

時間制限を用いた文法性判断課題 —基礎的検討と時間制限の設定方法について—

草薙 邦広
名古屋大学大学院

概要

本稿の目的は、外国語教育研究、及び第二言語習得研究にてしばしば用いられる、時間制限を設けた文法性判断課題に関して、その手法としての理論背景、手法の特徴、及び諸問題を概観した上で、より望ましい時間制限の設定方法についての検討を行うことである。特に、文法性判断課題における認知的プロセスのモデル、そしてそのモデルから仮定される文法性判断課題における反応時間の分布についての知見から、二つの時間制限設定方法（並行集団における事前実験の反応時間に基づく設定法、及び対象集団における実験後の反応時間に基づく設定法）を提案する。

Keywords: 文法性判断課題, 時間制限, 明示的・暗示的知識, 反応時間, 閾値処理

1. 理論的背景

第二言語における文法の運用能力、またはパフォーマンスは、典型的には母語話者との比較に顕著なように、課題や状況によって一貫性がなく、非均質的、または状況依存的であると考えられている。そのため、外国語教育研究、及び第二言語習得研究の分野では、学習者の文法運用（のぼらつき）に影響を及ぼすタスクの要因についても実証的研究が行われている。表1にはそのような要因と研究例をいくつか挙げる。

表 1

学習者における文法の運用に影響を及ぼすタスク要因の例

要因	研究例の一部
時間	Bialystok (1979), Ellis (2004), Kusanagi & Yamashita (2013)
注意 (形式 vs 意味)	Jiang (2007), 草薙・梁・坂東・福田・杉浦 (2012)
干渉負荷	Carrol & Ruigendijk (2012), 福田・草薙 (2012)
運用のモード (受容 vs 産出)	Grüter, Lew-Williams, & Fernald (2012), Paradis (2010), Schmidt (1980), Urano (2012)

例として、時間という要因について見てみよう。当然ではあるが、時間制限を設けることによって、学習者のパフォーマンスは低下する。しかしながら、すべての文法項目におけるパフォーマンスが一様に低下するわけではなく、時間というタスク要因は、対象とする文法項目の特性や、学習者自身が持つ要因（熟達度など）などとの交互作用を持つことが知られている（e.g., Kusanagi & Yamashita, 2013）。

また、実験参加者が、課題中において集中的に形式へ注意を向けるであろう、文法性判断課題（grammaticality judgment task, GJT）では、非常に高いパフォーマンスを示す学習者の場合であっても、意味理解を目的とするような読解課題では、文法形式への逸脱に鈍感さを示す文法項目があることも報告されている（e.g., Jiang, 2007; 草薙・梁・坂東・福田・杉浦, 2012）。しかしながら、母語話者の場合、時間制限や干渉負荷の影響が、学習者ほど明確に観察されることはあまり無いのである。

このような学習者における文法パフォーマンスの非均質性を説明するために、二種類の異質な言語知識の存在を想定する研究者もいる。用語や、定義に関わる概念などは、研究者や時代により実に多様ではあるが、近年では、以下のような二分法の枠組みが有名であろう。

- (1) 明示的ないし顕在的 vs. 暗示的ないし潜在的知識
- (2) 宣言的 vs. 手続き知識
- (3) 非統合的 vs. 統合的知識

例えば、最も包括的な概念のひとつと考えられる明示的・暗示的知識の枠組みにおいては、第一に、異種の知識と実際の言語運用におけるプロセス（言語処理）の対応関係を想定する。明示的知識は、典型的には、メタ言語（長期記憶における宣言的表象）を伴う分析的・統制的な処理を可能にすると考えられており、暗示的知識は、母語話者の言語運用に特徴的なように、流暢で淀みない、自動的、または安定的な処理を齎すと考えられている（概観としては、Bialystok, 1979; Ellis, 2005; Jiang, 2007; Paradis, 1994）。

このような異種の言語処理過程が、言語運用を誘出する特定のタスクの条件（時間・注意の方向・干渉負荷、または産出・受容の差など）の影響を受けるがために、学習者における文法パフォーマンスは、一貫性がなく、非均質性である、と捉えるのである。即ち、学習者の文法パフォーマンスを全体的に見た場合、これら異種の知識とそれに対応するプロセスは、混成的（heterogeneous）であるとも考えることができよう。

しかしながら、その異種の知識における関係性自体については、第二言語習得における主要な議論のひとつであるものの、知識の獲得過程における共通性、知識自体の同一性または重複性、二つの知識の変化や転移（または統合）の可能性などについて、研究者間で一致した見解があるとは言い難い。¹ここまでの議論の概略を図1にまとめる。

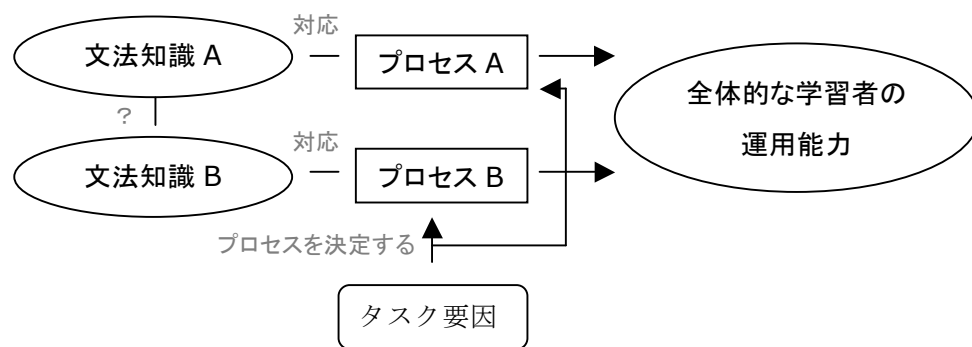


図 1. 文法知識，プロセス，及びタスク要因の関係性

上記のような前提に立つ場合，異種の文法知識の様態をそれぞれ調査するためには，異なるプロセスを誘発するタスク要因の対立を観察し，タスク間のパフォーマンスの差を比較することがなる。つまり，プロセス A を促すであろうタスクの成績は文法知識 A の様態を表し，プロセス B を促すであろうタスクの成績は文法知識 B の様態を表す，と期待するのである。

2000 年代中頃からの Ellis らによる一連の研究（書籍化され，Ellis et al, 2009 としてプロジェクト全体を概観できる）では，複数のテストにおけるタスク要因の共通性に基づいて，二種類のテストバッテリー（明示的知識を測定するテストバッテリー，及び暗示的知識を測定するテストバッテリー）を作成し，その妥当性を検証している。近年はこのテストバッテリーを用いた研究も増えている。

Ellis らが開発したテストバッテリーの中でも，時間制限を用いた文法性判断課題（timed grammaticality judgment task, TGJT）については，暗示的知識を測定する手法のひとつとして取り上げられることが，近年特に頻繁に見られるようになった。しかしながら，手法自体の検討，特に，手法自体の最も重要な点であるともいえる，時間制限の設定方法についての検討が十分になされているとは言い難い。そこで，本稿では，手法として GJT の概観，及び GJT における時間制限の設定方法についての検討を試みる。第二節は，手法の概観と問題点の考察に，第三節は，時間制限の設定方法の提案について，それぞれ割り当てられている。

