language performance and behaviors were monitored and responses recorded on a protocol by the students in the self-monitoring group and monitored by the instructor in the external-monitoring group. The time required for implementing the protocol or treatment was also measured and was considered to be an important pedagogical factor. At an alpha level of .05, the difference in achievement scores between the groups was not statistically significant.

Self-regulated learning and learning strategy theory emphasize the role of self-observation and self-evaluation in learning. Arguments for the effectiveness of self-regulation are based on the assumption that conscious reflection on one’s performance increases the frequency...
and accuracy of target behaviors or performance. Self-regulation can be broken down into many subprocesses such as self-monitoring, self-instruction, self-evaluation, self-correction, self-reinforcement (Mace, Belfiore, & Hutchison, 2001), self-observation, self-judgment, and self-reaction (Zimmerman, 1989). The line separating these processes is often unclear (Benson, 2001) because as Schunk (2001) observes they are interrelated and dependent. This paper will use the term self-monitoring (SM) because it is the most widely used and, in my opinion, best describes the process discussed in this study.

O’Malley and Chamot (1990) define SM as “checking one’s comprehension during listening or reading, or checking the accuracy and/or appropriateness of one’s oral or written production while it is taking place” and contrast this with self-evaluation, which is “checking the outcomes of one’s own language learning against a standard after the learning has been completed” (p. 232). Wenden (1991) offers similar definitions that differentiate SM and self-evaluation based on the time lapsed between production and assessment. This time period, however, does not change the nature of the evaluation process. Evaluation occurs both during and after learning. Therefore, for the purposes of the current study, SM and self-evaluation are considered a unitary construct referred to as SM and defined as conscious observation and evaluation, which is usually recorded, of one’s own performance or behavior on a learning task.

The proposition that SM enhances learning has been widely discussed in general education as well as language education. SM has been investigated in relation to many aspects of learning, with behavior modification receiving the most attention. Researchers into behavior disorders in children have extensively reported improved classroom behavior or performance due to SM (Lam, Cole, Shapiro, & Bambara, 1994; Reid & Harris, 1993). However, not all studies of SM and behavior modification have yielded positive results. SM was ineffective in increasing productivity in a study of adults with mental retardation (Shapiro & Ackerman, 1983).

Mixed results are also reported from studies examining the impact of SM on academic outcomes. Mace and Kratochwill (1985) showed that SM significantly reduced L1 verbal nonfluencies in college students’ speech, and Lan’s (1993) experiment on the effects of SM on college students’ statistics course grades resulted in the SM group outperforming the instructor-monitoring and control groups. However, in a similar experiment involving children studying math (Schunk, 1983), the posttreatment achievement scores of the SM group of children were comparable with those of the externally monitored group. SM also failed to improve per-

Whereas SM studies in education have been both behavioral and cognitive in nature, SLA researchers have focused solely on cognitive aspects. A cross-sectional speech monitoring study by van Hest (2000) shows that the ability to monitor one’s L2 speech errors is a valid predictor of acquisition. In a review of monitoring and self-repair in L2 speech, Kormos (1999) concludes that the SM of L2 speech aids acquisition and is a positive correlate of proficiency. In two descriptive studies, Charles (1990) and Cresswell (2000) emphasize the importance of SM through the technique of writing notes or annotations. Checking one’s reading comprehension is also considered useful and is recommended for developing reading skills (Block, 1992; Casanave, 1988). It has been extensively argued that SM and similar strategies are characteristics of good learners and enhance learning (Blanche & Merino, 1989; O’Malley & Chamot, 1990; Wenden, 1991). Despite these arguments for the effectiveness of SM, there has been little experimental evidence supporting claims that SM improves second language performance.

My search for studies that have manipulated SM and strategies similar to SM in the classroom yielded only two. Other studies offered anecdotal, theoretical, or correlational evidence. One of the empirical studies was by O’Malley, Chamot, Stewner-Manzanares, Russo, and Kupper (1985), who found statistically significant differences on speaking but not on listening posttests between treatment and control groups. The treatment group but not the control group was taught and directed to self- and peer-monitor their notes on the main points and cohesive markers of short speeches. However, only the treatment group was taught how to identify main points and cohesion markers. In other words, only the treatment group was explicitly taught structures and linguistic cues that would end up on the posttreatment test. This confounded strategy and structure instruction, strongly biasing the results in favor of a treatment effect. The time spent on strategy instruction and practice was 7 hours of class time during the fall semester. The total number of class hours in the semester was not mentioned in the study.

The treatment group in another experimental study (Viswat & Jackson, 1994) was also given more instruction than the control group, but in the form of “additional prelistening questions, which were designed to lead students to make use of the title, pictures, and information in the introductory statement to predict” (p. 241). Furthermore, significant teacher effects owing to different teachers teaching the strategies were
observed. Viswat and Jackson thought that this may have contributed to the large effect size of .57 in one of the treatment groups and felt that the study’s “results [were] not conclusive” but “promising” (p. 247).

Oxford (1992) and O’Malley and Chamot (1990) have documented much of the theory and research into L2 learning strategies such as SM, but most of these rest solely on theoretical arguments. The lack of experimental investigations into SM, and the bias due to the treatment groups’ added exposure to the target concepts, as in the two studies discussed above, points to the need for more research before claims can be made for the effectiveness of SM.

Lan (1993) and Schunk (1983) avoided the method bias present in the Viswat and Jackson (1994) and O’Malley et al. (1985) studies by providing external- or teacher-monitoring to the control group and comparing the results with those of the self-monitoring group. The control group was externally monitored by the instructor while each student in the treatment group monitored himself or herself, thus eliminating confounding effects due to one group receiving additional instruction. The present study, unlike previous monitoring or strategy instruction research in SLA, similarly controlled for bias arising from additional instruction.

