

博士論文

外国語の文法知識における一元性の検証：
文法性判断の正答率・反応時間・主観的変数を対象に

草薙邦広

名古屋大学大学院
国際開発研究科

審査委員会
山下淳子（委員長）
杉浦正利
三輪晃司

研究科教授会合格決定
2018年3月7日

目次

第一章 序章	1
1.1 問題の所在	1
1.2 研究の目的	2
1.3 本論の構成	3
第二章 外国語における文法知識	5
2.1 文法知識がもつ性質の概観	5
2.2 文法知識の二分法に関する構造的仮定	13
2.3 構造的仮定の実証と明示的知識および暗示的知識がもつ特性	17
2.4 文法知識の二分法に関する運用	19
2.5 文法知識の二分法に関する心的状態	21
2.6 文法知識の二分法に関する知識の獲得過程	23
2.7 運用と心的状態の関係	26
2.8 運用と獲得過程の関係	28
2.9 心的状態と獲得過程の関係	29
2.10 第二章のまとめ	30
第三章 外国語における文法知識の数理	32
3.1 外国語教育研究における数理的アプローチ	32
3.2 構造的仮定と因子分析モデルの適用に対する批判	35
3.3 明示的知識および暗示的知識と観測の数理的構造	45
3.4 課題の困難度および速さと正確さの二律背反	49
3.5 外国語の文法知識に関する一元観	52
3.6 第三章のまとめ	57
第四章 本実験	59
4.1 数理モデルと研究仮説	59
4.2 実験参加者	64
4.3 実験具	65
4.4 刺激	66
4.5 刺激の測定論的質	67
4.6 実験手続き	68
4.7 変数とデータ構造	69

4.8	解析モデルと解析環境	69
4.9	外れ値の除外	72
第五章 結果		75
5.1	観測の記述とベイズ推定	75
5.2	研究仮説 1：拡散過程モデル	82
5.3	研究仮説 2：主観的変数に対する判断結果と反応時間の影響	85
5.4	第五章のまとめ	90
第六章 総括		92
6.1	結果の総括と解釈	92
6.2	本論の限界	95
6.3	一元観がもたらす効用と今後の研究	100
6.4	結語	101
参考文献		103
謝辞		112
付録		113

図の一覧

第二章

2.1	文法知識が観測に対して影響を及ぼすことを示すパス図	8
2.2	二重能力パラダイムがもつ構造的仮定	17

第三章

3.1	SAT 現象を示すイメージ	50
3.2	SAT 現象と個人差	52
3.2	証跡と観測の関係	55

第四章

4.1	拡散過程モデルの概要	59
4.2	拡散過程を示す概略図	61
4.3	閾値母数と漂流比率の変化が正答率と反応時間分布に及ぼす影響の模式図	62
4.4	実験プログラムの手続きを示す模式図	65
4.5	実験に使用された刺激文における語数の分布	67
4.6	刺激のセットにおけるテスト情報量曲線	68
4.7	観測における D^2 の度数分布と、 χ^2 分布 ($df=2$) の確率密度曲線	73
4.8	弁別力と判断基準の散布図と相関楕円	73

第五章

5.1	三変数の観測における度数分布	76
5.2	実験参加者レベル（左）と項目レベル（右）における三変数の散布図行列	77
5.3	正答率に関するトレース図と事後分布から得られたサンプルの分布	78
5.4	規則反応率に関するトレース図と事後分布から得られたサンプルの分布	78
5.5	反応時間の母数に関するトレース図と事後分布から得られたサンプルの分布	79
5.6	反応時間における度数分布と推定された母数による理論分布（全施行）	80
5.7	反応時間の度数分布と推定された理論分布（正答と誤答別）	80
5.8	反応時間に関する母数における事後分布のサンプルから得られた分布	81
5.9	対数変換された反応時間の度数分布と推定された理論分布	81
5.10	拡散過程モデルによる理論分布と指数正規合成分布による理論分布の比較	83
5.11	実測値とシミュレーションにおける分位点一分位点プロット	84
5.12	固定効果における母数のトレース図と事後分布から得られたサンプルの分布	87
5.13	変量効果における母数の事後分布から得られたサンプルの分布	89

5.14	モデルによる予測値を示す三次元グラフ	89
5.15	個人レベルにおける平均規則反応率の比較	90

表の一覧

第二章

2.1 R. Ellis による明示的知識および暗示的知識がもつ特性の一覧.....	18
--	----

第四章

4.1 実験に使用された刺激文における語数の分布とその比率	67
4.2 ラッシュ・モデルと二母数モデルのフィット	67
4.3 全実験参加者の弁別力 (d') と判断基準 (c) に関する記述統計	72
4.4 分析対象となる実験参加者の弁別力 (d') と判断基準 (c)	74

第五章

5.1 個人レベルにおける標本の記述統計	75
5.2 項目レベルにおける標本の記述統計	75
5.3 個人レベルと項目レベルにおける三変数の相関係数行列	76
5.4 拡散過程モデルにおける推定母数と対数尤度	82
5.5 拡散過程モデルにおける推定母数の五数要約	83
5.6 実測値とシミュレーションにおける分位点の比較	84
5.7 階層ベイズモデルにおける各母数の事後分布から得られたサンプルの要約	86
5.8 実測値とモデルによる予測の混同行列	89

第一章 序論

1.1 問題の所在

昨今、ある種の外国語教育政策、個人の外国語能力に付随する社会的および経済的価値、外国語学習の費用対効果、また、それらに関連する数多くの言説が、広く社会の衆目を集めるようになってきている。そして、そうであればあるほど、本来は研究上の一概念に過ぎない外国語の文法知識といったものに対してであっても、社会の関心はそれ相応に集まるものである。それと同時に、当該の概念に対して関心をもつ、さまざまな人々の見方や認識は、教育政策を始めとして、教育における意思決定や経済的活動のみならず、個人の学習行動に対してでさえも、なんらかの影響を及ぼす。

たとえば、「外国語の文法知識が二種類に分けられる」という一種の見方は、外国語教育研究や、その関連分野の研究に限らず、その見方がもたらす帰結をも含んで、我々の社会において広く見受けられるものである。文法知識 A と文法知識 B なるものが科学的に実在し、「文法知識 A のみが、我々の社会において高い効用をもたらす、一方の文法知識 B は低い効用しかもたらさない」、またはその上で「文法知識 A を伸長させる方法は、教育方法 A のみであって、教育方法 B は、文法知識 B を伸長させるのみである」といった具合である。このような見方の帰結は、当然、「より高い効用を与える教育方法 A を選択すべきである」といった主張であり、これらのような主張が、実にさまざまなレベルの意思決定に影響すると想像することは難しくない。

効用のみならず、費用、資源、需要といった、ほかの経済学的視点によっても、さまざまな言説のパターンが発現しうる。たとえば、「文法知識 A は、より高い費用を要する学習環境 A のみによって得られる」、そしてその上で「学習環境 A をすべての生徒に提供することは費用の上で困難であるのだから、学習環境 B を与えるのがよい」、または、「文法知識 A を必要とする個人は我が国では少数派であって、大多数の個人は文法知識 B を有するのみでよいのであるから、学習環境 B を与えるとよい」といった言説である。

本論は、上記のような言説の類に批判を加えることを目的としているわけではない。著者は、それらの点に関して、純粹に社会学および経済学的観点によって議論されるべきであると考えているし、そしてそれを期待している。しかし、そうではなく、本論がその対象とすることは、上記のような言説の支えになっているであろう前提、「外国語の文法知識が二種類に分けられる」という見方、まさにそれ自体についてである。

「外国語の文法知識が二種類に分けられる」という見方は、もっとも一般的な場合、第二言語習得研究の用語であるところの明示的知識ないし顕在的知識 (explicit knowledge) と暗示的知識ないし暗黙的知識 (implicit knowledge) によって、一種の説明を与えられている。この二分法とその亜種は、当該の研究分野において支配的であるとまではいえない