2. 時間制限を設ける文法性判断課題の概観

2.1 時間制限と文法性判断の認知的プロセス

生成文法の枠組みに限ったことではないが，言語研究一般で用いられる文法性判断課題は，話者の文法性判断の結果を観察することによって，特定の個別言語における規則，及び体系を記述するために行われるものである。しかしながら，第一言語獲得，第二言語習得，そして外国語教育研究の分野では，運用能力の発達過程，ないし特定の文法項目の

獲得・習得困難性を記述するために行われる。言語研究一般で用いられるものとは異なり、第二言語に関わる研究では、GJT における正答率 (accuracy) を文法知識、ないし文法運用能力を反映するものとして解釈する。これまで研究史上、GJT は実に頻繁に用いられてきたといえようが、少なからず手法自体の問題性に関わる議論がなされてきた (以下の様な文献が参考になる : Chaudron, 1983; Davis & Kaplan, 1998; Ellis, 1991; Hedgcock, 1993; Mandell, 1999; Tremblay, 2005, 中でも Chaudron, 1983 は白眉である)。

近年の第二言語習得研究で用いられる GJT は、正答率のみならず、文法性判断における反応時間をひとつの変数として研究に取り入れる場合がある (e.g., Jiang & Nekrasova, 2007)。しかしながら、基本的には、処理のスピードよりも文法知識 (表象, 制約) の調査を取り扱うものであるため、正答率のみに基づく議論が一般的であろう。GJT は、課題自体の特徴として、被験者が集中的に形式へ注意を向けることが不可避であるため、メタ言語的知識の有無を測定するものであり、実際の言語運用に関わるような言語知識は測定できない、という批判も長い間なされてきた。代表的なものは、実際の言語産出においては、メタ言語的知識は言語運用をモニターする役割しか持たぬのであり、言語運用とは無関係である、という主張であろう (Krashen の議論など)。このような批判もあって、実際の言語運用を可能にすると考えられる、暗示的知識と呼ばれるような種類の知識を測定する必要性が高まってきているといえよう。

Bialystok (1979) は、そもそも形態素習得順序の研究や、中間言語可変性 (interlanguage variability) に関する研究などを背景としたものであるが、課題における制限時間という要因に明示的な焦点を当てた先駆的研究のひとつである。Bialystok は、GJT において、形式に焦点を当てた統制的・分析的処理は、自動的・全体的処理よりも、相対的により時間を要するという前提に立ち、課題に時間制限を設けることにより、前者の処理を妨げられるのではないかと考察した。即ち、GJT に対して時間制限を設けることで、暗示的知識の測定を試みたのである。この研究が、TGJT を用いる研究手法の原型のひとつであろう。

更に、Ellis (2004, pp. 256-257) は Bialystok の上記の知見を精緻化し、実験参加者が取る認知的プロセスに焦点を当て、文法性の判断を行う順次作動的なプロセスのモデルを提案している。まずは、そのモデルの概要を図示する (図 2)。

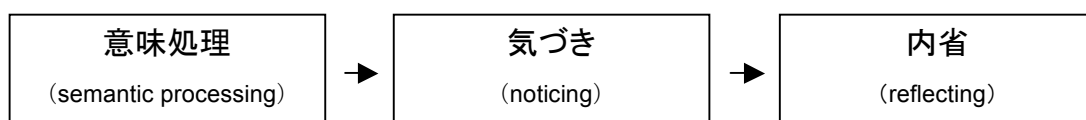


図 2. Ellis による文法性判断課題のプロセスモデル
(図は Ellis, 2004 を参照に著者が作成)

このモデルにおいては、被験者は刺激文が与えられた際、最初に意味処理（Semantic Processing）をするとされている。このプロセスでは、文の意味自体に注意が置かれ、意味を理解することを目標に処理されている。通常、このプロセスは自然な言語運用と同様であり、暗示的知識の使用を伴う、自動的で高速な言語処理と考えられている。

このプロセスにおいて、非文法的、または逸脱である要素があった場合、解釈が不可能になったり、または、解釈に著しく認知的負荷がかかり、処理に支障が発生すると予測できよう。その際に、注意を意味から形式へ向けるプロセスが気づき（Noticing）である。このプロセスは、意味理解を停止し、次のプロセスである内省（Reflecting）に橋渡しをする役割を持っていると考えられる。このプロセスが作動した段階で、文中に非文法的または不自然な形式が存在すると知覚することになる。

最後に行われるプロセスである内省は、文中のどの構成要素が不適切であるかを探索し、どのような原因（制約の違反など）によって、その文が誤りであるか理解しようと試みる。当然ながら、この内省のプロセスは、意識的・統制的な処理であり、リソースとしては、長期記憶に貯蔵され、言語的報告が可能なメタ言語的知識を用いる、と考えられている。

時間制限は、これらプロセスの内、内省のプロセスの作動を妨げると考えられよう。まず第一に、最後に作動する内省のプロセスは、文法性の判断というタスク自体に必ずしも必要なプロセスではない。更に、内省は、相対的に時間をより大きく消費するプロセスであるため、時間制限を設けた場合には、内省のプロセスを行うことが出来ず、気づきのプロセスのみによって課題に反応せざるを得なくなる、と考えられるからである。即ち、TGJT は暗示的知識のみをリソースとして使用する、気づきまでのプロセスによるパフォーマンスであり、untimed GJT は、明示的知識もリソースとして使用する、内省のプロセスまでによるパフォーマンスであると捉えられよう。このような関係を整理すると図3のようになろう。

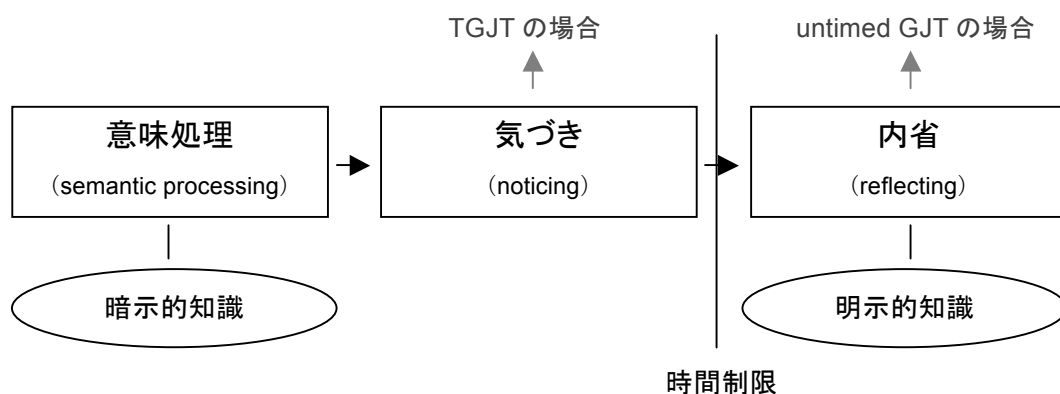


図3. 文法性判断課題のプロセスと時間制限の関係
(Ellis, 2004 を参照に著者が作成)

2.2 実験実施に当たって

ここからは、具体的な手法の特徴を概観するが、その前に、類似した用語の整理を行いたい。以下のような用語自体については、研究者間で一致した見解は無いと考えられるため、その点については特に留意されたい。しかしながら、このような種類の区別は重要であることは言うまでもなからう。本稿では、便宜上、大きく分けて GJT の種類には、untimed GJT, speeded GJT と timed GJT (これまで紹介した TGJT) の三種類があると考えられる。以下にこれら三つの用語の差異を記そう。