Based on current SM theory, I assumed that the SM group would show greater improvement in class performance and thus score significantly higher on an end-of-term achievement test. To test this assumption, both groups were taught the same materials at the same rate and given the same test. The test scores were then compared to determine the impact of SM.

The time needed for the SM or strategy intervention was another variable considered. Time is an important factor in determining the efficacy of SM, yet in the SLA literature examined, the proportion of time required for SM or strategy training and administration was never reported. In a critical review of learner training, Rees-Millar (1993) addressed this issue, suggesting that if an excessive amount of time is required for strategy training, it may be better to use instructional resources for traditional tasks. The intervention in the O’Malley et al. (1985) study mentioned above required 7 hours in a semester. Although the total number of class hours in a semester was not mentioned, 7 hours is still a considerable amount of class time.

Time is an important factor in teaching decisions and needs to be included in discussions on the practical implications of SM intervention. Therefore, I measured the time required for the SM training and
implementation. This would not only provide an estimate of the time needed for SM, but also information on its relationship to the effect size of the intervention. To summarize, the following research questions were investigated in a tertiary education setting in Japan: (a) What is the effect of the SM of language-related behavior and performance on class test scores? (b) What is the amount of class time required to implement and employ SM?

**Method**

*Participants*

Each participant was a student at one of two schools: a Japanese co-ed university and a liberal arts college for women. The treatment group was comprised of two first-year English classes from the co-ed university and one first-year English class from the women’s college. The control group was also drawn from two first-year English classes at the co-ed university and one first-year class from the college. All the courses were required, but the students had some choice of which section to enroll in. The women’s college assigned students to classes based on their scores on a large standardized test, the General Tests of English Language Proficiency (G-TELP) (G-TELP Testing Services Center, 2002). Scores on this test for the treatment group were one standard deviation higher than the control group. As we shall see later, this difference had no bearing on the achievement test. All the students were in the humanities and relatively similar in academic achievement and educational background. None had had experience self-monitoring. For these reasons, the groups were considered equal for the purposes of the study.

There were 25, 21, and 17 students in the classes that made up the treatment group \((n = 63)\) and 26, 19, and 6 students in the control group \((n = 51)\) for a total of 114, excluding 6 students who did not write the test. From the women’s college, there were 21 students in the treatment group and 19 in the control group.

*Design and Procedure*

The SM protocol was designed for students in the treatment group to monitor their task behaviors and academic accuracy during class. Each student in the treatment group was given the protocol in the first class of the term. Course content and tasks were taken from *Fifty-Fifty Book One* (Wilson & Barnard, 1998), a beginning level grammar, speaking, and
listening text. The protocol items required the students to reflect, monitor, and record their task-related behaviors and outcomes. Behavior-related items that were monitored included (a) note taking of new vocabulary and structures, (b) the percentage of English spoken during speaking exercises, (c) the amount of eye contact during speaking exercises, (d) the students’ effort, (e) the homework starting and finishing times, and (f) attendance.

Academic-related factors that were monitored included (a) the perceived difficulty of homework, (b) the homework score, (c) the students’ perceived mastery of the homework, and (d) their perceived mastery of the text’s grammar, listening, and speaking exercises (these three aspects were evaluated separately). The students also wrote the chapter number and the chapter’s target structures (e.g. prepositions) on the protocol. Each student quantified his or her language-related behavior and academic performance as a percentage or as a point on a scale.

At the start of the course, the students were informed that the SM protocol would be evaluated solely on its completion and accuracy and would be worth 10% of the final grade. Points were not deducted from the protocol score if the homework was unfinished, or if the students failed to understand or achieve the target behaviors or tasks. However, they would lose points for falsely recording completion of their homework. I randomly checked approximately 70% of the protocols after every class and found that they were satisfactorily completed.

The same behavior- and academic-related factors that were self-monitored in the experimental group were also monitored by the control group. However, the control group’s behaviors and outcomes were monitored by the teacher (i.e. externally). To externally monitor, the teacher randomly selected students and asked them questions regarding the same behavioral and academic items that were on the SM group’s protocol. This was done at the beginning or end of class, depending on the specific item. For example, homework completion was monitored at the beginning of the class while perceived mastery of language structures covered in class was monitored at the end of the class. Due to time constraints, it was not possible to externally monitor the entire control group during every class, but all students were familiar with the monitoring items listed above and every student’s work was monitored periodically. If they were not being monitored, they were observing the other students being monitored.

Both groups met once a week for 90 minutes and covered the same material during the same week over a 14-week term. Throughout the course the monitoring was consistent and proximal to the tasks, both
important considerations in SM theory (Schunk, 2001). Test achievement, the dependent variable, was measured one week after the term finished. Written test items were constructed to measure the students’ knowledge of the speaking task target behaviors, vocabulary, and grammar structures covered during the course. There were 85 mostly closed items (three were open items related to direction and prepositions; there were no true/false or multiple choice items). The students took the same test at the same time at each institution. The control and experimental groups’ test scores were compared to see how the students’ self-monitoring and reflecting on the concepts studied and practiced in class affected their test scores.