いものの、事実上の主流派であるか、または控えめにいったとしても、広く研究者に受け入れられているものである。しかしながら、必ずしもこの二分法について、誰もが納得するような一貫した学術的見解があるわけではない。むしろ、その知見の細部に目を配ると、研究者によって認識が異なる点が多く存在することが、長年に亘って、第二言語習得研究者によって主張されている (e.g., R. Ellis, 2004, 2005; R. Ellis et al., 2009; Hulstijn, 2005)。また、明示的知識および暗示的知識の二分法については、外国語教育に関する多数の学術的概念と同様に、その存在についての経験的根拠が十分なものであったとはいえない。実際に、第二言語習得研究では、これらの二種類の知識に関する標準的な測定方法と、その測定モデルに関してですら、2000年代半ばより議論が絶えない状況であり、その測定の公共性は、事実上ないに等しい (e.g., R. Ellis, 2004, 2005, R. Ellis et al., 2009; Gutiérrez, 2012, 2013; Suzuki, 2017)。つまり、奇妙なことに、私たちは、外国語教育に関する基礎研究を担当する分野において、当該の概念についての標準的な測定方法すら確立していないという状況と、そして当該の概念を無批判的に前提とする言説が、さまざまな場面において俎上に載せられる状況に、同時に居合わせているというわけである。

このように、「外国語の文法知識が二種類に分けられる」という見方自体は、それ自体が純粋に自然科学的探求の発露であったとしても、社会におけるその直接的および間接的な帰結の重さに見合うような、強固な基盤をもつなどと決してみなせるものではない。

1.2 本稿の目的

さて、上記の見方、つまり「外国語の文法知識が二種類に分けられるか」を検証するといったところで、本論は、「外国語の文法知識が二種類に分けられる」という命題の科学的真偽に触れるものではない。そもそも、本論は科学的真偽を一様に求める、というような大仰な姿勢による研究ではない。そうではなく、本論の主たる目的は、外国語の文法知識に関して、教育という所与の目的に対する整合性を最大化するような、新しい外国語の文法知識に関する認識論の方針を示すことである。

具体的に、本論では、広く社会に受け入れられている上記の二分法がもつ含意に反して、(a) 文法運用に関する観測は、外国語の文法知識が二種類ではなく、一種類であるか、または構造的に一元的であるかのように振る舞う、さらに、(b) 構造の一元性を仮定した数理モデルが、観測に対して比較的優れた近似を与える、ということを経験的に報告する。このことから、外国語の文法知識を、構造的に一元的であるとみなす見方が、その帰結を考慮した上で、より妥当であるということを経験的に結論としている。

さらに、外国語の文法知識が一元的であるとする見方、つまり本論で後に提唱するところの一元観は、既知の文法運用に関する観測的事実をより広範囲に説明することができるだけでなく、教育業務や教育に関わる意思決定、より具体的にいえば、教育評価、

カリキュラム開発，教材作成などにおいて高い有用性を示す可能性があるということを主張する。

また，当該分野の主流から外れるような本論の主張は，著者が外国語教育研究における数理的アプローチと呼ぶ，帰結主義的，経験主義的，そして形式主義的な特徴をもつ一種の研究姿勢の上に成り立っている。このアプローチの是非については議論が別れるところであろうが，本論では，当該分野において比較的新しいこの研究アプローチがもつ視野や，その手法について示すことも副次的な目的として見据えている。

1.3 本論の構成

本論の構成は以下の通りである。

第二章では，「外国語における文法知識」と題して，外国語教育研究，そしてその関連分野である第二言語習得研究や認知心理学といった諸分野から，多角的に「外国語の文法知識が二種類に分けられる」という見方を生んだ先行研究について紹介する。特に，外国語の文法知識に関する知見が，文法知識の，(a) 運用，(b) 運用に付随する心理状態，(c) 獲得過程という3つの観点から簡潔に整理できることを示す。

第三章には，「外国語における文法知識の数理」という題を与えている。この章の最初に，外国語教育研究における数理的アプローチについて，その研究姿勢の要点を概説する。その後，第二章で述べた明示的知識および暗示的知識の二分法について，その数理的構造が内包するさまざまな問題点を適宜取り上げながら批判する。さらに，心理測定モデル (psychometric model)，正確さと速さの二律背反 (speed-accuracy tradeoff) といった概念に触れつつ，どのように数理的アプローチを当該の問題に対して適用できるか，より具体的には，文法知識研究において，事実上の手法的スタンダードとなっている文法性判断課題 (Grammaticality Judgment Task; GJT) より発生する観測に対して，どのように優れた近似を与えることができるか，について議論する。その後，明示的知識および暗示的知識の二分法の代替的認識論となる，一元観を提唱し，外国語の文法運用に関する観測の構造を，一元的であると捉える見方について詳しく議論する。

第四章「本実験」は，比較的大規模な文法性判断課題を題材とした実験の詳細についてである。

第五章「結果」では，文法性判断課題に付随する正答率，反応時間，そして主観的変数の構造が，まさに一元観を支持する結果を示すということを，マルコフ連鎖モンテカルロ法 (MCMC) による階層ベイズモデルを援用することによって明らかにする。また，文法性判断課題における正答率，正答の反応時間分布，誤答の反応時間分布といった観測が，拡散過程モデル (diffusion model) といった数理モデルによって適切に近似を与えることが可能であることも示す。拡散過程モデルは，明示的知識および暗示的知識の枠組みが想定するような，文法知識の二元性を仮定しないことにその特色がある。

第六章は全体の総括に当てている。この章では、結果の総括、この結果に関して想定される反論とその反駁、本論の限界点と、一元観がもたらすであろう社会的帰結について論じる。

なお、本論では、全体の論の流れと可読性を重視し、厳密な数理的説明や数式の掲載は、必要とならない限り避けることとしている。また、そのような技術的説明に注力することは、数理的アプローチを標榜こそしているものの、あくまでも外国語教育研究の一部である本論の範囲を超える。このような数理上の、または技術的な諸側面に関する説明については、数理的アプローチと、本論が提唱する一元観に関するさらなる研究に譲ることとする。

第二章 外国語における文法知識

2.1 文法知識がもつ性質の概観

この章では、外国語の文法知識に関する先行研究を、特に第二言語習得研究と認知心理学の観点から、多角的に、そして批判的に概観する¹。

さて、文法知識と一口にいったところで、文法知識という用語を使用する学術分野は、外国語教育研究に限らず、第二言語習得研究、言語テスト研究、そして言語学といったように多岐に渡る。また、文法知識という用語が指示するものが、必ずしも、これらの分野間において同一であるとみなすことはできない。むしろ、従来の外国語教育研究と第二言語習得研究のみに視野を絞ったとしても、「外国語における文法知識とはどのようなものであるか」という抽象的な問いに対する答えは、究極的には研究者に視点によるところである。

しかしながら、この問いに対する、想定されうる数多くの答えの中であって、その最小公倍数を与えられるような性質のいくつかは、以下のようにまとめることができるだろう。また、本論では全編に渡り、用語としての文法知識を、特筆しない限り、以下の性質のすべてを満たすものとして取り扱っている。

- (a) 個人、集団、そして状況に帰属する特性である
- (b) 直接的な観測が不可能か、または困難な潜在変数 (latent variable) である
- (c) 外国語の運用、特に文法的側面に関連する運用の原因になる
- (d) 外国語の技能を構成する要素 (構成技能; component skills) のひとつである
- (e) 言語の構造的および機能的側面に関する表象や処理として捉えられる
- (f) 教育や学習の成果変数とされ、さらにそれが社会において広く容認されている

ここからは、上記それぞれの観点を順に概観していきながら、外国語における文法知識がもつ性質について多角的に検討していく。

(a) 個人、集団、そして状況に帰属する特性である

外国語における文法知識を、少なくとも個人や集団に帰属する特性であるとみなすことは、文法知識という用語を使用する諸分野において、暗黙の前提とされている。

集団ではなく個人に帰属する特性は、個人内、たとえば時系列において、相対的に安

¹ 本論は、第二言語習得研究、認知心理学、心理統計、数理心理学など、さまざまな分野の知見を総合的に援用している。しかしながら、本論に関する研究は、あくまでも学際性を是とする外国語教育研究のディシプリンの上で行われている。本論がもつ具体的な研究姿勢については、第三章を参考にされたい。