- Unitimed GJT (通常の GJT) : 時間に関わる制約が無く、実験参加者は、任意の時間をかけて課題を受けることができる。
- Speeded GJT : 実験参加者は判断を急ぐように設定されている。一般的には、「出来るだけはやく」という文言で指示が与えられ、一文について素早く判断をするようになっている (Jiang, 2012 などを参照のこと)。
- Timed GJT (TGJT) : 時間制限が設定されており、時間制限内に判断できない場合は、強制的に次の文の施行へ移行する。また、得点を 0/1 として扱うとするなら、無回答 (時間制限内に判断が無い場合) は、その施行について得点をしなかったものとして成績に計上される (即ち、誤答と同様に扱う。Ellis, 2005; Kusanagi & Yamashita, 2013 を参照のこと)。

まずは、speeded GJT と TGJT の違いに注目されたい。定義の包含関係から、TGJT は、speeded GJT の一種類であるとも考えられよう。即ち、TGJT は speeded GJT の同様に、判断を急ぐように指示を与えられており、加えて、時間制限内に判断が無かった場合の条件が与えられているのである。Speeded GJT は、通常の GJT である untimed GJT とは異なる判断プロセスを促すものであるが、TGJT は、それに加え、更にプロセスの「結果」に予めフィルターをかけるものと捉えられよう。

TGJT は、その性質の制約上、コンピュータ上で行われる。TGJT をコンピュータ上で行う専用の実験プログラムなどは、広く流通してこそのないものの、市販の心理学実験ソフトウェアなどでも作成可能であろう。また、Kusanagi and Yamashita (2013) が Hot Soup Processer を用いて独自のプログラムを開発しているように、プログラミング言語を用いれば、技術的に開発がさほど困難であるわけではない。コンピュータ環境で行われる TGJT は、図 4 の様な流れで行われる (あくまでも一施行の一例である)。

このような施行を繰り返し、課題の成績を計的な分析にかけることが一般的であろう。方法によっては、正答率と同様に反応時間のデータも得ることができようが、ほとんどの研究目的の場合、反応時間のデータを用いる機会は多くないと考えられる。

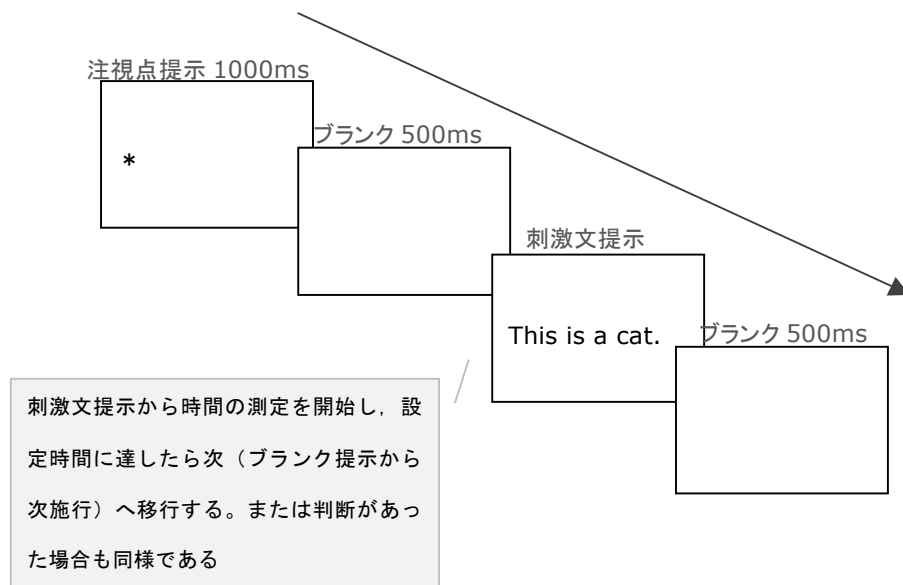


図 4. TGJT の流れにおける例

しかしながら、課題における成績の処理に関しては、後の分析の必要性を考慮し、「無反応」と「時間内の誤答」を分けて分析できるような形式で処理を行うことが望ましいだろう。また、仮に、ある実験参加者の成績において、時間内の誤答が無く、全て無反応であれば、時間制限自体を見なおす必要があるだろう。

国内において、プレゼンテーション用ソフトウェア（Microsoft 社の PowerPoint など）とマークシート等を併用した研究が何件か報告されている。このようなソフトウェアでは、スライドの投影時間を任意の時間に設定することが可能であり、実験文をそのスライドに提示することができる。その後、任意の時間によって実験文の提示が消え、実験参加者は、そのスライドに表示された文についての判断を行い、マークシートに回答を記入する、という手法である。

この実施方法は、「提示時間」に制限を与えることはできるが、「回答時間」を制御するものではないため、もしも提示された一文を記憶することが出来れば、実験文の提示が消えた後も、実験参加者は、自分が望む任意の時間をかけて回答することが可能であると考えられる。また、そのような受験方略を実験参加者が使用する場合、測定の対象とする文法知識よりも、音韻的側面における文字列の記憶保持力や、読解速度自体の影響が大きくなる可能性もあろう。厳密には、このような実施方法は、TGJT や speeded GJT とは操作化の仕方が異なっている、ということに注意する必要があるだろう。

2.3 手法上の諸問題

これまでは、一般的な TGJT の紹介であった。ここからは、実験実施方法に関わる手

法上の具体的問題点をいくつか述べる。

2.3.1 文法文と非文法文の取り扱い

まず、第一に、文法文と非文法文の違いについてである。GJT では、文法文を容認することによって得られる成績、及び非文法文を検知することによって得られる成績の二通りの成績が存在する。これまで、時間制限を用いない通常の GJT についても、文法文と非文法文の扱いは、理論的な観点からも絶え間なく議論がなされてきた (e.g., Birdsong, 1994; Hedgcock, 1993)。これまでの研究では、GJT における文法文 (容認することによる) と非文法文 (検知することによる) の成績、及び反応時間は一致しないと報告している場合が多く (e.g., Loewen, 2009), その差を説明する議論に寄与する確固たる知見は未だ見られない。

TGJT の場合であっても、Ellis らの一連の研究では、文法文の成績は、他の暗示的知識を測定するテストとの相関関係が弱く、テストバッテリーの妥当化のためのモデルに適合しない、と報告されている (この点の指摘については Isemonger, 2007 も参照のこと)。島田 (2010) は、この問題に取り組んだ先駆的研究であり、時間制限の有無と文法性に相互作用がある可能性について指摘している。Urano (2012) や Kusanagi and Yamashita (2013) も同様に、文法文と非文法文のスコアの解釈の取り扱いについて、慎重な態度を示している。

先に紹介した TGJT における認知的なプロセスの見地に立つ場合、Loewen (2009) や Kusanagi and Yamashita (2013) が示唆するように、「そもそも非文法文は、タスク上の理論的背景と適さない」として、非文法のスコアを分析に含めない、という取り扱いもひとつの考え方であろう。今後、このような問題点に対して、更なる議論が必要となるであろうが、現状では、文法文の成績と非文法文の成績の取り扱いに対して、敏感であるべきである、とまとめられよう。

