

流暢性と語彙フレーズ, hedges, との関係について

～日本人英語学習者コーパスからの示唆～

西川 恵

東京外国語大学大学院博士前期課程修了

1. これまでの経緯と現状

1-1. これまでの経緯

東京外国語大学 21 世紀 COE プログラム「言語運用を基盤とする言語情報学拠点」では、現在「TUFSS 言語モジュール」と称する多言語 e-learning 教材を開発中である。この教材開発作業と並行した基礎研究として日本人英語学習者 (JLE) の英語習得プロセスを解明するための学習者言語分析も必要であると考え、近年学習者言語研究のための方法論として注目されている学習者コーパスを作成することとなった。

基礎研究開発開始初年度であった 2003 年度は、今後の学習者言語分析研究の土台となる英語学習者コーパス、特に JLE コーパスに関する先行研究のレビューを行った (植田他, 2004)。JLE の言語データを扱った学習者コーパスはまだ絶対数が少なく、規模が小さいものがほとんどで、書き言葉データがそのほとんどを占めているという問題点があり、よって今後は既存のコーパスのデータ拡充に加えて、自然習得状況で英語を学んだ学習者のデータコーパス、様々なコンテキストにおける話し言葉データコーパス、初級学習者のデータコーパスなどの新しいコーパスの開発が必要であると報告した。

2004 年度は、これらの学習者コーパス作成に向けて実際に JLE から言語データを収集し分析する作業を開始した。言語データは、高校生、大学生、英語圏からの日本人帰国子女から収集された話し言葉のデータである。日本で英語を学習する高校生、大学生による言語データが COE プログラムの教材開発に役立つことはここで説明するまでもなからう。また、帰国子女の言語データも自然習得環境で高レベルの習得を達成した一日本人英語学習者の言語データとして大変貴重な資料であり、日本国内で英語を外国語として学習する一般的な日本人英語学習者にも応用可能な、教材開発のみならず英語教育全体への示唆が得られるであろう。2005 年現在もデータの収集は続いている。

1-2. 本報告書について

データ収集は続いているが、現段階では収集後文字化の作業を終えるにとどまっておりデータを分析する段階には至っていない。そこで、本報告書では上の JLE コーパスの中から帰国子女コーパスを取り上げ、言語資料を観察し、その特徴を2点挙げ、どのようなコーパス研究のための仮説構築が可能なのかをまず検討してゆく。ここでは2つの構築可能な仮説を挙げる。これらの仮説を基に今後は研究を発展させていく予定である。

さらに、今後これら2つの仮説を実際に検証する際、どのようなデータをどのように分析するのか、どのような結果が想定されるのか、その結果をどのように解釈可能なのか、今後改善すべき問題点などを議論する。最後に流暢性と語彙フレーズ及び hedges との関係がどのように英語教育に貢献するのか、TUFSS 言語モジュールの教材開発作業のために役に立つ可能性があるのかを考える。

1-3. 2004 年度のデータ収集経過報告

帰国子女コーパスの言語データ提供者は3人の日本人帰国子女¹である。少年と蛙についての物語の絵本シリーズを使用し、英語ネイティブスピーカー（以下 NS）と自由会話（約30分）、および絵本を使った物語描写（約10分）をさせた。データ収集に入る前にウォームアップ活動として5分程度 NS と会話した後で自由会話に入る。自由会話の後、絵本を使った物語描写をする。被験者ははじめに絵本を終わりまで見る時間を与えられ、その後はじめのページから英語で物語を描写する。2004年6月に行われた第1回のセッションに続き、約3ヶ月毎にデータ収集が行われた。本報告書では3人のうちの1人の物語描写のタスクについて扱う²。

1-4. 収集データから観察された特徴

帰国子女による発話は一般的に流暢性³が高いとされることが多い。これは話し言葉を媒体に自然習得したことによると思われる。この3人についても例外ではなく、流暢性が高いとみなされるものであろう。

流暢性以外の大きな特徴としては、

- ① 帰国後時間が経つにつれ kind of, sort of, like, stuff といった hedges, の使用回数が増加する傾向にあること、

- (1) and then he's *sort of* sad and mad.