定したもので、つまり変動が小さいものであって、同時に、個人間においてある程度の差や分散が観測されるものである。つまり、個人内における変動の大きさよりも、個人間における変動の大きさが、相対的に十分大きい、といったことが観測上期待されなければならない。これは級内相関や分散成分の相対的大小として理解してもよい。

たとえば、よく知られる性格 (personality) という心理学上の概念を例にしよう。その科学的実在性や、概念の導出過程についてはどうであれ、ある性格をもつとされる個人は、その性格で表されるような同種の性質を任意の時点間、安定的に保持している²。さらに、その保持された性格は、他人と弁別可能である、といった前提がある。一方、ある同一の状況に置かれている個人間における変動が比較的小さく、さらに状況間における変動の方が相対的に大きい場合、それは個人に帰属しているのではなく、状況に対して帰属するものとみなされることもある。このことについては、心理学における特性不安 (trait anxiety) と状況不安 (state anxiety) の区別 (Spielberger, 1966) にも、その典型例を見出すことができる。

集団についても同様であり、集団に帰属する特性は、集団内において、ある程度変動が小さいものであって、それと同時に集団間の弁別に対してなんらかの寄与をする、といった要件によって特徴づけられる。たとえば、言語学において、文法知識の帰属先をスピーチコミュニティに対して求める場合がある。生成文法では、広く知られるように、同一のスピーチコミュニティに所属する母語話者間の文法知識の差は、積極的に捨象されている³。

外国語教育研究や第二言語習得研究においても同じように、ある種の限定的な研究目的をもつ場合を除けば、学習目標なり習得対象であるところの文法知識は、母語話者というスピーチコミュニティに帰属しているとみなされ、母語話者間における文法知識の差は、積極的に捨象される傾向がある。また、第二言語習得研究の一部、特に言語習得における母語の影響に関心をもつ研究では、文法知識の様態が、母語を同じくする学習者の集団や、あるデモグラフィック情報によって分類された小集団に帰属するとみなされることがある。現実における用語の使用例に目を向けると、「日本語を母語とする学習者の文法知識」 ("*grammatical knowledge of Japanese EFL learners*") や「中学生の文法知識」

² 社会心理学には、根本的な帰属の誤り (fundamental attribution error)、または対応バイアス (correspondence bias) と呼ばれる概念がある (e.g., Jones & Harris, 1967; Gilbert & Malone, 1995)。これは、ある行動に関して、その原因を状況や環境ではなく、個人の特性に帰属させる過剰な傾向を示す。ある行動に関する原因の帰属先が何であるか、という観点は、既存の理論をメタ的に捉える重要な観点のひとつになりうる。これらの観点は、帰属理論 (attribution theory; e.g., Kelley, 1973; Kelley & Michela, 1980) などとも関連する。本論が使用する帰属および特性という用語は、この帰属理論によるものである。

³ 生成文法の場合、理想化 (idealization) というその特徴的な概念が示すように、スピーチコミュニティ内における個人間の差を認めないわけではなく、それをほとんどの場合、研究の対象外としている。捨象は、このように、研究者による研究実践上の主体的な行為であると本論はみなしている。

("grammatical knowledge of junior high school students") といった表現を、過去の研究における題目の中に見出すことは容易である。

なお、本論では、少なくとも歴史的経緯の上で、文法知識を、個人や集団に帰属する特性であると同時に、さまざまな種類の状況に帰属するものとしてみなす見方の発現が、文法知識の二分法における導出の背景にある重要な点であると考えている。このことについては、次節以降に詳しい。

(b) 直接的な観測が不可能かまたは困難な潜在変数 (latent variable) である

脳機能イメージング技術を駆使するような一部の研究を除けば、文法知識を、自然科学の要領によって直接的に観測することは、現在の科学技術では不可能であるか、または極めて困難であるとみなされている。実際に、一部の分野を除けば、文法知識とされるものが、なんらかの記号、値、映像などといったものを観測可能な形で示すことはない。むしろ、文法知識は、動機づけ強度、不安傾向、自己効力感といった無数の心理学的概念や、学習意欲、読解技能、そして熟達度といった外国語教育研究に関連する数多くの概念と同様に、心理統計が呼ぶところの潜在変数⁴として扱われている。これを構成概念 (construct) と呼ぶこともできる。

外国語の文法知識が構成概念であるという一種の見方が、外国語教育研究において普及したのは、2000年代の第二言語習得研究、特に R. Ellis を始めとした研究グループの活躍 (R. Ellis, 2004, 2005; R. Ellis et al., 2009) によるところが大きい。とりわけ、R. Ellis (2005) は、副題に「心理統計的研究」("A psychometric study") という表現を冠していることからわかるように、文法知識を心理統計の視点から捉えようとした先駆的研究である。この研究の後に、文法知識を構成概念であるとみなし、積極的にその概念的特性や、測定モデル、そして測定の全体的評価であるところの構成概念妥当性 (construct validity; Messick, 1995) を検討する研究が続いた (e.g., Bowles, 2011, R. Ellis & Loewen, 2007; Gutiérrez, 2012, 2013; Spada, Shiu & Tomita, 2015; Suzuki, 2017; Suzuki & DeKeyser, 2015; Vafae, Suzuki, & Kachisnke, 2017)。次節にて紹介する明示的知識および暗示的知識といった理論的枠組みにも関連するが、複数課題からなるテストバッテリーの成績を観測変数として潜在変数モデルを適用した上で、文法知識の値を数理的に推定する試みに、その手法的新規性が見られたことも特筆すべきである。これらの研究によって、文法知識は、少なくとも理論的には、得られた観測によって推定される潜在変数として扱われるようになった⁵。

⁴ 直接的に観測できる変数を観測変数といい、観測変数より推定され、直接的に観測できない変数を潜在変数という。ただし、その区別に吟味が必要であることはいうまでもない。

⁵ 言語テスト研究では、古くから文法知識の測定に関する研究は盛んであった。しかしながら、単一の文法テストの成績に対して、潜在特性モデル (latent trait model) ないし項目反応理論を適用する例はあったが、テストバッテリーの成績に対して潜在変

(c) 外国語の運用, 特に文法的側面に関連する運用の原因となる

文法知識は, 上記のように, 潜在変数として扱われるものであるが, 潜在変数としてその値の推定が解析上可能になるためには, 潜在変数の値こそが複数の観測変数の原因である, という一種の条件を設けなければならない。つまり, 潜在変数であるところの文法知識の値が高ければ, 外国語の運用, 特に文法的側面の運用に関する複数の観測変数の値が高い, といった因果関係を想定しなければならない。この関係を, パス図を使用して簡略的に表すとすると, 図 2.1 のようになる。この図では, 誤差を省略している。

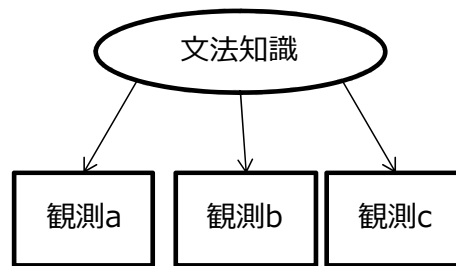


図 2.1 文法知識が観測に対して影響を及ぼすことを示すパス図

たとえば, 集団の中で, ある個人 x がもつ文法知識の値が, 真値として相対的に高いとする。このことは, 個人 x がより豊かな文法知識をもつといった解釈を可能にする。このとき, ある文法的側面の運用に関する観測, つまり, 文法テストの得点, 作文における誤りの数, または文法テストにおける回答時間の短さなどに, 上記の文法知識の値が独自の影響を及ぼしている, ということが上記のモデルの含意である。これは, 「豊かな文法知識をもつ個人の文法テストの得点は高く, 豊かな文法知識をもたない個人の文法テストの得点は低い」といった関係であると考えてもよい。しかし, この関係は連関関係でなく, 文法知識から観測へ向かう関係でなければならない。このような前提は, 反映モデル (reflective model) という考えた方⁶に依拠している (e.g., Anderson & Gerbing, 1988; Edwards & Bagozzi, 2000)。反映モデルでは, 潜在変数および観測変数の両方が連続量である場合, 一般に知られる因子分析モデルを適用することができる。よって, 文法知識は, 因子分析における因子のようなものである, と理解することもできる。詳しく