**Results**

After collecting and examining the protocols at the end of the course, I found that 95% of them had been properly completed. Average attendance for all the classes ranged from 84% to 92%. An achievement test was administered to both groups and its reliability was checked. The Kuder-Richardson 21 reliability estimate for the achievement test was .88, indicating consistent item performance. The control group’s scores on the test were lower ($M = 44.8, SD = 12.49, n = 51$) than the SM or treatment group’s ($M = 49.0, SD = 12.13, n = 63$). An independent samples $t$ test assuming equal sample variances shows that at an alpha level of .05 the SM intervention did not result in a statistically significant difference between the two groups, $t(112) = 1.80, p = .07$ (two-tailed), $d = .34$. The 95% confidence interval of the difference of the means ranges from –8.76 to .41. The dependent variables were normally distributed and Levene’s test statistic ($p = .89$) indicates homogeneity of variances.

As reported earlier, the G-TELP placement scores for the two classes in the women’s college differed by an average of one standard deviation. The correlation between the G-TELP and the achievement test for the higher scoring class was $r = .16$ and $r = .20$ for the lower scoring class. These low correlations indicate that the two tests were unrelated and thus the one standard deviation difference in the G-TELP scores of the two women college classes at the beginning of the course should not have affected the class test scores. Any differences were likely due to differences between the test types, the G-TELP being a norm-referenced test measuring proficiency and the classroom test being a criterion-referenced test measuring topics covered over the length of the course. The SM group required 30 minutes training during the first class and approximately 12 to 15 minutes for completing and checking the protocol in each of the
subsequent classes. Over a 14-class, 21-hour term, this required approximately 4 hours or 20% of class time.

**Discussion**

The main purpose of this study was to determine the impact of SM on class achievement scores. The data in this and other empirical studies in both general education (e.g. Schunk, 1983; Shapiro & Ackerman, 1983; Susser, 1981) and SLA (e.g. Viswat & Jackson, 1994; O’Malley et al., 1985) support the hypothesis that SM has no significant effect on achievement. However, the confidence intervals and effect size (d = .34) show that there is a .95 probability that a small positive effect in addition to no effect may also be observed in the population.

The present study also set out to measure the amount of time required for the SM intervention. Perhaps more important than the failure to reject the null hypothesis—that SM intervention is of no use—was the approximately 4 hours or 20% of class time required for implementing and using the protocol. This result is consistent with the 7 hours needed for training in the O’Malley et al. (1985) study. In practical terms, educators considering SM or other metacognitive strategy instruction should anticipate the possibility of large time costs against minor or no increases in student achievement.

The sample size was admittedly small. However, the study was based on current theoretical arguments that SM causes significant changes in achievement. If this had been true, the sample size should have been sufficient. Assuming SM intervention results in effect sizes the same or greater than the current study, replication studies would need a sample larger than 270 students ($n_1, n_2 > 135$).

My concern that the higher standardized test scores (G-TELP, 2002) of part of the treatment group would influence their class test scores was alleviated by the extremely low correlations between the tests. In other words, the scores of the class achievement test were not influenced by differences in the institutional placement test scores of the students. The combination of behavioral and cognitive outcomes in the protocol was not a limitation, but measuring the separate effects of these two factors might tell us if one has a greater impact on achievement.

The sheer number of arguments for SM and strategy instruction give the appearance that they have been extensively researched, yet two important strategy researchers, O’Malley and Chamot (1990) acknowledge that “there has been little confirmation of the effectiveness of strategy
training with second language tasks” (p. 224) and to date there still has not been much empirical research. The results of the current experimental study also cannot confirm the effectiveness of SM strategy training.

The findings presented here indicate that further empirical investigations into the impact and time requirements of SM are needed before accepting the hypothesis that SM training is effective and practical in the classroom. Teachers must realize the possibility that there may be little or no effect from SM and that a considerable amount of time may be needed to implement it in the classroom.

Acknowledgments
I would like to thank the editor and the two anonymous referees of JALT Journal for their advice on the revision of this paper.

Terry Vanderveen teaches at Kanagawa University. His research interests include motivation and evaluation. He can be contacted at <teratkana@yahoo.co.jp>.

References


George Braine’s book *Teaching English to the World* (TEW) is part of a current line of publications with international perspectives on English language teaching (ELT) (e.g., Burns, 2005; Lin & Martin, 2005). However, unlike Burns’ (2005) collection, which documents and contributes to how teachers are increasingly part of the globalization of English, and that of Lin and Martin (2005), which brings a clearly critical perspective to bear on ELT policies and practices, Braine’s book does exactly what its subtitle suggests: it describes the history, curriculum, and practice of ELT in 15 diverse countries in Asia, Europe, the Middle East, and South America. As with Braine’s (1999) anthology, the focus of TEW is on nonnative English speaking (NNES) teachers and their experiences learning and teaching English. The authors follow a standard format, beginning with an introduction, an overview of the history of ELT in their country, then an outline of the past and present curriculum for English at various levels and in several contexts, and finally a biography (or autobiography) of one English teacher, usually the author. This consistency is helpful as readers go through various chapters, and it enables the authors to describe the historical and current situation in their countries personally.

Following a preface and introduction by the editor, the different chapters are arranged alphabetically using the countries’ names in English, so TEW starts with Kanavillil and Cristina Rajagopalan’s summary of ELT in Brazil and ends with Yasemin Kirkgoz’s outline of the situation in Turkey. In between we learn and read about ELT in China (by He An E), Germany (Claus Gnutzmann), Hong Kong (Icy Lee), Hungary (Peter Medgyes), India (Premakumari Dheram), Indonesia (Junaidi Mistar), Israel (Ofra Inbar-Lourie), Japan (Oda Masaki and Takada Tomoko), Lebanon (Kassim Shaaban), Poland (Joanna Radwanska-Williams and
Liliana Piasecka), Saudi Arabia (Khalid Al-Seghayer), Singapore (Antonia Chandrasegaran), and Sri Lanka (Minoli Samarakkody and George Braine). While space does not permit a summary of each chapter here, numerous themes echo throughout TEW, such as the importance of influential teachers in the authors’ lives, and the involvement of the British Council in many contexts.