2.3.2 成績の比較について

第二の問題点は、untimed GJT と TGJT の成績の比較に関することである。TGJT を用いる研究では、通常、untimed GJT と TGJT の成績の比較を行う。両課題の性格からして、untimed GJT の方が高い成績を示すことが当然のように期待されるわけであり、単純な成績の比較のみで得られる知見は限られている。時間制限を任意に強く (短く) 設定する方が、当然、成績の下がり幅が大きくなり、両課題の成績の差が統計的に有意になりやすいため、とも考えられる。寧ろ、研究背景の節で述べたように、時間制限というタスクの要因と、学習者要因、及び刺激に用いる文の特性との交互作用こそが調査されるべきであろう。

例えば、刺激 A と刺激 B を用いて、untimed GJT と TGJT の成績を比較するとする。

被験者内二元配置の分散分析を行うとした場合、従属変数は課題の成績、独立変数は刺激の種類（刺激 A 及び B）と課題の条件（時間制限の有無）となろう。もしも、交互作用が無い場合（図 5）には、刺激 A であっても刺激 B であっても、時間制限条件の成績が低い、というような情報しか得られない。これでは、文法習得の研究が主な関心を寄せる、ある言語項目における習得困難性の記述、またはある学習者群における文法発達段階の記述などといった研究目的に適したデータとはなり難い。

一方、刺激の種類と課題の条件に、交互作用が見られる場合（図 6）、刺激 A の成績は時間制限の影響を受けずに、刺激 B のみ時間制限の影響を受けていることになる。このことから、刺激 B は、明示的知識（またはメタ言語的知識）としては習得が容易であるが、暗示的知識として習得（または知識が統合、またはそれに伴う処理が自動化）するには、刺激特有の困難性を持つと予測することができよう。このような調査結果から、刺激 A と B に見られる言語的要因（構造や規則性、典型性、頻度など、または母語との対称分析といった言語間の要因）の差異などを分析し、習得のメカニズムの解明を試みることになる。同様に、学習者要因についても課題の条件との交互作用を調査するべきであろう。

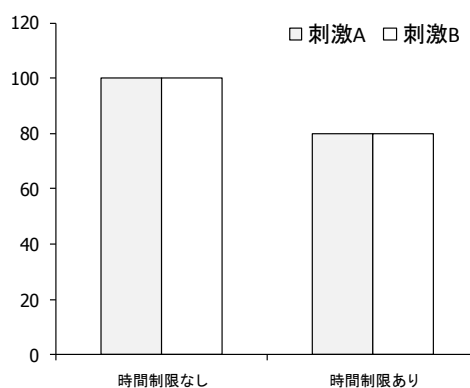


図 5. 交互作用が無い結果（仮想データ）

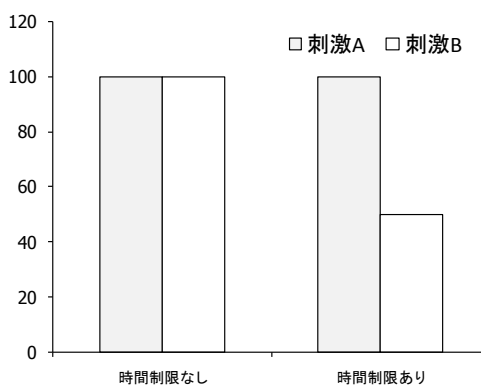


図 6. 交互作用のある結果（仮想データ）

しかしながら、untimed GJT を明示的知識、TGJT を暗示的知識の成績、として安易に解釈することには注意が必要である。untimed GJT の成績には、通常、暗示的知識の貢献と考えられている、即時的な判断による成績も含まれる場合があるからである。

この点については、Suzuki, Kubota, Itagaki, and Takeuchi (2006) が指摘しているように、明示的知識の貢献による言語運用は、暗示的知識による貢献と考えられるパフォーマンスを取り除いたもの、と解釈して然るべきであろう。具体的には、同じ刺激に関する限り、本来、明示的知識の貢献と考えられる成績は、TGJT と untimed GJT の成績の差分である、と考えるのである。この考え方の方が、実験結果の記述としては、より妥当であろう。しかしながら、分散分析に代表されるような統計的な分析を用い、時間制限による成績の変動や、刺激の種類などによる変動を議論の軸とする場合、両課題における成績の差分を取る処遇の是非については問わずともよいだろう。

2.3.3 他課題との関係性

三つ目の問題は、暗示的知識を測定すると考えられる他の課題との関係性である。心理言語学的な実験手法において、俗に「オンライン課題」(コンピュータネットワーク上の、という意味でないことに留意されたい) と呼ばれるものは、全て、暗示的知識に基づくパフォーマンスを測定している、というような見解がなされることがあるが、この点については注意が必要である。

例として、文法習得の分野においては、視線計測、または自己ペース読み課題を用いた誤り検知の課題 (anomaly detection tasks) が用いられることが多い。このような課題は、「実験参加者が暗示的 (ないし統合的) 文法知識を持つ場合、誤り (逸脱) に対して自動的で、即時的な反応を起こす」という前提にもとづいている (e.g., Jiang, 2007, 2012)。操作化したレベルにおいては、誤りのある刺激の箇所に、読解中の注視時間の遅れ、注視回数の増加、ないし自己ペース読み課題では、ボタン押しによる反応時間の遅れが生じるか、ということを観察するものである。分析は、基本的に、実験参加者の注視時間の平均値を比較することによって行われる。この課題の特色は、具体的に誤りのある箇所における即時的な反応を記録することができること、更に、意味理解を目的として読解するようにデザインされるため、より自然な言語運用を測定できることである。

その点において、TGJT においては、判断の根拠となった文中の具体的な箇所は不明であり、更に、自然な言語運用とは考えにくく、寧ろ、課題開始時から形式へ注意を向けることが前提となっていることに留意すべきである。つまり、即時性、注意の方向性の違い、言語運用の自然さという観点では、自己ペース読み課題とは根本的に異なっている。

しかしながら、TGJT の手法的な強みは、そのデータ形式であろう。上記の通り、自己ペース読み課題は、実験参加者グループの平均値を分析の単位とするため、個人のデータに基づいた議論はし難い。一方、TGJT の成績は、個人が持つひとつの変数であり、条

件が整う限り、信頼性係数の算出、相関分析、因子分析、重回帰分析、構造方程式モデリングなどの手法を用いることができる。

例えば、読解力の構成技能に関わる研究では、Erçetin and Alptekin (2012) や Zhang (2012) が、読解能力の構成技能として TGJT を用いて、主成分分析、重回帰分析や構造方程式モデリングを行っており、第二言語運用能力を予測する、ひとつの新しい観点を提供している。

また、指導法の効果などを検証する研究では、時系列における課題の成績を比較するデザインも頻繁に使用されている。このような繰り返しのデザインにおいても、TGJT を用いることができよう。

その他にも、プライミングを用いた判断課題、干渉負荷を用いた判断課題、口頭における課題、及び読み上げ課題など様々な課題があるが、このような課題全てが、同様の知見を生むわけではない。時間制限の有無のみならず、口頭・紙面の差、受容・産出の差、注意の方向性、干渉の有無など、様々な操作化による知見を、統合的に検討することが必要であろう。実験の目的に適したデザインを選択することが重要であり、TGJT の特性を有効に利用した研究が、今後更に望まれるところである。

2.3.4 形式に対する注意の問題

第四の問題は、前節と部分的に関連するが、課題中の形式に対する注意の問題である。GJT 一般に共通して、実験参加者が課題中に形式へ注意を向けることは不可避であり、本質的に、実験参加者がメタ認知的な受験方略を用いる可能性は比較的高いといえる。受験方略には、例えば極論であるが、刺激の統制やプログラム上の提示順の法則を見抜き、全く刺激を読まずに回答するような場合も考えられよう。