数モデルを適用する例は見られなかった。また, それ以前の第二言語習得研究において, 文法知識は, もっとも一般的な場合, 古典的テスト理論の下で任意のテストの合計得点として取り扱われていた。それと同時に, 1990 年代までは, 特定の文法項目の文法知識について関心をもつ研究が多かった。

⁶ 反映モデルは, 潜在変数モデル (latent variable model) におけるひとつのモデルである。反映モデルは, 潜在変数を観測変数の原因とみなすことによるモデリングを指し, 形成モデル (formative model) と対立する。形成モデルは, 観測変数を潜在変数の原因とみなすことによるモデリングである。心理学では一般的に, 個人の能力を反映モデルによって扱う場合が多い。一方, 社会学においては, 社会経済的地位といった種の潜在変数について, 形成モデルを使用する場合が多い。

は、第三章を参照されたい。

本論では、このように、文法知識を潜在変数として取り扱い、さらにその性質を、便宜的に連続量であると仮定している。また、文法知識が原因となり、その影響を受けていると考えられる言語行動 (language behavior)⁷の一部を、本論では、運用 (performance) と呼ぶ。

このようなモデリング・コンセプトは、明示的に議論されるかどうかに関わらず、第二言語習得研究や言語テスト研究においても、一般的に見られるものである。読解技能は、一種の言語行動やその結果を表すさまざまな観測変数、たとえば、読解問題の正答率、読解速度、そして視線計測値に対する原因であるとか、ライティング技能が、産出語数、誤りの数、完成までの時間に対する原因である、といった知見に、その類例を数多く見出すことができる。ただし、これらの概念と観測の関係については、現状において、十分に検証されたものばかりであるわけではない。

(d) 外国語の技能を構成する要素 (構成技能; component skills) のひとつである

上記のように、文法知識に限らず、外国語に関する技能の数多くは、多様な側面における外国語運用の原因となる潜在変数として扱われている。一般的に、四技能として集約される、ライティング技能、スピーキング技能、聴解技能、そして読解技能はもちろんのこと、語彙知識や、方略的知識、メタ認知能力、または教育研究において広く受け入れられている社会技能 (social skills) などは、文法知識とともに、総合的な外国語運用の原因となる、またはそれを構成する潜在変数であると考えられている。それだけではなく、読解研究や言語テスト研究に顕著なように、形態的知識、音韻認識能力、推論能力、などといった、より細かな構成概念を研究者が設ける場合もある。

我々が関心をもつべき問題は、これら複数の構造や連関についてである。外国語運用の原因となる潜在変数同士の構造や連関を調べることで、外国語の運用メカニズムを明らかにしようとする一種の方法論を、外国語における構成技能アプローチ (component skills approach; Carr & Levy, 1990, Jeon & Yamashita, 2014; Melby-Lervåg & Lervåg, 2014; Van Gelderen et al., 2004; Yamashita & Shiotsu, 2017) と呼ぶ。構成技能アプローチは、主に読解研究において盛んに応用されており、文法知識が潜在変数として明確にみなされ始めた2000年代半ば以前より、着実に研究を積み重ねていた。構成技能アプローチは、因子分析モデルや構造方程式モデリング (SEM) といった多変量解析の普及に伴って発展した。このアプローチは、心理学における古典的な知能研究 (intelligence research) の方法論に

⁷ 本論は、用語としての「行動」を、一般に、観測可能なものを意味するものとして扱っている。本文で述べているように、「運用」と「行動」は観測が可能であり、「知識」や「認知機構」は観測不可能である。また、直接的に観測されない文法知識は、観測可能な言語行動に影響を及ぼすであろう潜在変数のひとつである。本論は、徹底的行動主義者がそうするように、内省的データも一種の行動とみなしている。ただし、内省的データについても運用とみなすわけではない。

非常によく似通っている⁸。

さて、外国語における構成技能アプローチでは、外国語の運用に影響を及ぼす潜在変数を、構成技能 (component skills; Carr & Levy, 1990) と呼ぶ。構成技能アプローチにおけるモデリングでは、典型的な場合、潜在変数間の因果関係をモデルに取り入れる。そのため、ある構成技能が影響を及ぼす他の構成技能を下位技能 (sub-components) などと呼び、それに対応する上位の技能と区別することがある。それぞれの構成技能は、その名が表すように、熟達度といった単一の統合的な技能に対する下位技能としてモデリングされる場合も多い。文法知識も例外ではなく、熟達度を構成するひとつの構成技能としてみなされる場合がある (R. Ellis, 2006)。

過去の第二言語習得研究では、特に文法知識または語彙知識が読解技能に及ぼす影響の強度について検証する研究が数多くなされた (e.g., Shiotsu & Weir, 2007; Yamashita & Shiotsu, 2017)。読解技能と他の構成技能の関係については、潤沢な研究実績があるため、これらの研究を統合したメタ分析も報告されている (Jeon & Yamashita, 2014)。このように、文法知識は、熟達度ないし言語運用を支える構成技能のひとつであり、他の構成技能との相対的な関係、つまり、その潜在構造⁹について論じられることがある。

また、文法知識の二分法によって、二種の構成概念を想定するとき、これらはともに構成技能としてみなすべきであるし、実際にそのようにみなされてもいる (R. Ellis, 2006)。

(e) 言語の構造的および機能的側面に関する表象や処理として捉えられる

文法とは、一般的にいて、言語の構造的および機能的側面に関する規則や規範であるのだから、文法知識という用語が参照するドメインは、言語の構造的および機能的側面である。しかし、言語の構造的および機能的側面というドメインは、実に曖昧なものであって、実質的に、この用語が参照するドメインは、言語のあらゆる側面に対して開かれている。たとえば、もっとも基本的な言語学的分類に沿うと、(a) 統語論 (syntax), (b) 形態論 (morphology), (c) 音韻論 (phonetics), (d) 意味論 (semantics), (e) 語用論 (pragmatics) といった領域を含む。一般的には、統語論や形態論に関するドメインのみを指して、文法知識と呼ぶ場合が多い。

用語の使用に関する限り、外国語教育研究者や第二言語習得の研究者は、統語的知識であるとか、または語用論的知識といった用語を使用して、それが参照するドメインを積極的に限定するか、そうでなければ文法知識ではなくて、言語的知識といった用語を

⁸ 因子分析モデルは、知能研究の発展と切り離せない関係がある。因子分析モデルが開発された後、スピアマンによる2因子、サーストンによる9因子、近年ではガードナーによる多重知能など、潜在変数間の構造に対するアプローチには、因子分析モデルにおける推定因子数を中心とした理論化という傾向が見られる。このような点は、構成技能アプローチと相同している。

⁹ 潜在変数間の因果関係をここでは潜在構造と呼んでいる。

使用して、より包括的な側面をもつドメインを参照することが多い。ただし、多くの研究者は、言語的知識の区分として、文法知識と語彙知識があるものと認めている。このような区別は、構造的および機能的規則の適用範囲が語以上であるかなどといった言語学的視点からもたらされたものであり、多様性に富んでいる。同時に、語彙主義と呼ばれるような言語観や、語彙主義から派生したさまざまな言語習得観の下では、構造依存的な統語規則のみではなく、語彙目録が他要素に要請する機能的範疇などに着目することもある。