TEW is commendable for many reasons, and teachers, researchers, and teacher trainers will find it to be a current and relevant resource. The overviews of ELT curricula are short but helpful, and the personal accounts of meeting the challenges of English learning and teaching are encouraging. What comes across is just how complex and diverse ELT is in specific countries, as it involves numerous levels of language proficiency, schools (including kindergartens, elementary, middle, and secondary schools, colleges and universities, all discussed in chapters here), and types of students (e.g., academic vs. vocational). It is also interesting to read about situations where English is the main or only foreign language (such as in Japan or Saudi Arabia) versus those where it is just one of several foreign languages (e.g., in Lebanon and Poland). Also communicated through this book is the way in which English and teachers of English have been involved, in various contexts, both in continuing and challenging some of the hegemonic views of English and ELT. Readers in Japan, for example, will identify with the impact of tests on English learning and teaching in places as similar or different as Hong Kong, Israel, and Lebanon.

While some people might view this collection as one largely by and for NNES teachers, as a native English speaker I believe this book is of relevance to various audiences. Teachers preparing to go to one of the countries above, for example, should consider the related chapter as required background reading, and teacher trainers who want to understand any of these ELT contexts will also find this anthology useful. TEW offers a wealth of information and numerous anecdotes on the ELT profession around the world, including the sobering realities of low salaries and high drop-out rates among ELT professionals (e.g., in China, Israel, and Hungary), as well as encouraging stories of how relevant English education is both to students and teachers. The importance of professional development is evident throughout the book, and I valued insights from writers such as Inbar-Lourie and Mistar, who share something of their failings as well as their successes. I also appreciated the references to research and works about ELT in various languages (apart from English), and the (auto)biographies of English teachers, with whom I found myself connecting on numerous levels—as a learner, teacher, researcher, and
teacher educator. Reading these surveys of a wide range of countries and ELT situations I also developed a renewed and growing respect for teachers in developing countries (e.g., Indonesia), who struggle with many realities that I have not faced in my own experiences in Africa, Asia, and North America.

A book with such a broad title may be criticized for reflecting the ELT situation in a rather limited way in only 15 countries, particularly through the eyes and experiences of just one (or, in three cases, two) writers per country. Braine admits that South America is only represented by one country (Brazil), and there is no chapter on a country in Africa or North America. As with any edited collection, some chapters in TEW are better written (and more easily read) than others, and the gems here include those by well-known authors like Gnutzmann and Medgyes, as well as those by individuals new to me, such as Al-Seghayer and Shaaban. Although the format is standard, some of the teachers' stories are written in the first person, while others are in the third person. There also seemed to be occasional anti-U.S. sentiments and in a few cases misguided or off-putting comments about native English speaking teachers (as in the chapter on Japan). Yet if TEW's chapters are read as personal narratives representing the views of the authors, readers will find much from which they can learn. I therefore highly recommend it.

References

“Real language problems need practical solutions” declares the back cover of Michael Swan’s *Practical English Usage*, and practicality pervades the book, from its straightforward approach to almost every aspect of the language to its example sentences that sound so natural you feel as if you are eavesdropping on a conversation. Now in a fully revised third edition, this classic reference work deserves a place on every teacher’s bookshelf, as well as in libraries for higher-level learners to consult. In his introduction, Swan sets out what the book does and does not do: “[It] is not a complete guide to the English language . . . its purpose is . . . to give learners and their teachers the most important information they need to deal with common language problems” (p. x). The information mainly addresses modern standard British English and “the book is not intended as a systematic guide to American usage” (p. ix). However, where American style varies it is noted; one entry covers the main differences between American and British usage.

For easier reference, this edition has a 6-page contents overview at the front (replacing the former alphabetical list of entries) and a 35-page index at the back. Another new feature is “Don’t say it! 130 common mistakes.” This contains 35 examples each at basic, intermediate, and advanced levels, plus 25 for very advanced learners, with corrections and keys to the relevant explanatory section. Each list of errors could form the basis for an in-class discussion and review. Entries have been updated, thus the entry on correspondence, which covers formal and informal letter writing, now has a second section on e-mail and text messages.

Of the 634 alphabetical entries, more than half are devoted to grammar. However, this is not a book to consult for prescriptive rules, which Swan says “are made by people who believe that they can improve a language or protect it against change” (p. 291). Swan’s rules are descriptive of what actually is happening to the language. A section on variation and change details some trends such as comparative *more* replacing -er forms, or *who* replacing *whom*. For complex items, such as modal auxiliary verbs, the reader can consult the general entry for an overview, and then look
up a particular verb for details; *may* and *might*, for example, take up eight sections over six pages. Spoken grammar is also covered in entries on contractions, short answers, tags, and reply questions.

In addition to American English, other varieties and styles included are standard English and dialects, correctness, spoken and written English, formality, abbreviated styles, slang, and taboo and swear words. In the latter category, the words are listed with their literal meanings and starred according to their strength. Swan cautions students to be aware not only that their use might offend their interlocutors, but also that “a foreigner who uses swear words may give the impression of claiming membership of a group that he or she does not belong to” (p. 565). Teachers of reading will appreciate the extensive section on news headlines, in which specific features of their grammar and vocabulary are decoded.