- (2) to take off *like* the fancy clothes and *stuff*

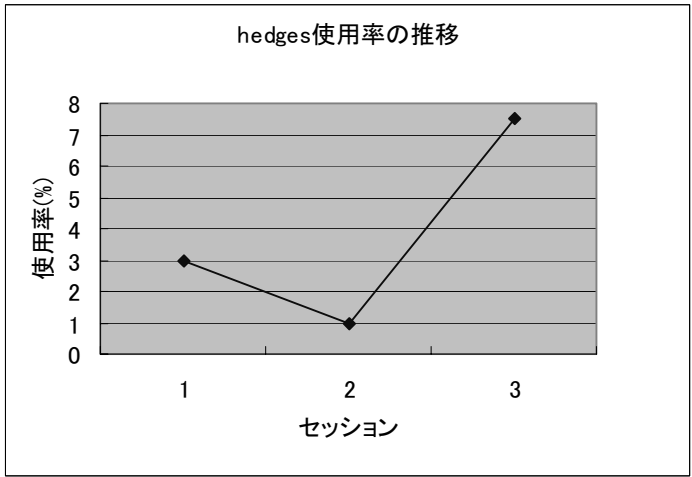
が挙げられる（グラフ1参照）。Hedges とは曖昧な概念を説明したり、はっきりとした

¹ 3人（全員女子）のうち2005年12月現在1人は中学3年生、2人は中学2年生。データ収集開始当時は中学2年生と1年生であった。

² データ収集を開始したのが帰国直後からであったこと、3人のうち発話総語彙数、使用語彙バラエティーともに最も高かったことという2つの理由からこの1人を選んだ。

³ 何をもち「流暢性」が高いとするのか、その定義については後に説明を加える。ここではあくまでも直感的に漠然と「流暢である」と判断されるレベルのものである、ということである。

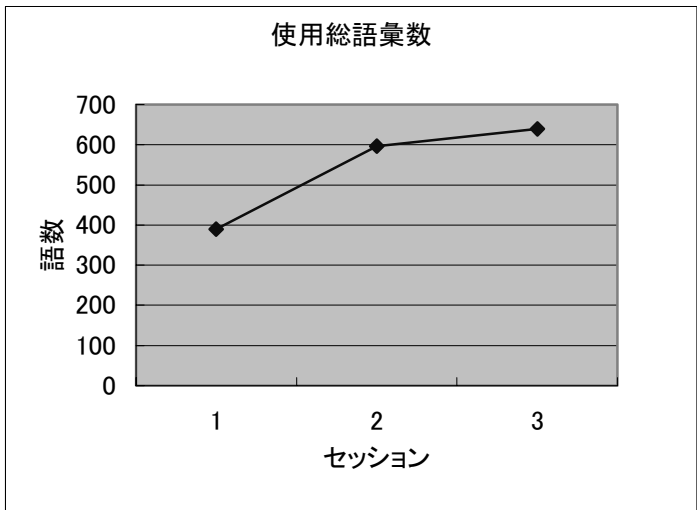
断言を避けたい場合に使用されるもので、日本語では「～のような、～みたいな」にあたるものと考えられる。



グラフ 1：hedges 使用率の推移

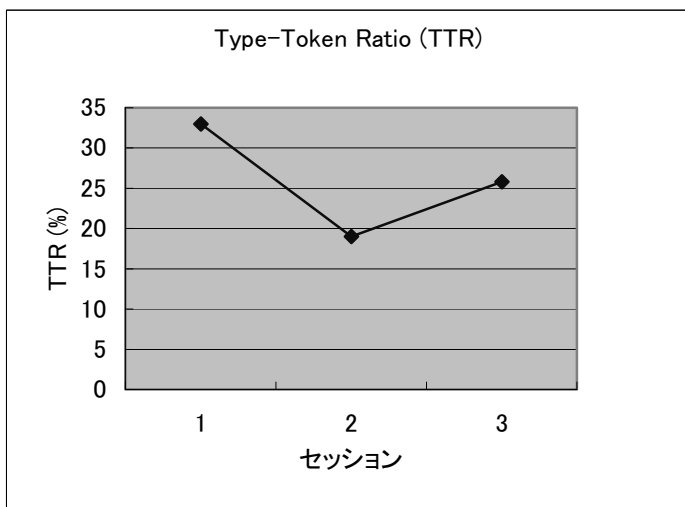
さらに 2 点目の特徴として、

- ② 発する総語彙数は増えても使用語彙のバラエティー⁴は急激に低下する傾向にあること、
がある（グラフ 2 および 3 参照）。



グラフ 2：使用総語彙数の推移

⁴ 使用語彙バラエティーは、同じ語彙の複数回の使用を 1 回とした延べ語彙数 (type) を総語彙数 (token) で割った type-token ratio (TTR) で測定した。