このような場合は、実験実施方法、特に刺激の統制が十分でない場合が多く、早急に見直し、再実験を行うか、分析からそのような反応の実験参加者の成績を除外するべきであろう。特に同種類の刺激が繰り返し提示されるようなデザインの場合、フィルターの含有率を高く設定し、提示順にもブロックランダムイズを組む、といった必要性があるろう。

例として、Kusanagi and Yamashita (2013) は、ひとりの実験参加者につき、216 文の判断があるデザインであるため、二度、別日に分けて実験を実施している。216 文の内、最終的に分析に用いられたのは、フィルターと文法文を除く 72 文である。また本刺激（四種類）とフィルターの構成比がブロック毎に一定になるようにブロックを組み、ブロック内外でランダムイズを行なっている。このような工夫も重要な点であるだろう。

2.3.5 文中における誤りの位置

五つ目の問題点は、刺激の統制における逸脱の位置である。非文法文においては、文中の誤り（非文法文であると確定して判断しうる語）となる特定の位置がある。この語

の位置が TGJT において統制されてない場合、時間制限の効果のばらつきを生む可能性がある。具体的には、例文(i), 及び(ii)を参照頂きたい。

- (i) * Some of the *man* who like books are professors.
- (ii) * Some of the men who like book *is* professors.

(i), 及び(ii)の例文はともに非文法文であるが、非文法文であると確定して判断できる語の位置が大きく異なる (*man* と *is*)。 (i)の例では、即時的な判断ができる実験参加者の場合、最初から4語目を読んだ時点で判断をすることになるが、(ii)の場合では、被験者はほぼ文の最後まで読まなければならない。時間制限の設定方法にも関連するが、このように、誤りの位置の差が大きくなるような刺激の統制は望ましくないだろう。

3. 時間制限の設定方法について

上記のような諸問題の他に、最も重点的に検討すべきは、時間制限の設定方法自体である。残念ながら、TGJT においては、時間制限の設定方法こそが最も重要なポイントであるにも関わらず、この問題に取り組んだ研究は、ほぼ皆無である。

テストにおける時間制限の設定方法は、教育心理学及び心理統計の分野では 1950 年第から近年に至るまで、洗練された知見の蓄積があるものの (e.g., Cronbach & Warrington, 1951; Evans & Reilly, 1972; Oshima, 1994; van der Linden, 2011), 第二言語習得研究, 及び外国語教育研究では、あまりこのような研究例は見られない。そもそも、上記のような研究の知見が、そのまま TGJT に援用可能であるとは限らないだろう。

3.1 これまでの設定方法の概観

第二言語習得研究や外国語教育研究に関する限り、TGJT における時間制限の設定方法は、これまで大きく分けて三つの種類が提案されている。ひとつ目は、予め、全刺激に同じ時間を設定する方法である (便宜的に一定時間方式と呼ぶ)。一定時間方式を採用した研究としては、Bialystok (1979), Han and Ellis (1998) などが挙げられよう。この設定方法は、気づきと内省のプロセスには時間的な差があるため、内省による判断をさせない時間の閾値を設定しようと試みたものである。しかしながら、この手法は Ellis (2004) が指摘するように、実験参加者にとっては、短すぎる、または長すぎるといった状況があり得る可能性が強い。また、刺激文自体の長さや複雑さによって、意味処理、及び気づきにかかる時間に、ばらつきが生じる可能性もある (p. 257)。

二つ目の設定方法は、島田 (2010) の研究によるものである。この設定方法を用いた研究に、島田 (2010) の他、浦野 (2011), Urano (2012) などがある。島田は、上記の一定時間方式の問題点を考慮し、刺激の項目の特性を時間制限に反映する方法を提案して

いる。具体的には、刺激文の文長と、構造的複雑さの両方を加味し、刺激自体の観測的特性を、算定式に当てはめ時間制限を設定するものである（便宜的に、算定式方式と呼ぶ）。島田が提案した算定式は、以下のとおりである。

$$\text{時間制限} = 0.6 \times \left(\text{総単語数} - \left(\frac{\text{冠詞数} + \text{指示詞数}}{2} \right) \right) + \text{複雑度}$$

（島田, 2010, p. 44）

文の複雑度は、2段階で評価している。疑問文、二重目的語は1とし、複文の場合は2と計上する。この手法の優れた点は、第一に、刺激毎に異なる設定時間を設定するため、一定時間方式の問題点を克服している点である。また、第二に、その設定の簡便さであろう。この設定方法は予備実験などをせずとも、刺激の特性（語数など）の情報のみによって時間制限を設定できる。しかしながら、島田のこの算定式は根拠が明白ではない（「試行錯誤の結果、（中略）考案した」 p. 44）。例えば、実測値における、一文の読解時間や、文法性判断課題の反応時間を、この算定式がどの程度予測できるか、といった実証研究も必要であろう。

三つ目の設定方法は、実測値の反応時間に基づく方法である（反応時間に基づく方法と呼ぶ）。この方法は、Ellis（2004）にて提案されたものであり、この設定方法を用いた研究は数多い（e.g., Bowles, 2011; Ellis, 2004; 2005; 2006; Erçetin & Alptekin, 2012; Gutiérrez, 2012; Loewen, 2009; Mirzaei, Domakani, & Shakerian, 2011; Kusanagi & Yamashita, 2013; Zhang, 2012）。これは、事前に、同刺激を用いた予備実験を、母語話者に対して *speeded* 条件で行い、各刺激文の平均反応時間に、二割増しの時間を時間制限とするものである（例外的に Kusanagi & Yamashita, 2013 は $M + SD$ と設定している）。

この手法は、実測値の反応時間に基づくため、ある程度、刺激自体の特性を反映することができる。しかしながら、「母語話者を基準とするべきである」ということの理論的背景や、「何故二割増しであるのか」といった根拠は、必ずしも明確ではない。もしも仮に、読解速度自体が、文法知識とは完全に独立したものとしても（そして母語話者よりも遅いものとしても）、母語話者における読解速度の二割増しの時間が、殊更望ましいとは考えにくい。また、暗示的知識の作動速度自体が母語話者よりも二割程度遅い、というような観点は奇妙である（自然な言語運用というのであれば、暗示的知識自体は母語話者と同等と考えるべきであろう）。更に、基準とする母語話者のデータを取ることも自体が、国内では困難である場合も考えられるため、一定時間方式や算出式方式と比べて簡便であるとは言い難い。

また、母語話者のある項目における反応時間の平均値が 3.00 秒である場合に、二割

増しの条件で時間制限を設定すると、時間制限は、3.60 秒になり、0.60 秒の増加になるが、実質的に 0.60 秒といった微量の増加に意義があるとは考えにくい。一方、10.00 秒の場合は、12.00 秒となり、2.00 秒増加することになる。即ち、母語話者の反応時間自体が長い刺激の場合は、増加量も著しく大きくなる。仮に、刺激の文字数などが総じて短く、母語話者の反応時間も短くなると予測されるような実験の時間制限は、母語話者と同様の反応時間となるため、増加分が小さくなりだろう。逆に、刺激文が長くなるにつれ、母語話者の反応時間より離れていくということにもなる。このように、ただ機械的に二割増しをするような設定方法ではなく、母語話者の反応時間や学習者の反応時間の分布を比較し、研究目的に適した設定を心がけるべきだろう。