Areas I found of particular interest were the multiple entries on determiners, those little words that often cause big problems for my students; the lists of basic meanings of common and useful prefixes and suffixes, which could provide source material for exercises to expand vocabulary and improve reading comprehension; and the eight-page section on discourse markers, where I learned, among other things, that *e.g.* actually stands for *exempli gratia*. There are also entries on politeness including “social” language which contains many useful conversational gambits, distancing verb forms, and softening expressions. Topic areas include age, dates, names and titles, telephoning, and telling the time. Spelling, punctuation, and pronunciation are also addressed. A large number of entries deal with lexical problems, such as easily confusable words or expressions. The following are a few examples from over 80 entries: *accept* and *agree*; *besides, except, and apart from, efficient and effective; maybe and perhaps; say and tell; sensible and sensitive*. Learners may find these the most helpful sections, since this information is not often found in dictionaries. Over a hundred words are dealt with individually, including warnings where there may be “false friends.”

The last point above reflects an occasional Eurocentric bias often found in publications from Britain, and has less relevance to Japanese learners, being primarily an issue for speakers of languages that contain words that look the same in English. Perhaps there is a need for a guide to those English loan words whose meanings have changed in Japanese. One area I would like to see expanded is the entry on idioms, collocations, and fixed expressions. At only two pages, this seems slight, given the current emphasis on lexical approaches to language learning. These are, however, minor points and do not constitute any weakness of the book.
Even a work as comprehensive as this cannot cover everything, so Swan frequently recommends that students consult a good dictionary, or sometimes a grammar reference work, for more detail. Throughout the book, there is extensive cross-referencing which can lead to fascinating, if at times distracting, reading. All in all, I found it an awe-inspiring book for its sheer scope and attention to detail. The style is clear, giving it the distinction of being a reference book that is also a good read. I anticipate it being a source of vital information for years to come.


Reviewed by
Cheryl Kirchhoff
Meiji Gakuin High School

Hughes’ collection of essays by leading researchers in the field of speech discourse is a stimulating book for teachers who are thinking critically about teaching spoken English. The essays pursue issues of pronunciation, World Englishes, vague language in speaking, analysis of classroom discourse, and the role of speech in language learning theory. Hughes’ previous book, _Teaching and Researching Speaking_ (2002), complained that, “The role of mode and the status of speech in language acquisition paradigms has been remarkably undertheorized” (p. 24). Her new book responds to that need and showcases authors who are leading the way to new theories and applications for spoken English in TESOL. The aims of the book are twofold: one, to highlight current research and issues on spoken discourse, and two, to discuss how these insights will affect language teaching. Most of the essays are faithful to these aims and discuss current qualitative research, implications, and ideas for application to classroom teaching.

The twelve essays are divided into four parts, each focusing on an issue of spoken language research. Part 1 deals with native speaker attitudes towards language learners, particularly the state of World Englishes and negative attitudes that hinder communication. Authors in Hong Kong describe their experiences over the last 10 years as the World Englishes
(WE) paradigm has been replaced with a priority on native-like accent. Unique sociopolitical influences behind this change are uncovered leaving the authors with doubts about the future of WE. The second essay shows evidence that communication problems between native and non-native speakers can be caused by negative attitudes of the native speaker. Successful communication requires both interlocutors to make an effort. The researcher found that native speakers with negative attitudes toward nonnative speakers use various avoidance strategies that hinder communication. Both essays suggest that all English users, particularly native speakers, should be made aware of worldwide varieties of English. This view is echoed by Canagarajah (2006) in “TESOL at forty: What are the Issues?,” in which he describes the need to teach students to be able to “shuttle between communities” in our multidialectal world (p. 26).

Part 2 focuses on elements of prosody: intonation and speech rhythm and how they relate to turntaking in conversation. Ann Wennerstrom, in her essay on intonation, reviews reasons why intonation is essential for comprehension and comprehensibility. Intonation organizes discourse “much as the formatting conventions of subtitles and paragraph indentations do in written discourse” (p. 75). Intonation also shows the speaker’s intentions at sentence endings, for example, rising intonation with echo questions or a flat extension when hesitating in a word search. Her microanalysis of a Japanese learner revealed how the learner’s analytic approach to speech hindered intonation and comprehensibility. She suggests explicit instruction of lexical chunks and intonational phrases to improve comprehensibility. Teachers’ “foreign talk” to learners was also analyzed and found to contain unnatural intonation. Recent research on rhythm in speech is presented in an essay by Ee Ling Low. The former way of describing languages as “stress timed” and “syllable timed” is now acknowledged to be less a dichotomy and more a continuum. Investigation of speech rhythm revealed that nonnative speakers have far fewer reduced vowels in their speech when compared to native speakers.

Part 3 highlights several aspects of teaching spoken discourse including the need for prioritizing speech over writing in language teaching, the overlooked but important role of vague language in speaking, and new approaches to analyzing classroom discourse. Anna Mauranen’s essay advocating a greater priority to be placed on spoken language is particularly thought provoking. She recounts how research into aspects of speech (pragmatics, discourse analysis, conversational analysis) have “revitalized linguistics and challenged the adequacy of sentence-based models which have developed from analyzing written language or
invented sentences” (p. 143). Increasing globalization has resulted in a growing need for teaching spoken communication. New corpus studies on spoken language and the analysis of conversation have given evidence that speaking is highly organized and requires more skills than writing. She challenges teachers and linguists to acknowledge that in order to equip learners with the skills needed for real-world language use, spoken language needs to have priority over written language.