さらに、文法知識を、一種の知識表象 (knowledge representation) である、と考える表象主義 (representationalism) ないし象徴主義 (symbolism)¹⁰も、かなり一般的な枠組みであって、文法知識に関する表現の中で、表象という用語を好んで使う研究者も多い (e.g., Hawkins, 2001; White, 2003)。文法知識に関する表象は、心的なものであるところの記号や象徴、または命題であるかのように扱われる。たとえば、英語の *men* は、複数の男性を意味するので、この単語の知覚にともなって、語であるところの *men* には複数という表象が付与されていると考えられたり、または、「主語の数が単数であり、文の時制が現在時制、さらに主語の人称が三人称であれば、動詞を屈折させる」といった規則や制約自体が、表象であるとみなされたりもする。前者のように、単純な命題 (p) によって構成されるものを宣言的知識または宣言的記憶 (declarative knowledge/memory)、後者のように条件を伴う命題 ("if p , then q ") を、手続き的知識または手続き的記憶 (procedural knowledge/memory) と呼ぶこともある (e.g., Squire, 2004)。また、近年の神経科学や記憶研究では、長期記憶の下位分類として、陳述記憶と非陳述記憶を設ける場合もある (Squire & Zola, 1996)。

本論は上記のように、文法知識に関して、表象主義による捉え方を中心としているが、連合主義 (connectionism; e.g., Fodor & Pylyshyn, 1988; Schneider, 1987) による捉え方を好む研究者もいる (e.g., N. Ellis, 1998, 2005)。連合主義は、表象主義とは異なり、自然言語や記号類によって明示的に表現できるような知識を想定しない。過去の第二言語習得研究において、表象主義は、主として生成文法の生得的言語習得観と、言語習得に関する生物学的基盤、つまり生得的言語習得装置 (LAD) の存在を仮定する姿勢と結びついていた。一方の連合主義は、言語に特化した生得的言語習得装置の存在を仮定せず、むしろ汎用的な認知能力による知識の獲得を言語習得に適用する見方をもっている。

さらに、認知科学を基盤とする第二言語習得研究者は、表象と処理 (processing) を区別する。前者は、心的機構において、より静的なものであり、後者は表象に伴う動的な

¹⁰ 表象主義および象徴主義は、実にさまざまな分野において使用される用語であり、その厳密な定義は定まるものではない。

本論では、以降に示すように、ある認知過程を説明するモデルが、想定上の構造化された、具体的なイメージ、記号、象徴などを含むか、といった観点においてこの用語を使用している。また、より一般的な意味において、連合主義も表象主義に立脚するとする考え方もあり、本論における簡易的な区別は、必ずしも多くの研究者の賛同を得られるものではない。

過程や操作を指す。表象と処理の二項対立は、第二言語習得研究においても重要な理論的観点でもある。たとえば、ある特定の個別言語を母語とする学習者が、ある文法の機能的側面において、母語話者とは大きく異なる観測を示すとき、その原因が、表象欠陥 (representation deficit; Hawkins, 2001) であるか、そうでなければ処理困難性 (processing difficulty/constraints; e.g., Prévost & White, 2000; Lardiere, 2000) であるかといった議論 (e.g., Jiang, 2004) は、第二言語習得研究における主たるテーマのひとつである。しかし、これらのような一部の研究を除けば、文法知識は、表象や処理がもつ両方の側面を含むものとして括られることも多い。

本論では、文法知識という用語が参照するドメインを、言語学的な観点から制限することと、表象と処理を区別することについては行わない。また、文法知識は、表象や処理の両方がもつとされる側面を含む概念として取り扱うこととしている。その理由は、本論の主張が、上記の区別を特段必要としないことにほかならない¹¹。

(f) 教育や学習の成果とされ、さらにそれが社会において広く容認されている

現代社会において、文法知識ということばが指示するものは、上記のような外国語教育研究やその関連分野で論じられているものと、必ずしも一致するわけではない。むしろ、「外国語における文法的側面に関する知識や技能」といった漠然とした意味によって理解され、さらに教育に関する話題の中で日常的に使用されるということは、特筆すべき特徴である。すなわち、文法知識ということばは、社会における通用概念であるだけでなく、教育用語としても頻繁に使用されるものである。

たとえば、外国語における一般的な標準テストは、文法知識を測定するとされる固有のセクションをもつ場合が多い。また、教育実践におけるカリキュラムやシラバスにも、文法知識を伸長させる、といった文言が見られることもある。それだけでなく、一般的な学習教材においても、文法知識の伸長を効果として謳うものがあったり、教育従事者および学習者自身が、「文法知識を伸ばしたい」というように、ひとつの意思決定上の観点とすることもある。

どのような教育的および指導的枠組みであったとしても、整合性の高いカリキュラムデザインや、実際の学習活動を構成するためには、その帰結をもたらす成果が、事前に明示化されていなければならない。同時に、そのように明示化された成果は、教育評価

¹¹ ただし、著者自身は、本論の主張と大きな齟齬を発生させるというわけではないものの、語彙主義、その中でも生成意味論 (Generative lexicon; Pustejovsky, 1991) がもつ言語観、つまり、有限であるところの語彙が有限の表象をもち、それらの表象がもつ要素の演算機構によって意味が動的に構成 (composition) されるという考えをもっている。そればかりか、実際に、この考えにもとづく研究を行って来ている (草薙, 2017a)。この考えは、明確に表象主義のものであり、さらに、表象と処理を区別する立場であるとして差し支えないし、有限であるところの表象の存在を積極的に仮定する点において、言語学的な観点によって、ドメインを常に制限しないわけではない。

と密接な関係をもち、さらに、それは教育や学習の質向上に資する重要な観点でもある。一般に、教育や学習の帰結を定量的にあらわすものを、教育ないし学習における成果として位置づける。これまでの国内における英語教育実践において、少数の例外を除けば、文法知識を成果の一部として捉えていたことは、疑いようのない事実である。

このようにマクロな視点によれば、社会全体において、そしてよりミクロな視点によれば、実際の教育現場において、文法知識という概念は、さまざまな局面において、教育という営為に深く溶け込んだものであるといえよう。

2.2 文法知識の二分法に関する構造的仮定

ところで、文法知識とは上記のように、実に複雑な概念であるが、これがさらに二つの種類に分けることができる、という見方は、実にさまざまな研究に見られるものである。第二言語習得研究に視野を限っても、このような文法知識の区分に関連する用語を無数に挙げることができる (e.g., Jiang, 2004, 2007)。それらのような無数の用語の中にあつて、もっとも代表的なものは、明示的知識と暗示的知識による二分法であり、この二分法は、さまざまな理論や用語が示す概念を取り入れた包括的なものであると考えられる (e.g., R. Ellis, 2004, 2005; R. Ellis et al., 2009)。

たとえば、明示的知識および暗示的知識の研究に関する第一人者である R. Ellis が過去に行つた主張 (e.g., R. Ellis, 2004, 2005, 2006) によると、当該の概念は、(a) 認知心理学における明示的学習および暗示的学習 (Sorace, 1985), (b) 手続き的知識 (Anderson, 1983), (c) Krashen のモニター・モデル (e.g., Krashen, 1981, 1994), (d) 自動化 (McLaughlin, Rossman, & McLeod, 1983; Segalowitz & Segalowitz, 1993), (e) Dickerson, R. Ellis, Tarone などによる中間言語可変性 (interlanguage variability; e.g., Dickerson, 1975; R. Ellis, 1985; Tarone, 1979, 1983, 1988) といった 1980 年代や 1990 年代におけるさまざまな研究の影響を受けた、複合的な産物であると理解することができる。しかしながら、同時に、その包括性によって、さまざまな点において曖昧性や問題点を残す類型であることには十分に注意しなければならない (Hulstijn, 2005)。また、本論の大部分において、第二言語習得研究における明示的知識および暗示的知識について述べるときは、R. Ellis や彼を踏襲する研究の枠組みに視野を絞って議論を進めている。さらに、過去の第二言語習得研究における通例的な用法がどうであれ、ここでは、第二言語習得研究における明示的知識および暗示的知識は、認知心理学一般でいわれているところの宣言的知識と手続き知識、陳述記憶と非陳述記憶、または後述する人工言語学習研究において使用される明示的知識および暗示的知識といった用語と同一ではない、と著者は考えている。

第二言語習得研究における明示的知識および暗示的知識という二分法の導出、成立、そして普及の過程を明らかにすることは、容易になされるものではない。ただし、大局的見地に立った場合、この枠組みに対して、もっとも直接的な影響を与えた研究は、1980