グラフ 3 : Type-token ratio の推移

観察された特徴①から, hedges の使用は流暢性と何らかの関係があるのではないかと考えた。高い流暢性は語彙フレーズの使用と関係しているという研究報告はこれまでもされている。詳しい説明は後に行うが, 先行研究によると語彙フレーズを使用することでチャンキングによって記憶容量の負担を減らすことが可能となり流暢性が高くなるという。Hedges と流暢性との関係については, 現段階ではまだ先行研究があまりなく明らかになっていない。しかし, 語あるいは語句を記憶して使用し記憶容量の負担を軽減しているという点では hedges と流暢性の関係は語彙フレーズと流暢性の関係と非常に似ていると考えられよう。

また特徴①は特徴②から裏付けられるとも捉えられる。英語喪失が進む過程で語彙検索や産出に時間がかかるようになり, 暗記している hedges を使用することで流暢性を保とうとするのではないかと考えられる。

さらに, 特徴①および②から流暢性そのものの定義を考える必要があるのではないかと考えた。沈黙することなく言葉を発していることが「流暢である」と判断されることが多いが, はたして本当にそれだけで「流暢性」が高いと言えるのだろうか。セッションが進むと発する語彙数は増えているので, 高い流暢性を発話語彙数とすると流暢性は高まっている, とされる。しかし, 使用されている語彙のバラエティーは急激に低下している点を考慮すると, 定義を使用語彙のバラエティーとすれば発話語彙数が増加しても流暢性は低下している, となる。

次のセクション 2 では流暢性とチャンキングそして語彙フレーズに関する先行研究をまとめ, あわせて hedges と流暢性についても語彙フレーズと流暢性との関連との類似点を挙げながら関係の可能性を考える。発話語彙数と使用語彙バラエティーについては, 日本

人帰国子女の言語喪失研究分野から語彙喪失と文法喪失に関する研究を紹介し、流暢性の定義と絡めながら研究仮説を導き出す。

2. 語彙フレーズと流暢性に関する先行研究及び hedges と流暢性との関連性

帰国子女コーパスのデータ提供者である3人は一般的な日本人英語学習者と比べると高い流暢性が観察されると前セクションで述べた。さらに hedges の使用も観察されたがこれは流暢性と関連しているのだろうか。このセクションでは語彙をひとまとまりで処理する過程、つまり語彙フレーズの使用のメカニズム、そしてその流暢性との関係についての研究を紹介し、その後 hedges との類似点についても議論する。

2-1. 語彙フレーズと流暢性

2-1-1. チャンク

人間の言語情報処理と記憶のメカニズムは密接に関係しており、その記憶容量には限界があるとされている。一度に記憶できる容量は memory span とよばれ、数であれば 7 ± 2 個程度であると言われている (Miller, 1956) が、記憶すべき対象にまとまりを作ることによって要領よく記憶することが可能となる。例えば、NBCCBSABC を9個のアルファベットの連鎖という処理単位 (processing units) で記憶しようとするのが難しいが、NBC-CBS-ABC といったまとまりを作ると処理単位が3個となり記憶しやすくなる。この処理単位の形成を、チャンキング (chunking) と呼ぶ。チャンキングの大きな利点は、処理の効率性 (processing efficiency) にある。言語処理過程においては、個々の語一つ一つではなく、いくつかの語をまとめて一つの単位として処理することが、それを認識、貯蔵、検索、産出する時間を節約することにつながり、自動性 (automaticity) と流暢性 (fluency) に貢献する (Ellis, 2001)。

Pawley and Syder (1983) は、英語母語話者は頻繁に使用する語彙チャンクを一つのまとまりとして長期記憶に貯蔵し、実際の言語使用の場面での言語処理負担を軽減させているのだと指摘する。言語の規則を理解し、文生成を創造的に行うことに加えて、記憶している無数の語彙チャンクを自在に検索、産出することで流暢に話すことができるのだという。また、このような語彙チャンクは語用論的機能とも密接に結びついており、場面に応じて適切な語彙チャンクを選択できる能力は言語運用上重要であると考えられている。Sinclair (1991) は、前者のチョムスキー的言語観に立った文生成の創造的な側面には自由選択原理 (open-choice principle) が、後者の語彙チャンクを記憶して使用している側面には非選択的な語彙原理 (idiom principle) が働いていると説明している。文生成は自由選択原理により基底の規則体系システムに基づいて創造的になされるが、談話における語彙選択には自由選択原理ではカバーできない制約があり、記憶している多数の語彙チャンクを単独で使用する非選択的な語彙原理が働くとされる。伝統的に言語学では自由選択原理を研究対象としてきたが、近年、語彙原理も同様に重要であると考えられている。