The fourth and final part of the book focuses on assessment, evaluating learners’ speaking progress, and measuring learners’ mental functions. Neither article offers useful insight for assessing speaking in the classroom. John Levis’ essay criticizes existing pronunciation assessment methods as haphazard or for having relative standards. He suggests evaluating for intelligibility and fluency, but gives no clear direction on how to measure these qualities. The final essay describes sociocultural theory and suggests new ways to assess psychological processes involved in language learning.

For teachers desiring to keep up with current research trends in speech pedagogy this book presents a stimulating collection of essays. While not a book of practical teaching applications, it does give plenty of food for thought about deeper or broader issues concerning teaching speaking.

References

Reviewed by
Byron O’Neill
Kyoto Notre Dame University

The Language of Language: Core Concepts in Linguistic Analysis (2nd ed.) is a student-centered, interactive textbook based on author Madalena Cruz-Ferreira’s lecture notes for an introductory undergraduate linguistics course at the National University of Singapore. It consists of 12 chapters that can be organized into 6 sections—an introduction, morphology, phonology, grammar, meaning, and language usage. The authors state in the preface that their goal is to “encourage informed thinking about issues relating to language structure and use, by discussing as broad a sample as possible, in a book of this size, of the kinds of activities that linguists busy themselves with” (p. xiii). They achieve this by assuming no prior knowledge of the field and by being selective in what they include.

Each chapter is divided into seven parts. The first part is the “Chapter Preview,” where a series of open-ended questions set the tone. An “Introduction” then serves to create a link from any previous chapters before presenting the “Main Content,” which is separated into numbered subsections for ease of reference. Multiple “Activity” boxes provide thought-provoking tasks that invite readers to review and interact with the information presented so far. “Food for Thought” presents relevant quotations from renowned linguists. Finally, a short bibliography for “Further Reading” is followed by a list of “References.”

The first section of the book serves as an introduction to the study of language by describing its universality and diversity. Chapter 1 begins by defining language as a mode of communication between humans and the various forms and functions that this involves. A brief examination of some key features of language is followed by an introduction to the scientific examination of language. Chapter 2 explores how languages are related to each other through discussions of variation, the classifications of language families, language shift, and universal grammar.

The next two sections discuss how linguistic meaning is presented through words, which acts as a foundation for a discussion on speech
sounds. Chapter 3 is an introduction to morphology through an examination of English word classes and morphemes. Chapter 4 looks at how and why new words are constructed. Chapter 5 defines phonetics and phonology before showing how speech sounds are biologically produced. The different types of sounds are discussed before the reader is shown how they could be transcribed and analyzed with the International Phonetic Alphabet (IPA) and Distinctive Feature (DF) approaches. Chapter 6 closely examines the role phonemes play in everyday language use and introduces the reader to phonemic systems and analysis.

The fourth section addresses sentence-level grammar. Chapter 7 begins with syntax and constituency, and follows up with phrases, phrase structure, and recursion. Tree diagrams are used to illustrate some of the more difficult concepts. Chapter 8 deals with sentence structure by discussing syntactic form and function, obligatory and optional sentence constituents, and verb categories.

Semantics and pragmatics are explored in the next section. Chapter 9 initially looks at the concept of meaning through four parameters: sense vs. reference, denotation vs. connotation, literal vs. metaphorical, and compositional vs. idiomatic. Semantic shift and the relationship between meaning and sound, spelling, other words, and between sentences is also discussed. Chapter 10 looks at the functions that language performs with information on speech acts, the Cooperative Principle, politeness, and the organization of conversation.

The final section gives an overview of discourse and language communities. Chapter 11 starts by examining the structure of information within and across clauses before analyzing discourse patterns. Chapter 12 serves as an observation of groups of language speakers categorized as native, bilingual, and The Others, a euphemism used by the authors to describe nonnative users of language. Language loss, language endangerment, and language death are also covered.

*The Language of Language: Core Concepts in Linguistic Analysis (2nd ed.)* explores the nature of the English language, and suggests ways in which students and linguists can study it. My only criticism is that it focuses much more on the former than on the latter. This book would be more valuable with less emphasis on the mechanics of language, which the authors feel is a necessary prerequisite to understanding the science of its examination. Perhaps this is so, but for those interested in the topics suggested in the subtitle, a more balanced approach might be more effective.
Henry Widdowson, the father of applied linguistics, has written *Text, Context, Pretext* to give a wake-up call to the field of discourse analysis. In this volume, Widdowson compellingly argues that discourse analysis in its best known faces is being confused with interpretation, displaying a *functional fallacy*, as he refers to it. As a result of poorly thought out methodologies and guiding philosophies, the field is in disarray, and runs the risk of being sidelined altogether, if it does not present a more objective and systematic treatment of analyzing language.

On the face of it, the application of this book to classroom teaching is tenuous. While Widdowson’s work (1984; 1990) in applied linguistics has been influential, the current volume makes no direct reference to language teaching. Even his past work (especially Widdowson 1979; 1984) has had limited impact on language teaching, as Widdowson’s writing tends to be quite dense, with long-winding arguments not organized in a more reader-friendly manner. For example, I feel that Widdowson’s important dichotomy of *genuine texts* versus *authentic practice* is still misapplied in the ELT literature and remains the minority view in discussions regarding these topics for this very reason. Potentially, this book could be relegated to a similar status due to its style of presentation. On the other hand, how to treat texts, what is context and how it shapes language, what agendas speakers have—all of these are very important questions that teachers need to keep in the back of their minds as they prepare their lessons. Such questions inform the relevance, the authenticity, the heart of curriculum design by asking: what shall be the content of our teaching?