年代後半から 1990 年代になされた中間言語可変性研究であり、この研究分野を基盤として、R. Ellis を始めとした研究者が、さまざまな研究の知見を、明示的知識および暗示的知識の二分法に対して統合するような形で付け足していったものであると、著者は考えている (e.g., 草薙, 2015; Kusanagi & Yamashita, 2013)。

その根拠のひとつは、明示的知識および暗示的知識に関する第一人者である R. Ellis は、この枠組みを提唱する以前に、中間言語可変性に関する研究者であったということと、中間言語可変性研究が、その名の通り、運用に関する個人内分散ないし課題間分散 (inter-task variance), または課題効果 (task effect) に対して多大な関心を寄せていた、という歴史的経緯である。中間言語可変性研究の流行は、社会言語学の流行を受けてのことでもあったが、当時の第二言語習得研究における、もっとも主要な関心は、第二言語学習者の運用に見られる個人内分散を、いかに体系的に、そして合理的に説明できるかということであった (e.g., Dickerson, 1975; R. Ellis, 1985; Tarone, 1983)。

個人内分散に対する関心は、当時の第二言語習得研究において、非常に先進的な視点であったと考えられる。1960 年代における初期の第二言語習得研究は、第一言語獲得と第二言語習得という二項対立、ないし母語話者と学習者という二項対立を中心としたものであった (Hulstijn, 2005)。誤用分析や対照研究といった古典的な研究手法にも数多くの例が見られるように、母語話者がもつ文法知識と学習者がもつ文法知識は、二項対立に据えられ、明らかに文法知識は、集団に帰属するものとして扱われていた。また、特に学習者において、文法知識を個人に帰属させる視点は、文法知識の発達という重要な観点にも繋がった。

このような歴史的経緯の中にあって、「ある課題の条件下における学習者の文法知識」ないし「ある心理的条件または、社会的、言語的環境下における文法知識」といった中間言語可変性研究の新しい見方には、文法知識を集団や個人に帰属するものから、状況へ帰属するものとして捉え直した、という一種の認識論的転換が見られる。R. Ellis (1985) では、この認識論的転換が実に明確に示されている。ある学習者が産出した、以下の 2 つの文を参照されたい。これは R. Ellis (1985) が示した例である。

(1. a) *Moses lives in Lusaka.*

(1. b) * *Moses lives in Lusaka, but he work in Kafue.*

R. Ellis (1985, p. 120)

さて、この 2 つの文を見てみると、(1. a) の文では、その単純な文構造において、発話者が、動詞 *live* を適切に屈折させている。しかし、一方の (1. b) の文では、後続する 2 つ目の節における動詞、*work* についてのみ屈折がなされていない。当該の研究では、こ

のような個人内における運用の可変性を、文脈可変性 (contextual variability) や、状況可変性 (situational variability) と呼んでいる。これらの用語が明示するように、文法知識について、集団や個人という帰属先から、状況に対する帰属先へと変換した見方があることは明らかである。

上記に見られるような、特性の帰属先に関する認識論的転換は、前節において紹介したように、心理学における不安研究などとまったく相同しており、人文社会科学研究において、しばしば見られる現象であるといっても過言ではない。帰属理論における共変モデル¹²の通り、当該の認識論的転換が起きる条件は、上記のような個人内における運用の変動が、相対的に個人間における変動よりも大きい、といった観測によるものであると考えられる。たとえば、条件 A と B における、ある個人 x の運用に見られる差が、条件 A における x と、別の個人 y における運用の差よりも大きい、といった具合である。当時、第二言語習得に関する課題や測定法が多数開発されたことも、このことの背景にあったと考えられる。たとえば、1970 年代の形態素習得研究 (e.g., Bailey, Madden, & Krashen, 1974; Dulay & Burt, 1973) では、自然順序 (natural order) という用語が示すように、現象としての第二言語習得は、ある種の普遍的性質に特徴づけられていると主張されていた。しかし、普遍的性質を示すかのように主張された形態素習得順序についても、5 種類の課題を同時に分析した Larsen-Freeman (1975) によれば、課題間におけるスコアの順序はまったく一致しておらず、逆に課題間による差が非常に大きいということが主張されている。中間言語可変性研究の理論的動機は、このような課題間分散の発見に起因すると考えられる。

ここで、個人内分散ないし課題間分散を説明するための理論的枠組みについて注目してみる。たとえば、中間言語可変性の主要な研究者のひとりであった Tarone (1983) は、3つのパラダイムと称する見方を提案している。この見方は、当時の中間言語可変性研究の中においても中心的なものである。

ひとつ目は、単一能力パラダイム (the homogenous competence paradigm) と呼ばれるものである。このパラダイムでは、言語運用に関して、単一の能力のみが仮定され、個人内分散および課題間分散は、能力とは独立した運用上の問題、ないし処理制約 (processing

¹² 帰属理論における共変モデル (e.g., Hewstone & Jaspers, 1987; Kelley, 1972) は、ANOVA モデルとも呼ばれることがあり、一般化可能性理論と同様に、さまざまなファセット (相) がもつ分散成分の相対的な大きさに注目する。このモデルの下では、分散の原因について三種類を想定する：(a) 一致性 (consensus), (b) 固有性 (distinctiveness), (c) 一貫性 (consistency)。一致性とは、異なる人物間におけるある行動の共分散である。固有性とは、ある状況において、ある行動が安定している程度である。一貫性とは、時系列間における個人内の共分散である。一致性と固有性が低く、一貫性が高い現象について、人は個人に帰属すると考える傾向がある。または一致性、固有性、一貫性が高い行動は、刺激に帰属すると考えられる。さらに、一致性が高く、固有性と一貫性が低い行動は、状況に既存すると考えられる。中間言語可変性研究は、まさに可変性という用語が示すように、一貫性の低さを繰り返し主張していることに注目されたい。

constraints) などに帰されるものとされた。このパラダイムは、当時、第二言語習得研究においてその影響が根強かった生成文法の思想を顕著に表しているともいえる。

2つ目のパラダイムは、混成能力パラダイム (the capability paradigm)¹³ と呼ばれるものである。このパラダイムは、Tarone 自身が提唱しているものである。このパラダイムにおいて、能力とは、本質的に混成的 (heterogeneous) なものであり、タスクや社会的文脈に対応するスタイルによって影響される、有限離散的であるが、定数をもたない¹⁴ 言語規則の集合体と捉えられる。このパラダイムは、特に社会言語学的知見を踏まえたものであった。

最後のパラダイムは、二重能力パラダイム (the dual competence paradigm) である。このパラダイムは、Krashen のモニター・モデル (1981) を踏襲し、習得された知識 (acquired knowledge) と学習された知識 (learned knowledge) という 2 つの排他的な要素こそが、個人内分散および課題間分散の源であるとみなす¹⁵。このモデルは、個人内分散を説明するために、個人が、2 つの構成概念をもつと仮定する。ここでの構成概念、つまり、それぞれ異種の知識とは、表象の心的格納庫 (mental storage of representations) を意味しており、さまざまな言語規則を示す表象群が、二種類の心的格納庫 (習得された知識 vs. 学習された知識) にそれぞれ格納されていると考えられた。その上で、ある社会的要因、課題要因、そしてスタイルなどは、行動ないし心的機能としてのモニターが利用可能かどうかという条件に対して影響を及ぼし、モニターは、学習された知識に格納されている表象のみを利用する、というように説明された。

二種類の心的格納庫が存在するという、当時の議論の中において主流派ではなかった、この二重能力パラダイムに見られるモデリング・コンセプトは、次節でより詳細に説明するところの明示的知識および暗示的知識の二分法の中に残り続けることになった。このコンセプトについては、先述したように、個人内分散を説明するためのものとして、2 つの構成概念を導出したことにそのもっとも顕著な特色がある。ここで、二重能力パラダイムのコンセプトを表すパス図を図 2.2 に掲載する。このコンセプトの主な特徴は、少なくとも数理的に、現在においてもっとも一般的な明示的知識および暗示的知識の二分法に仮定されている条件と適合する。

このように、「外国語の文法知識は二つに分けられる」という見方、つまり件の二分法

¹³ この訳は定訳でないことに注意されたい。

¹⁴ ここでの含意は、言語能力が無限数存在することではなく、究極的には計上可能であるというを示している。定数をもたないとは、具体的に何種類の言語能力かを固定しないということである。