2-1-2. multi-word item (MW I)

語彙原理の重要性がクローズアップされる中、近年語彙研究では multi-word item (MW I) の役割の大きさが注目されるようになった。一般的な MW I の定義は、Moon (1997) によれば、①その MW I がどのくらい慣例的に使用されているか (institutionalization)、②どのくらい固定化されているか、あるいは言い換えができないか (fixedness)、③ MW I を構成している個々の語レベルで MW I 全体の意味を推測できるか (noncompositionality) の 3 基準を満たすものとされる (表 1 を参照されたい)。この分類の中でも特に、prefabs として分類されている談話機能をもつ MW I の語彙学習における重要性が特に近年注目されている。

表 1 Moon (1997) にもとづく

MW I	例
compounds	Prime Minister, short-circuit, brown-eyed など。
phrasal verbs	単音節の動詞 (come, go, get, put, take など) + 前置詞 (of, in, on, down など) からなるもの。
idioms	spill the beans, have an axe to grind, kick the bucket など。
fixed phrases	of course, at least, in fact, by far, good morning, how do you do, excuse me, you know など上の 3 つのカテゴリーには入らないものや white as a sheet などの直喩, it never rains but it pours などのことわざもこれに分類される。
prefabs	the thing/fact/point is, that reminds me, I'm a great believer in ... のような決まり文句。固定化の度合いは異なるが、ほとんどが特定の談話機能を果たし、慣習化されている。

2-1-3. 語彙フレーズ (lexical phrase)

Nattinger and DeCarrico (1992) は語彙フレーズ (lexical phrase) という観点から MW I を整理し、英語教育への応用可能性を説明している (表 2 を参照されたい)。彼らによれば、言語項目は①syntactic strings: 統語能力によって生成される言語の全ての文法構造の基礎となるもの、②collocations: 共起性はあるが語用論的機能はないもの、③lexical phrases: 語彙と文法の間位置するもので、言語形式と語用論的機能を両方持つ語彙チャンク、に大別されるという。Moon (1997) の MW I の分類で fixed phrases と prefabs とされているものが、Nattinger and DeCarrico (1992) が語彙フレーズとするものにあたると考えられ、例えば、by the way (topic shifter) や not only X but also Y (relators) のように形と機能を有し、語彙と文法の間位置するものを指す。語彙フレーズの特徴は、何らかの語用論的機能を持っているという点である。よって、Prime

Minister や spill the beans のように共起性はあるが語用論的機能がないコロケーションやイディオムは語彙フレーズには含まれない。語彙フレーズには, what on earth など語の入れ替えが不可能な非生産的な語句と, 入れ替えが可能な Could you ...?, a year/month... ago などの生産的な語句がある。

表2 Nattinger and DeCarrico (1992) にもとづく

語彙フレーズ	説明	例
polywords	色々な機能を持つ固定化した短いフレーズ。	by the way (topic shifter) , so to speak (fluency device) , strictly speaking (evaluator) , you know (clarifier) , once and for all (summarizer) nevertheless, moreover, however (relator)
institutionalized expressions	比較的長い固定化した表現。ことわざ, 格言, 社会的活動に必要な言い回しなど。	how do you do? (greeting) , nice meeting you (closing) , give me a break (objection) , there you go (approval) , long time no see (greeting)
phrasal constraints	短～中程度の長さで, 1, 2ヶ所が色々な語句で入れ替えられるようになっている。	a [day/year/long time] ago (temporal relator) as far as I [know/can tell] (qualifier) , you [creep/jerk] (disapproval) , the [sooner/better/happier] the [sooner/better/happier] (comparator) , for [instance/example] (exemplifier)
sentence builders	スロットが埋めることで文を生成できるようにしている文の枠組み。	I think that X (assertion) , not only X, but also Y (relators) , my point is that X (summarizer) ,let me start with X (topic marker) , it's only in X that Y (qualifier) , [would/could] you [help/lift that]? (request) , it seems that X (assertion) , that reminds me of X (topic shifter)