The book is divided into 10 chapters, the first 5 dealing with the concepts outlined in the title of the book. In Chapter 1, *text* and *discourse* are differentiated. Discourse is the process of interpretation, of meaning negotiation, between people. The language that emerges from this process is text. This reiterates the position that Widdowson (1979) takes earlier, that *text*-ness does not reside as an intrinsic quality of the text, but is instead a quality conferred by its users.
In Chapter 2, Widdowson tackles Halliday’s systemic functional (S/F) linguistics. While S/F is credited for its description of language components, Widdowson insists that these components do not code particular functional interpretations. The reason for this is that such evaluations can only be made with due reference to context, the content of Chapter 3. In this chapter, a historical analysis of how the concept of context is treated by significant contributors, as is how relevance theory (RT) only chooses to pay attention to particular aspects of context in its judgments of what texts “mean.” Corpus analysis (CA), the examination of what kind of language patterns, especially lexical, tend to occur in similar contexts, is the subject of Chapter 4. In S/F, the particular language used in a particular setting for a particular purpose is referred to as genre, which I would suggest is conceptually similar to co-text, but of which Widdowson makes no mention. Finally, “pretext,” the agenda or intention that language users bring to the table when negotiating meaning, and its impact are discussed in Chapter 5.

The final five chapters discuss how different approaches to text and discourse analysis have fallen short of being systematic, and thus have lost credibility. Chapter 6 criticizes critical discourse analysis (CDA) for being selectively attentive to particular features of text and context—in effect, for using a politically-motivated pretext in its methodology. CDA brings so much data to bear on a text, that researchers are forced for practical reasons to focus on individual language features to study. The resulting discriminate choosing of features is reminiscent of CDA’s failing, and this dilemma is described in Chapter 7. “Analysis and interpretation” is the focus of Chapter 8: what do these concepts mean? Where many linguists criticize literary critical interpretation, Widdowson defends it as an approach that unabashedly makes private interpretations. He goes on to contend that since these interpretations are universally different, they cannot be subjected to devaluation for simply being academic, nor can CDAs be elevated for being exclusively sociopolitical.

In Chapter 9, the methodologies of the different approaches to text and discourse analysis are subjected to a final review and are found lacking given their selective attention to language features which support the view that they have brought to the text. Widdowson concludes in Chapter 10 by calling for more scientific (i.e. systematic and objective) methodologies, which in turn would be more credible and authoritative for public consumption.

To my recollection, I have not read a more comprehensively scathing review of ELT-related research anywhere, as this volume has not exactly
followed the guidelines for winning friends. But because it does not pull
any punches, it will influence thinking. Coupling wit with incisive logic
supported by specific and numerous illustrations, this volume scores
some mighty hits on current approaches and eminent names in the field
of linguistics. Reading this book is like having someone pull the curtains
back on the political and ideological battles being waged in academia.
The book is an education in these epic scenes, but I am glad not to be one
of the actors referred to in its pages.

*Text, Context, and Pretext* is not light reading, but reading it will reward
the committed language researcher, student of linguistics, and even class-
room teacher. I had to read Chapter 3 several times before I could make
sense of it (and that only after proceeding to read Chapter 6 and 7 and
then going back). Yet, the book was stimulating for its innovative treat-
ment of concepts and aggressive use of style, and I would recommend it
as a must-read for anyone with an interest in how language is and should
be evaluated, in a spectrum ranging from the sociopolitical (for instance,
Lakoff, 2004) to the educational and language learning realms.

**References**

Lakoff, G. (2004). *Don’t think of an elephant: Know your values and frame the debate—

Press.

University Press.

Press.
Information for Contributors

All submissions must conform to JALT Journal Editorial Policy and Guidelines.

Editorial Policy

JALT Journal, the refereed research journal of the Japan Association for Language Teaching (Zenkoku Gogaku Kyoku Gakkai), invites practical and theoretical articles and research reports on second and foreign language teaching and learning in Japanese and Asian contexts. Submissions from other international contexts are accepted if applicable to language teaching in Japan. Areas of particular interest are:

1. Curriculum design and teaching methods  
2. Classroom-centered research  
3. Cross-cultural studies  
4. Testing and evaluation  
5. Teacher training  
6. Language learning and acquisition  
7. Overviews of research and practice in related fields

The editors encourage submissions in five categories: (1) full-length articles, (2) short research reports (Research Forum), (3) essays on language education or reports of pedagogical techniques which are framed in theory and supported by descriptive or empirical data (Perspectives), (4) book and media reviews (Reviews), and (5) comments on previously published JALT Journal articles (Point to Point). Articles should be written for a general audience of language educators; therefore statistical techniques and specialized terms must be clearly explained.

Guidelines

Style

The JALT Journal follows the Publication Manual of the American Psychological Association, 5th edition (available from APA Order Department, P.O. Box 2710, Hyattsville, MD 20784, USA; by e-mail: <order@apa.org>; from the website: <http://www.apa.org/books.ordering.html>). Consult recent copies of JALT Journal or TESOL Quarterly for examples of documentation and references. A downloadable copy of the JALT Journal style sheet is also available on our website at <http://www.jalt-publications.org/jj/>.

Format

Full-length articles must not be more than 20 pages in length (6,000 words), including references, notes, tables, and figures. Research Forum submissions should be not more than 10 pages in length. Perspectives submissions should not be more than 15 pages in length. Point to Point comments on previously published articles should not be more than 675 words in length, and Reviews should generally range from 500 to 750 words. All submissions must be typed and double-spaced on A4 or 8.5”x11” paper with line spacing set at 1.5 lines. The author’s name and identifying references should appear only on the cover sheet. Authors are responsible for the accuracy of references and reference citations.