3.2 設定方法の提案

ここからは、上記のような設定方法の問題点を克服するために、より望ましい時間制限の設定方法について考察したい。考察を開始する前に、いくつかの検討方針をここにまとめたい。第一に、文法性判断の認知的プロセスに、よりよく適合した根拠を持つ方法が望ましい。第二に、一定時間方式に見られるような、刺激が持つ個別の特性を無視するような設定方法でなく、個別の刺激毎に固有の反応時間を設定することが望ましい。第三に、簡便ながらも、研究者の主観的操作に依存せず、更に研究目的によって、ある程度可変性を持たせられる方法が望ましい。

最初に、第一の観点に関してである。まずは、前述した文法性判断課題のプロセスについて再掲したい。Ellis (2004) によると、文法性判断は、意味処理、気づき、そして内省という三つの順次作動的なプロセスから成り、時間制限をかけることで、気づきまでのプロセス、及び内省までのプロセスという二つのパターンを引き出すことが出来ると考えられている。このことから、少なくとも理論的には、文法性判断課題における反応時間の度数分布は「双峰分布」を成すであろう、と想定すべきである（図 7 に双峰分布を持つ度数分布の例を挙げる）。

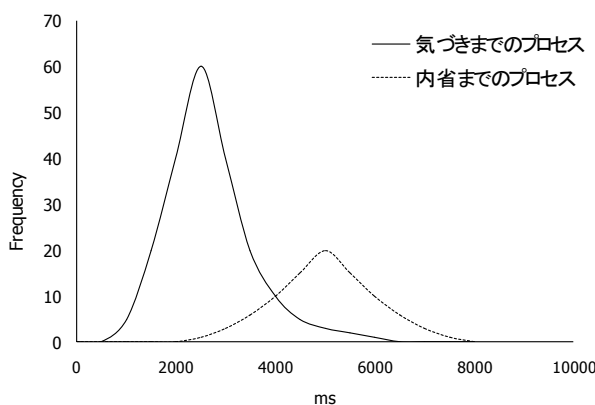


図 7. 双峰分布（仮想データ）

双峰分布は、典型的なケースの場合、複数の母集団に属するものが、同一の標本に紛れ込んだ時に見られる。この場合、まずは、気づきまでのプロセスと、内省までのプロセスが一つに紛れ込んだ分布である、と想定すべきであろう。そして、この双峰分布から、内省によるプロセスの反応を取り除くことにより気づきまでのプロセスによる反応のみを観察することができる、と考える。つまり、反応時間における外れ値の処理と同様に、他方のプロセスによる反応時間を取り除くような「閾値」を決定すればよいと考えるのである（反応時間における外れ値の取り扱いについては、橋本, 2011; Ratcliff, 1993 を参照のこと）。

この方法は、実測値に基づく手法のひとつであるため、時間制限（閾値）を決定するためには予備実験等を行う必要がある。具体的な方法は、データや研究目的に依存するであろうが、以下の二つのようなデータ収集方法が例として考えられるだろう。

まずひとつ目は、事前に、本実験の対象となる実験参加者と同一の母集団に属すると考えられる集団（並行集団）に対して、同一刺激を用いる *untimed GJT*, または *speeded GJT* を実施する方法である。このような予備実験によって、反応時間の分布、特に双峰分布であるかを確認し（双峰分布にフィットさせてみるのが望ましい）、その後閾値を設定する、という方法である。しかしながら、本実験と同等の規模の予備実験を必要とするこの方法は、算出式方式などと比べ比較的簡便であるとは言い難い。

もうひとつの方法は、単一の実験で済ませる方法である。つまり、本実験では、*untimed GJT* または、*speeded GJT* を実施し、事後（*post hoc*）の分析として、反応時間の分布を確認した後に閾値を決定し、閾値を超える反応時間を持つ反応を成績から除外するのである（外れ値の処理と全く同様である）。この手法は、上記の並行集団による方法よりも簡便であり、実際の反応の後に分析をするため、情報の損失がない（実験の段階では成績が決定しないため、時間制限が厳しすぎた、といった事態を招かない）。しかしながら、閾値の設定方法を明示的にしなければ恣意的で主観的な操作を含む分析になりやすいため、十分に注意するべきであろう。

反応時間における具体的な閾値の設定方法は、外れ値の処理と同様に無数にあり、研究目的や、データによって大きく異なり得るだろう。以下に、いくつか、現実的で簡便であると思われる閾値の設定方法を紹介しよう。しかしながら、橋本（2011）が述べるように、都合のいい手法を使い分けるのではなく（p. 143）、先行研究などに基づく理論的根拠と、実際のデータの取り扱い手順を明示化することが肝要であろう。

双峰分布における閾値の設定方法は、画像処理などの分野で用いられる閾値処理（この場合、慣習的に「しきいち」処理と読む）がひとつの参考になる。閾値処理は、カラー画像を単色画像に変換する処理などで用いられる。画像の明度における度数分布表を用いて、その度数分布の状況から閾値を設定するものである。

まずは、最も単純な方法は、実測のデータを、双峰分布にフィットさせ（正規分布

の場合、 $\mu_1, \sigma_1, \mu_2, \sigma_2$ という四つのパラミターで決定される。ex-Gaussian 分布の場合²、それに指数分布の τ が加えられる)、二つ目の峰の分布における最小値を予測することだろう。図 8 のようなデータの場合、2000ms が内省のプロセスを全く含まないパフォーマンスであると考えられよう(その地点までのデータは 100%の割合で前の分布に属するデータであるといえる)。逆に、図 8 における四角の部分は、両方のプロセスによるパフォーマンスが含まれる可能性がある(混成的であるともいえよう)。

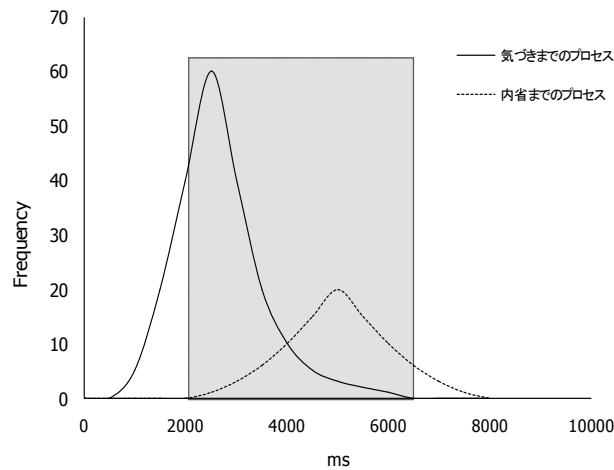


図 8. 閾値処理の例 1 (仮想データ)

図 8 のようなデータの場合、100%他方のプロセスを取り除こうとすると、対象とするデータの大半も失ってしまうことになる。そのような場合、双峰が交り合う部分の構成比を予め任意に決めて置き、閾値を決定するとよいだろう(極端に単純化すると、図 9 のようになる)。例えば、50%までであれば内省のプロセスを認める閾値、といった具合である。