¹⁵ 当時の研究において、本論が述べる帰属理論や個人内分散ないし個人間分散といった観点があつたわけではないことに注意されたい。歴史的に見れば、これらの研究は、社会心理学における帰属理論が成立する前であり、個人間分散、課題間分散などといった用語はほぼ使用されることはなかった。本論において、過去の理論をよりメタ的に捉える道具立てとして帰属理論を適用しているのであって、当時のこれらの概念が帰属理論によって導出されたわけではない。

に関する導出背景については、運用において、個人内分散が個人間分散よりも相対的に大きい、といった観測の発見によって、文法知識を状況依存的なものであるとする見方が徐々に広まり、その説明の中で、観測の背景に 2 つの潜在変数を仮定するひとつの方略が、当該の研究に関わる経緯の中で選ばれた、と理解することができよう。本論では以降、このことを「文法知識の二分法に関する構造的仮定」か、またはこれを略して構造的仮定と呼ぶ。この構造的仮定をより厳密に定義すると、個人の課題間分散を説明することを目的とし、複数の課題成績を観測変数とした二因子モデルである。

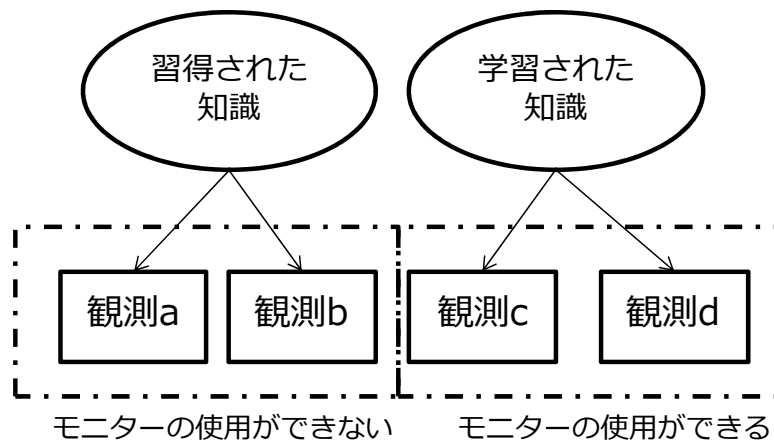


図 2.2 二重能力パラダイムがもつ構造的仮定

しかし、残念ながら、この構造的仮定は、当時、実証的に検討されることがなかっただけでなく、中間言語可変性研究は、大局的に見ると、1990 年代後半には下火になった。1980 年代や 1990 年代の第二言語習得研究は、現在のような高度な測定技術や解析技術をもたなかったため、非常に合理主義的なものに留まった。そして、上記の 3 つのモデルは、当時の測定および解析技術において、実質的に識別可能なものでもなかった。ただし、明示的知識および暗示的知識の二分法に限らず、上記の 3 つのパラダイムが取り扱ったような、運用における個人内分散に対して説明を与えようとする試みは、その後の第二言語習得研究に関するさまざまな枠組みの研究にも、その名残りをみることができる (e.g., Jiang, 2004, 2007)。

2.3 構造的仮定の実証と明示的知識および暗示的知識がもつ特性

上記の構造的仮定が息を吹き返し、多数の研究者に認識されるようになったきっかけは、時を経て 2000 年代、R. Ellis (2004), R. Ellis (2005), R. Ellis et al. (2009) などの研究である。R. Ellis と彼の共同研究者が行ったこの一連の研究は、Marsden Project と呼ばれ、R. Ellis et al.として一冊の書籍にまとめられている。これらの研究は、まさに上記

の構造的仮定についての直接的な実証を目指したものであった。実際に、R. Ellis (2005) は、潜在変数を 2 つと仮定した構造をもつモデルが、観測に対して比較的高い適合性をもつと報告し、これは明示的知識および暗示的知識という枠組みを支える直接的な証拠とされた。このモデルにおいて、(a) メタ言語テスト、(b) 時間的制約を設けない GJT は、明示的知識として推定された潜在変数に対して高い負荷をもつ観測変数群であり、(c) 時間的制約を設ける GJT、(d) 口頭産出テストは、暗示的知識として推定された潜在変数に高い負荷をもつ観測変数群であると主張された。さらに、その後、このモデルにおいて、その妥当性を支えるものとする証拠が次々と報告された (e.g., Bowles, 2011; Erlam, 2009; Loewen, 2009)。

さて、この構造的仮定に関する実証が十分であったかどうかは、次章の内容に譲るとして、ここでは、当該プロジェクトの理論的貢献について考えたい。一連の研究は、R. Ellis (2005) に見られるように、明示的知識および暗示的知識に関する幅広い特性の列挙より始まっている。R. Ellis が列挙した明示的知識および暗示的知識の特性を、簡易的にまとめると表 2.1 のようになる (R. Ellis, pp. 151-152; 著者訳)。この表が示すものは、現在においても、明示的知識および暗示的知識についての、主流の認識論的基盤になっていると考えられる。

表 2.1

R. Ellis による明示的知識および暗示的知識がもつ特性の一覧

特性	明示的知識	暗示的知識
意識の程度	高い意識の程度	低い意識の程度
時間	時間的制約がない場合に発現	時間的制約がある場合に発現
注意	形式へ注意	意味へ注意
体系性	ばらつきが大きい	ばらつきが小さい
確信度	不確かであるという印象	確かであるという印象
メタ言語	使用する	使用しない
学習可能性	大人が学習しやすい	子どもが学習しやすい

本論では、中間言語可変性研究から、これらの一連の研究に至る同一の関心の下、運用に関する個人内分散を取り扱うために、構造的仮定が先んじて導出され、この表に列挙される特性群は、構造的仮定の導出後に足されたものであると捉えている。このような経緯は、R. Ellis (2004, 2005) などにおいて、(a) 操作的定義といった用語がしばしば見られること、つまり構造が二因子となるように事前に相当の調整を行っていること、また、(b) 潜在変数の構造自体やモデルにおける制約についての議論を一切行っていないこと、などからも明白である。つまり、明示的知識および暗示的知識という概念の存

在とその構造を所与とした上で、それらの構成概念に対して、具体的な特性の記述を事後的に与えたということである。本論では、このことを構造の優先と呼ぶ。

よって、表 1 に見られるそれぞれの知識についての特性は、まさに本節の最初に述べたように、中間言語可変性の研究に限らず、認知心理学や認知神経科学といったさまざまな分野の概念を、構造的仮定に沿うようにするため、包括的に参照したものであると考えられる。たとえば、上記のうち、厳密に切り分けることは困難であるが、意識の程度と確信度といった概念は、認知心理学における明示的知識および暗示的知識に関する研究に由来する。一方、メタ言語、学習可能性、注意といった概念は、従来のさまざまな第二言語習得研究の知見に由来するものである。また、時間という観点に関しては、R. Ellis による独自の導出であり、これら、独立に導出されたさまざまな知見が、どのようにして十把一絡げに統合可能になるのか、という点についての根拠は薄弱である。R. Ellis が提案した上記の包括的二分法がもつ貢献は、学術史上、十分に揺るぎないものであったという評価については論を待たない。しかし、本論の実証部分では、結果的に上記のような見方は誤りであると主張することになる。

次節からは、R. Ellis による上記の分類の代替案として、明示的知識および暗示的知識がもつされるさまざまな特性を、(a) 運用、(b) 心的状態、(c) 獲得過程という 3 つの観点から整理できることを示しながら、文法知識がどのように二分されるか、またはされないかという点について考察していく。そして同時に、以降の議論は、R. Ellis による分類の要点を包含するものになると考えている。

2.4 文法知識の二分法に関する運用

文法の運用とは、もっとも一般的な場合、観測変数であるところの課題成績やテスト得点などである。外国語教育研究、第二言語習得研究、そして言語テスト研究では、学習の到達度ないし熟達度といった概念に対して、次元性という測定論的性質を付与してきた。つまり、ある一次的尺度上における値の大小といった具合である。これは、先述の通り、上記の分野が、学習者と母語話者という二項対立の上で発展してきたこと (Hulstijn, 2005) がその主な理由だと考えられる。たとえば、学習の到達度や熟達度は、母語話者に対する近似度である、といった具合である。