Submission Procedure

Materials should be sent as an e-mail attachment in Rich Text Format (preferred) or by post. Postal submissions must include a clearly labeled floppy disk or CD-ROM and one printed copy. Please submit materials to the appropriate editor indicated below.

Materials to be submitted— all of this can be sent by e-mail in one attached RTF file (preferred) or by post. If by post, be sure to include all of the material on a floppy disk.

1. Cover sheet with the title and author name(s).
2. One (1) copy of the manuscript, with no reference to the author. Do not use running heads.
3. Contact information sheet, including one author’s full address and, where available, a fax number and e-mail address.
4. Abstract (no more than 150 words).
5. Japanese translation of the title and abstract, if possible (no more than 400ji).
6. Biographical sketch(es) (no more than 50 words each).
7. Authors of accepted manuscripts must supply camera-ready copies of any diagrams or figures and a copy of the manuscript (RTF or ASCII).
Evaluation Procedures
Submissions will be acknowledged within one month of their receipt. All manuscripts are first reviewed by the editorial board to ensure they comply with JALT Journal Guidelines. Those considered for publication are subject to blind review by at least two readers, with special attention given to: (1) compliance with JALT Journal Editorial Policy, (2) the significance and originality of the submission, and (3) the use of appropriate research design and methodology. Evaluation is usually completed within three months. The first author of a published article will receive one copy of the journal and 10 off-prints with the option to order more at a set rate (contact JALT Central Office for price details). Review authors receive one copy of the journal.

Restrictions
Papers submitted to JALT Journal must not have been previously published, nor should they be under consideration for publication elsewhere. JALT Journal has First World Publication Rights, as defined by International Copyright Conventions, for all manuscripts published. We regret that manuscripts or computer disks cannot be returned. If accepted, the editors reserve the right to edit all copy for length, style, and clarity without prior notification to authors.

Full-Length Articles, Research Forum, and Point to Point Submissions
Please send submissions in these categories or general inquiries to:

jj-editor@jalt-publications.org
Steve Cornwell, Ed.D; JALT Journal Editor

Perspectives
jj-editor2@jalt-publications.org
Deryn Verity, Ph. D; JALT Journal Associate Editor

Reviews
The editors invite reviews of books, tests, teaching systems, and other publications in the field of language education. A list of publications that have been sent to JALT for review is published monthly in The Language Teacher. Please send submissions, queries, or requests for books, materials, and review guidelines to:

jj-reviews@jalt-publications.org
Yuriko Kite, Ph. D; JALT Journal Reviews Editor

Japanese-Language Manuscripts
JALT Journal welcomes Japanese-language manuscripts on second/foreign language teaching and learning as well as Japanese-language reviews of publications. Submissions must conform to the Editorial Policy and Guidelines given above. Authors must provide a detailed abstract in English, 500 to 750 words in length, for full-length manuscripts and a 100-word abstract for reviews. Refer to the Japanese-Language Guidelines for details. Please send Japanese-language manuscripts to:

jj-editorj@jalt-publications.org
Yoshinori J. Watanabe, Ph.D; JALT Journal Japanese-Language Editor

Address for Inquiries about Subscriptions, Author Reprints, or Advertising
JALT Central Office
Urban Edge Building 5F
1-37-9 Taito, Taito-ku, Tokyo 110-0016, Japan
Tel.: 03-3837-1630; Fax: 03-3837-1631
(From overseas: Tel.: 81-3-3837-1630; Fax: 81-3-3837-1631)
E-mail: jco@jalt.org URL: http://www.jalt.org
日本語論文投稿要領

JALT Journalでは日本語で執筆された論文、研究報告、実践報告、書評等を募集しています。
文体: 一般的な学術論文のスタイルを用い、章立ての仕方や参考文献のデータの書き方などは、Publication Manual of the American Psychological Association (5th ed.)の定める方式に合わせて下さい。不明の場合は、JALT Journalの英語論文を参考にするか、日本語編集者までお問い合わせ下さい。なお、JALT Journalの読者は現場の教師が主なので、特殊な専門用語や統計的手法は、わかりやすく定義するか説明を加えるなどして下さい。

提出するもの:
- 原稿3部（執筆者名は書かないこと）
- 執筆者の名前と所属機関名を書いた表紙
- 連絡先（住所、電話番号、ファックス、e-mailアドレス）
- 400字以内の和文要旨
- 英文のタイトルと、500〜750語の英文要旨（書評の場合は100語程度の英文要旨）
- 100字以内の執筆者略歴
- MS-Wordで保存したファイル（マッキントッシュ使用の場合はrtfファイルで保存したもの）

査読：編集委員会で投稿要領に合っているかどうかを確認したあと、少なくとも二人の査読者が査読を行います。査読者は執筆者の名前は知らされません。査読の過程では特に、原稿がJALT Journalの目的に合っているか、言語教育にとって意味があるか、独創性はあるか、研究計画や方法論は適切か等が判定されます。査読は通常二か月以内に完了します。

注意：JALT Journalに投稿する原稿は、すでに出版されているものや他の学術雑誌に投稿中のものは避けて下さい。JALT Journalは、そこに掲載されるすべての論文に関して国際著作権協定による世界初出版権を持ちます。なお、お送りいただいた原稿は返却しませんので、控を保存して下さい。

投稿原稿送り先またはお問い合わせ：
JALT Journal 日本語編集者 渡部良典
jj-editorj@jalt-publications.org