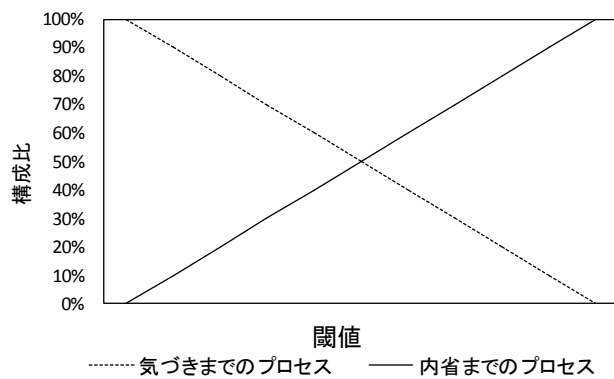


図 9. 閾値処理の例 2 (仮想データ)

また、画像処理の分野でよく用いられように、双峰の谷間になる点を閾値としてもよいだろう（図 10）。この方法は **mode** 法と呼ばれ広く知られている。谷間の位置は、データが双峰分布にフィットしている場合、数的に導き出すこともできる（計算式は本稿では割愛する）。

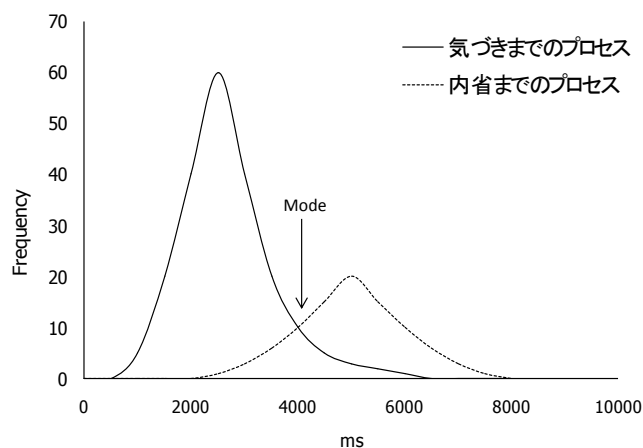


図 10. 閾値処理の例 3（仮想データ）

更に簡便な方法としては、予め、理論的知見から、全体の内、何%を気づきまでのプロセスによる反応が占めるとするかを仮定し、そこを閾値と決定する方法もありうるだろう。この方法は、画像処理の分野では **p-tile** 法と呼ばれる。また、一般的に行われる外れ値の処理も基本的には同種の操作といえよう。例えば、学習者は **speeded GJT** を行う場合、95%は気づきまでのプロセスで判断する（そして残りの 5%は内省を行う）、といった根拠が合理的に得られる場合（あくまでも仮定である）、反応時間の長い方から 5%を除外するのである。

双峰分布における閾値の設定に関しては、この他にも統計的により洗練された方法があるが、本稿の及ぶ範囲ではない。しかしこれらの設定は、より望ましい設定方法の観点として指摘した三つに、以下のような点において適合していると考えられる。

- ・ 「文法性判断という認知的プロセスにより適合した理論的根拠を持つ方法が望ましいこと」 → 提案した設定方法は、文法性判断課題のプロセス、という理論的観点から自然に導き出される現象の予測に適合している
- ・ 「刺激が持つ個別の特性を無視するような設定方法でなく、個別の刺激毎に異なる反応時間を設定すること」 → 並行集団の場合は標本サイズに依存するが、刺激毎に設定時間を決定することは可能であり、更に、母語話者を基準とする方法とは異なり、デザインによっては学習者の要因も反映することができる
- ・ 「第三に、簡便ながらも、研究者の主観的操作に依存せず、更に研究目的によってあ

る程度、可変性を持たせることが出来る方法が望ましい」→基準となる並行集団の実験手続きや、(特に事後の分析における場合に) 閾値の決定方法を明示的に示すことによって、再現性があり、更に、具体的なデータに応じた可変性がある方法である。簡便さについては、並行集団による方法に難点が残るものの、母語話者のデータを基準とする方法よりも、比較的簡便であろう。

4. 本稿のまとめと展望

これまで、本稿では、TGJT に関わる理論的背景、そして手法としての特徴を概観し、手法が持つ問題点について検討した。更に、時間制限の設定方法について取り上げ、これまでの時間制限設定方法の問題点を克服すべく、新しい時間制限設定方法の提案を試みた。その手法とは、並行集団における事前実験の反応時間に基づく設定法と、実験後の反応時間に基づく設定法であった。

しかしながら、本稿の提案は、あくまでも理論的な前提、及び TGJT を用いた研究の文献調査に基づいており、知見の実証的支持を未だ持たない。例えば、GJT の反応時間における双峰分布などといった前提はあくまでも、理論的見地から当然のように予測されるべきである、という考察によるものであり、これまで複数の実証研究に於いて繰り返し指摘されているような経験的事実ではない。今後、このような点を含め、TGJT それ自身の妥当性、信頼性を調査する実証的研究が、殊に望まれるであろう。

特に、時間制限の設定方法については、種々の設定方法の妥当化の研究が最も望まれるところであろう。例えば、他課題などの外的基準と整合性や、信頼性などを調査した研究は未だ皆無である。浦野 (2013) が述べるように、ERP や自己ペース読み課題から既知である実験結果 (e.g., 日本語母語話者における数の一致の逸脱へ対する鈍感さ) との整合性の検証など、今後、このような実証的評価が研究上の課題となってくるであろう。また、視線計測 (e.g., 坂東・梁・草薙・福田・杉浦, 2012; 草薙・梁・坂東・福田・杉浦, 2012) や瞳孔径計測 (e.g., 梁・草薙・坂東・福田・杉浦, 2012) を用いることによって、明示的、及び暗示的知識を測定する新しい試みもしばしば見られる。このようなデータが、TGJT の測定に対して新しい知見を提供することも十分に有り得るだろう。

結語となるが、手法自体へ対する深い考察と、実証的なデータの積み重ねが、今後、この手法の発展に強く寄与することを心から願うばかりである。

謝辞

本稿における内容の多くは、著者が名古屋大学国際開発研究科に対して 2012 年に提出した修士論文「日本語を母語とする英語学習者の明示的・暗示的文法知識—複数形に対する時間制限を操作した文法性判断課題を用いて—」に基いている。まずは、修士論文に対する指導を頂いた、名古屋大学の山下淳子教授に、深く感謝の意を示したい。また、本

稿に関して建設的なコメントを頂いた北海学園大学の浦野研教授，及び外国語教育メディア学会メソドロジー研究部会の関係各位の先生方にも心から謝意を表したい。しかしながら，本稿における不完全な点，または誤りは，全て著者のものである。

注

1. 明示的，及び暗示的知識の関係性についての議論については，本稿の及ぶ範囲ではない。しかし，概論として以下のような文献が参考になるであろう：Bialystok (1979)，Ellis (2004, 2005, 2006)，Ellis et al. (2009)，Gutiérrez (2012)，Han and Ellis (1998)，Jiang (2007)。
2. 一般に，反応時間を分析の対象とする心理言語学実験においては，反応時間のデータは，指数分布と正規分布を足したような性質を持つ ex-Gaussian 分布にフィットすることが知られている (e.g., Lacouture & Cousineau, 2008)。