2-2. Hedges と流暢性

2-1 で語彙フレーズと流暢性の関連性の可能性について述べたが, hedges と流暢性についても同様の関連性の可能性がある。Hedges と流暢性との関係については, 現段階ではまだ先行研究があまりなく明らかになっていない。しかし, 語あるいは語句を記憶して使用し, 記憶容量の負担を軽減するという点では hedges と流暢性の関係は語彙フレーズと流暢性の関係と非常に似ていると考えられよう。

また, hedges は対話者との意味交渉 (negotiation of meaning) をする過程での適切な語彙の検索時間 (lexical search time) が必要な際に現れることも多く (Poos and Simpson, 2002) , 何らかの機能を持つ語あるいは語句を記憶して使用しているというこの観点からも hedges と流暢性の関係は語彙フレーズと流暢性の関係と関連しているといえる。

Hedges と流暢性に関する先行研究を調査し整理することが今後の課題の一つである。

2-3. 語彙喪失と文法喪失

言語喪失研究分野⁵では、日本語を第1言語（L1）とし、第2言語（L2）として自然習得した英語を喪失してゆく過程を日本人帰国子女の言語データをもとに報告している研究がある。L1が日本語でL2が英語の場合、L2喪失過程においていくつかの特徴がみられるとされているが、その中で、文法よりも語彙の喪失の方が顕著であるという特徴が報告されている。（Yoshitomi, 1992; 1999; Tomiyama, 1999; 2000）。

実際に、今回の3人の帰国子女の言語データの特徴の一つとして、帰国後の経過時間が長くなるにつれ使用語彙のパラエティが急激に低下することが挙げられた。これはもう一つの特徴としてあげた、hedgesの使用回数が高くなることとも関連しているようである。

2-4. 構築可能な研究仮説

以上、Moon（1997）のMW IやNattinger and DeCarrico（1992）の語彙フレーズ概念と流暢性の関係をふまえて考えると、帰国子女は語彙の急激な喪失により流暢性が低下してゆくことを埋め合わせるために語彙フレーズやhedgesを効果的に使用している可能性がある。よって、流暢性と語彙フレーズおよびhedgesに関して、以下の2つの仮説が構築可能であると考えられる。

- a. 帰国後時間が経過すると語彙フレーズの使用率が高くなる
- b. 帰国後時間が経過するとhedgesの使用率が高くなる

はじめに述べたように現段階では帰国子女データを実際に分析する段階には至っていないが、セクション3ではこれらの仮説を検証するための研究計画案をたててみることにする。

3. 研究の方法

3-1. データ

今回取り上げた1人のデータを縦断的に追っていく。本報告書では物語描写タスクの3セッション分の経過を扱ったが、今後は自由会話タスクも加えた全セッションのデータを対象とする。タスクはNSとの自由会話（30分）と物語描写の2つの課題とし、3ヶ月ごとに収集するものとする。

3-2. 分析

Moon（1997）のMW IやNattinger and DeCarrico（1992）のまとめている語彙フ

⁵ 言語喪失研究とは人間が一度習得した言語を喪失してゆく過程の研究である。言語習得過程と表裏一体の過程であるとも考えられる言語喪失過程は、言語保持教育だけでなく言語習得過程や言語教育への示唆にも富んでいる。

フレーズの中のどのフレーズが、そしてどの hedge が何回目のセッションで何度使用されているのか、を計算する必要がある。語彙フレーズおよび hedges の各セッションでの使用回数をそれぞれ発話回数と総単語数で割り、語彙フレーズ使用率、hedges 使用率として計算する。この計算の際にはコンピューターソフト CLAN⁶を使用する。セッション毎の語彙フレーズ使用率、hedges 使用率の推移を折れ線グラフで表示する。

語彙フレーズと hedges に加えて、何をもって「流暢性」とするのかを検討するために、各セッションでの (1) 総単語数、(2) 使用語彙バラエティー (TTR)、(3) ポーズ率、(4) 言い直し率、(5) 繰り返し率、(6) 1 分間あたりの単語数、も計算し、多角的な視点から「流暢性」を測定する。

3-3. 結果と考察

結果の解釈はどの観点から「流暢性」を測定するかで大きく変わってくる。流暢性測定のために挙げた (1) から (6) までの項目を吟味し、セッションが経過するにつれてどの観点からみた「流暢性」が実際に低下するのかわきについてまず議論する必要がある。