参考文献

- 坂東貴夫・梁志鋭・草薙邦広・福田純也・杉浦正利 (2012). 「視線計測装置を用いた暗示的文法知識の習得困難性の測定—処理の「即時性」に着目して—」 第52回外国語教育メディア学会全国研究大会. 甲南大学.
- Bialystok, E. (1979). Explicit and implicit judgments of L2 grammaticality. *Language Learning*, 29, 81–103.
- Birdsong, D. (1994). Asymmetrical knowledge of ungrammaticality in SLA theory. *Studies in Second Language Acquisition*, 16, 462–473.
- Bowles, M. (2011). Measuring implicit and explicit linguistic knowledge: What can heritage language learners contribute? *Studies in Second Language Acquisition*, 33, 247–271.
- Carroll, R., & Ruigendijk, E. (2012). The effects of syntactic complexity on processing sentences in noise. *Journal of Psycholinguistic Research*, first view article. 1–21
- Chaudron, C. (1983). Research on metalinguistic judgments: A review of theory, methods, and results. *Language Learning*, 33, 343–377.
- Cronbach, L. J., & Warrington, W. G. (1951). Time limit tests: Estimating their reliability and degree of speeding. *Psychometrika*, 14, 167–188.
- Davis, W., & Kaplan, T. (1998). Native speaker vs. L2 learner grammaticality judgments. *Applied Linguistics*, 19, 183–203.
- Ellis, R. (1991). Grammaticality judgments and second language acquisition. *Studies in Second Language Acquisition*, 13, 161–186.
- Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning*, 54, 227–275.
- Ellis, R. (2005). Measuring implicit and explicit knowledge in second language acquisition: A

- psychometric study. *Studies in Second Language Acquisition*, 27, 141–171.
- Ellis, R. (2006). Modeling learning difficulty and second language proficiency: The differential contributions of implicit and explicit knowledge. *Applied Linguistics*, 27, 431–463.
- Ellis, R., Loewen, S., Elder, C., Erlam, R., & Philip, J. (2009). *Implicit and Explicit Knowledge in Second Language Learning, Testing, and Teaching*. UK, Bristol: Multilingual Matters Ltd
- Erçetin, G. & Alptekin, C. (2012). The explicit/implicit knowledge distinction and working memory: Implications for second-language reading comprehension. *Applied Psycholinguistics*, first view article, 1–27. doi: 10.1017/S0142716411000932
- Evans, F. R., & Reilly, R. R. (1972). A study of test speededness as a source of bias. *Journal of Educational Measurement*, 9, 123–131
- 福田純也・草薙邦広 (2012). 「統語的及び語用論的逸脱への注意における干渉負荷の影響：大文字・小文字混合提示条件の判断課題を用いて」第 80 回外国語教育メディア学会中部支部秋季研究大会. 名古屋大学.
- Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28, 191–215.
- Gutiérrez, X. (2012). Implicit knowledge, explicit knowledge, and achievement in second language Spanish. *The Canadian Journal of Applied Linguistics*, 15, 20–41.
- Han, Y., & Ellis, R. (1998). Implicit knowledge, explicit knowledge and general language proficiency. *Language Teaching Research*, 2, 1–23.
- 橋本健一 (2011). 「反応時間計測実験における外れ値の取り扱い—L2 心理言語実験の場合—」『より良い外国語教育研究のための方法：外国語教育メディア学会（LET）関西支部メソドロジー研究部会 2010 年度報告論集』135–145.
- Hedgcock, J. (1993). Well-formed vs. ill-formed strings in L2 metalinguistic tasks: Specifying features of grammaticality judgments. *Second Language Research*, 9, 1–21.
- Isemonger, I. M. (2007). Operational definitions of explicit and implicit knowledge: Response to R. Ellis (2005) and some recommendations for future research in this area. *Studies in Second Language Acquisition*, 29, 101–118.
- Jiang, N. (2007). Selective integration of L2 knowledge in adult second language learning. *Language Learning*, 57, 1–31.
- Jiang, N. (2012). *Conducting reaction time research in second language studies*. New York: Routledge.
- Jiang, N., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *The Modern Language Journal*, 91, 433–445.
- Kusanagi, K., & Yamashita, J. (2013). Influences of linguistic factors on the acquisition of explicit and implicit knowledge: Focusing on agreement type and morphosyntactic regularity in

- English plural morpheme. *ARELE*, 24, 205–220.
- 草薙邦広・梁志鋭・坂東貴夫・福田純也・杉浦正利 (2012). 「学習者は意味理解を目的にした読解課題中に文法形式の誤りを検出できるか—視線計測を用いて—」 第79回外国語教育メディア学会中部支部春季研究大会. 名古屋学芸大学短期大学部.
- Lacouture, Y., & Cousineau, D. (2008). How to use MATLAB to fit the ex-Gaussian and other probability functions to a distribution of response times. *Tutorials in Quantitative Methods for Psychology*, 4, 35–45.
- 梁志鋭・草薙邦広・坂東貴夫・福田純也・杉浦正利 (2012). 「文法性判断課題における第二言語の文法形式の誤りに対する注意の測定—瞳孔径計測データを用いて—」 第79回外国語教育メディア学会中部支部春季研究大会. 名古屋学芸大学短期大学部.
- Loewen, S. (2009). Grammaticality judgment tests and the measurement of implicit knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam and J. Philip. (2009). *Implicit and Explicit Knowledge in Second Language Learning, Testing, and Teaching*. Bristol, UK: Multilingual Matters Ltd. pp. 66–94.
- Mandell, O. (1999). On the reliability of grammaticality judgment tests in second language acquisition research. *Second Language Research*, 15, 73–99.
- Mirzaei, A., Domakani, M. R., & Shakerian, Z. (2011). Differential accessibility of implicit and explicit grammatical knowledge to EFL learners' language proficiency. *Iranian Journal of Applied Linguistics*, 14, 111–143.
- Oshima, T. C. (1994). The effect of speededness on parameter estimation in item response theory. *Journal of Educational Measurement*, 31, 200–219.
- Paradis, J. (2010). Bilingual children's acquisition of English verb morphology: Effects of language exposure, structure complexity, and task type. *Language Learning*, 60, 651–680.
- Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA (pp. 393–419.). In N. C. Ellis (Ed.), *Implicit and explicit learning of language*. San Diego, CA: Academic Press.
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114, 510–523.
- Schmidt, M. (1980). Coordinate structures and language universals in interlanguage. *Language Learning*, 30, 397–416.
- 島田勝正 (2010). 「文法性判断テストにおける問題文提示時間制限の有無と明示的・暗示的知識」『桃山学院大学 英米評論』24, 41–53.
- Suzuki, W., Kubota, Y., Itagaki, N., & Takeuchi, M. (2006). Explicit and implicit language knowledge on a grammaticality judgment task. *ARELE*, 17, 11–20.
- Tremblay, A. (2005). Theoretical and methodological perspectives on the use of grammaticality

judgment tasks in linguistic theory. *Second Language Studies*, 24, 129–167.

浦野研 (2011). 「普遍文法に基づいた第二言語習得研究における時間制限のある文法性判断タスクの利用」全国英語教育学会第37回山形研究大会. 山形大学.

Urano, K. (2012). Measuring Japanese learners' implicit and explicit knowledge of adverb placement in English. 第38回全国英語教育学会愛知研究大会. 愛知学院大学.

浦野研 (2013). 「第二言語学習者の暗示的知識の測定法—構成概念妥当性の視点から—」『外国語教育メディア学会 (LET) 関西支部 メソドロジー研究部会 2012 年度報告論集』36–45.

van der Linden, W. J. (2011). Setting time limits on tests. *Applied Psychological Measurement*, 35, 183–199.

Zhang, D. (2012). Vocabulary and Grammar knowledge in second language reading comprehension: A structural equation modeling study. *The Modern Language Journal*, 96, 559–575.