セッションが進むにつれ語彙フレーズの使用率が高くなれば仮説 a は棄却されない。つまり帰国後時間が経って語彙の検索に時間を要するようになり流暢性が低下するに伴って、語彙フレーズを使用するようになるといえる。この結果が出た場合は、総単語数や使用語彙バラエティー (TTR) と語彙フレーズ使用率との相関も測定し、語彙フレーズの使用が具体的には何に貢献しているのかなどを考えても有意義であろう。逆に、高くならなければ仮説は棄却される。つまり流暢性が低下しても語彙フレーズを使用するわけではない、ということになる。仮説 b についても a と同じアプローチで結果の解釈が可能であろう。セッションが進むにつれ hedges の使用率が高くなれば仮説 b は棄却されない。つまり帰国後時間が経って語彙の検索に時間を要するようになり流暢性が低下するに伴って、hedges を使用するようになるといえる。逆に、高くならなければ仮説は棄却される。つまり流暢性が低下しても hedges を使用するわけではない。

4. 課題、問題点と今後の方向性

今後の課題、問題点は、まず hedges と流暢性の関係を扱った先行研究を調査すること、「流暢性」の測定項目をさらに検討すること、学校教育の影響、被験者の年齢要因などが挙げられる。

この研究により語彙チャンクや hedges が流暢性の維持に貢献することが判明すれば、流暢性を高めるために語彙チャンクや hedges の教授が効果的であるといえることができる。語彙チャンクを中心とするアプローチ法 (Willis, 1990; Lewis, 1993) の教育的効果は未知数であるが、よりよい授業教材・授業案開発のためには、学習者コーパス研究が大きな役割を果たす点については疑う余地はないであろう。TUFS 言語モジュールの教材開発作業のための基礎研究としての JLE コーパス作成およびコーパスデータ分析の今後に大いに

⁶ <http://childes.psy.cmu.edu/> より無料でダウンロードが可能。

期待したい。

参考文献

- Ellis, N. C. (2001). 'Memory for language'. In P. Robinson (eds.), *Cognition and Second Language Instruction*, pp.33-68. Cambridge: Cambridge University Press.
- Lewis, M. (1993). *The Lexical Approach: The State of ELT and a Way Forward*. Hove, England: Language Teaching Publications.
- Miller, G. A. (1956). The magical number, seven plus minus two: Some limits on our capacity for processing information. *Psychological Review* 63:81-97.
- Moon, R. (1997). Vocabulary connections: multi-word items in English. In N Schmitt and M. McCarthy (eds.), *Vocabulary: Description, Acquisition and Pedagogy*, pp.40-63. Cambridge: Cambridge University Press.
- Nattinger, J. R. and DeCarrico, J. S. (1992). *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.
- Pawley, A. and Syder, F. (1983). Two puzzles for linguistic theory; nativelike selection and nativelike fluency. In Richards, J. and R, Schmidt (eds.), *Language and Communication*, pp.191-226. London and New York: Longman.
- Poos, D., & Simpson, R. C. (2002). Cross-disciplinary comparisons of hedging: Some findings from the Michigan Corpus Academic Spoken English. In Reppen, R., Fitzmaurice, S. M., and Biber, D., *Using Corpora to Explore Linguistic Variation*. pp.3-23. Amsterdam: John Benjamins.
- Sinclair, J. (1991). *Corpus, concordance collocation*. Oxford: Oxford University Press.
- Tomiya, M. (1999). The first stage of second language attrition: a case study of a Japanese returnee. In L. Hansen (Ed.), *Second language attrition in Japanese contexts* (pp.59-79). New York: Oxford University Press.
- Tomiya, M. (2000). Child second language attrition: a longitudinal case study. *Applied Linguistics*, 21 (3), 304-332.
- Willis, D. (1990). *The Lexical Syllabus: A new approach to language teaching*. London: Collins ELT.
- Yoshitomi, A. (1992). Towards a model of language attrition: neurobiological and psychological contributions. *Issues in Applied Linguistics*, 3 (2), 293-318.
- Yoshitomi, A. (1999). On the loss of English as a second language by Japanese returnee children. In L. Hansen (Ed.), *Second language attrition in Japanese contexts* (pp.80-111). New York: Oxford University Press.
- 植田恵, 西川恵, 丸井ふみ子 (2004). 「日本人英語学習者の学習者言語コーパス基礎調査」『言語情報学研究報告 5』(21世紀COE研究報告集)