もちろん、学習者と母語話者の言語運用には、観測上、大きな差があることは事実である。第二言語習得研究において、もっとも古典的な測定具である GJT (Johnson & Newport, 1989) を例にとると、GJT における母語話者の個人正答率は、100% やそれに近い値を示すことが多い。また、逆にそれを期待して項目が作成されることもある。たとえば、生成文法では、母語話者の、ある文に対する容認性は、構造に起因するところの文法性を検証する道具である。一方、一般的な学習者の個人正答率は、例外的少数を除けば、母語話者が示すような値に遠く及ばないことが多い。また、母語話者程度 (native-like) と

という言葉があるように、仮に学習者が母語話者の成績と遜色ない言語運用を示したとするならば、それは学習の最終到達 (ultimate attainment; e.g., Birdsong, 1992; Johnson & Newport, 1989) を示すか、そうでなければ、その個人は二言語話者 (bilingual) であるとみなすことも多い。

また、個人正答率といった課題成績に次いでもっとも広く容認されている、反応時間データや視線計測データといった心理言語学的変数も同様に、言語運用に類する変数である。さまざまな心理言語学的課題において、ほとんどの場合、母語話者は学習者に比べ、圧倒的に短い反応時間や速い読解速度を示す (e.g., Clahsen & Felser, 2006a, 2006b; Frenck-Mestre, 2002)。2000 年代以降の第二言語習得研究の一部、たとえば文処理研究 (grammatical processing studies) は、心理言語学や認知神経科学の手法を取り入れた上で、学習者と母語話者間における心理言語学的実験における結果の比較を集中的に行ってきた (Clahsen & Felser, 2006b)。また、心理言語学的実験に関わる変数だけでなく、非侵襲的な脳機能イメージングデータなどを使った研究も、同種の群比較を行っている (e.g., Abutalebi, Cappa, & Perani, 2001; Perani & Abutalebi, 2005)。学習者と母語話者間の脳機能上の差異などについては、脳機能局在論の観点から議論されることも多い。

また、観測における二群の中心傾向に限らず、それが中間言語可変性研究の主要なテーマであったように、学習者の運用における散布度も重大な関心の的であった。ほとんどの場合、学習者の運用は、母語話者のそれに比べて非常に大きな散布度を示すことがわかっている。より正確さを期すならば、ある観測変数において、母語話者群が示すであろう個人間分散および個人内分散は、学習者群が示すであろうそれらよりも遥かに小さい、といったように形式化できる。この観点は、R. Ellis によるところの分類の、体系性などにも通ずる。このことは、母語話者の文法知識が、特に生成文法に顕著なように、集団に帰属するものとみなされると同時に、学習者の文法知識が、集団や個人だけでなく、中間言語可変性研究に顕著なように、状況に対して帰属するものとみなされるようになったことからわかる。

このように、文法の運用に関わるさまざまな観測において見受けられる、学習者と母語話者の差は、それぞれの集団において特異的な構成概念の違いに帰着すると考えられるようになった。このような理論史上の展開を、Hulstijn (2005) は以下のように簡潔にまとめている。

There are good theoretical and educational reasons to place matters of implicit and explicit learning high on the agenda for SLA research. As for theoretical motivations, perhaps the most central issue in SLA theory construction in need of explanation is the differential success in one's first language (L1) and in one's second language (L2).

(Hulstijn, 2005, p. 129)

How can we explain universal success in the case of L1 acquisition and differential success in the case of L2 acquisition? Among the many explanations that have been proposed, (中略) ... one finds explanations that involve the notions of implicit and explicit learning.

(Hulstijn, 2005, p. 129)

母語話者と学習者の運用における差の説明とは、具体的にどのようなものだろうか。たとえば、学習者、特に入門レベルや外国語環境にいる者は、明示的知識のみをもち、母語話者は、暗示的知識のみをもつ。その上で、明示的知識は、低い課題成績および遅い反応時間の原因であり、一方の暗示的知識は、高い課題成績および早い反応時間の原因である、といった具合である。より一般的な観点においても、明示的知識のみをもつ者は低いレベルの運用に終始し、一方、暗示的知識をもつ者は母語話者程度の言語運用を示す、といった表現を見ることがある。

ここに見られる、学習者と明示的知識、母語話者と暗示的知識といった二項関係の類推は、外国語教育研究および第二言語習得研究において普遍的に見られるものである。運用に関わる平均値や分散の構造から、文法知識は、集団、すなわち学習者や母語話者に帰属するもの、という認識が構成されたように考えられる。一方、理論的な観点では、暗示的知識は、生成文法で定義されるような生得的な機構の発露による母語話者の言語知識である、または、連合主義的にパターン学習によって獲得された知識体系であり、そして明示的知識は一般的な学習過程によって得られる知識体系である、といった主張 (e.g., Ellis, 2005; Hulstijn, 2005) が見られた。さらに、学習者と母語話者の運用における観測上の差は、明示的知識および暗示的知識の二分法に関わらず、普遍文法へのアクセス可能性や、認知機構の本質的違いによって説明されることもある (White, 2003)。しかし、これらの本質的な構造は相同している。この学習者と母語話者、明示的知識と暗示的知識という二つの二項関係は、以降で説明する獲得過程にも通じるものがある。

2.5 文法知識の二分法に関する心的状態

心的状態は、主に認知心理学の研究に由来するところが大きく、前節のように、学習者と母語話者の二項対立に着目したものではなく、その導出過程は基本的に独立である。

認知心理学では、意識、とりわけ学習における意識の役割についての研究が盛んであり、認知心理学の研究者は、明示的知識と暗示的知識という用語を、意識という文脈において頻繁に使用する (e.g., Dienes, 2005, 2008)。しかしながら、これらの用語によって、外国語教育研究および第二言語習得研究の研究者が参照するものと、認知心理学の研究者が参照するものが、本来であれば同一であるかどうかには相当の注意が必要である (草薙, 2015; Tamura, Harada, Kato & Kusanagi, 2016)

明示的知識および暗示的知識という用語を使用する、認知心理学における最も典型的な研究分野は、人工言語学習研究である。これらの研究における明示的知識とは、いわば意識的知識 (conscious knowledge) を意味しており、知識表象としては、*"I know that I know P"* というように、その表象自身に再帰的要素を含んでいるものと仮定されている。これは、高階思考 (Higher Order Thought) などとも呼ばれるものに理論的基盤を置いている (e.g., Dienes, 2004; Perner & Dienes, 1999; Rothental, 2000)。ここでは、命題 *P* について、自分が知っていることを知っている、という意味で、命題 *P* を基準にして、より一段階高次であると考えるのである。一方の暗示的知識とは、いわば無意識的知識 (unconscious knowledge) であり、知識表象としては、単純に命題 *P* というように、再帰的要素を含まない。

無意識的知識の測定法や、それに関する実験手順については、これまで実にさまざまなものが提唱されている (e.g., Dienes, 2004, 2008; Dienes & Scott, 2005; Scott & Dienes, 2008)。現在においてもっとも広く受け入れられている方法は、判断行動 (judgment behavior) を伴う各種の課題成績を分析することであろう。当該課題の分析に際しては、下された判断が正答か誤答かといった課題成績のみならず、主観的変数 (subjective measures) という特徴的な変数とその解析法を導入する。

主観的変数は、課題における各試行について、その判断に随伴する心的状態を被験者に対して内省的に問うもので、複数の質的水準を設け、それを実験参加者に選択させることがその典型例である。これは、従来より意識の測定法として知られた、言語報告 (verbal report) や思考発話法 (think aloud method) を計量的に簡易化した方法論である。たとえば、確信度 (certainty) について、「正解であるとの確信がある」ないし「正解であるとの確信がない」といった二択を設けて尋ねる場合や、その判断に伴う内省的な心象について「規則」(rule)、「記憶」(recollection)、「直観」(intuition)、「親密さ」(familiarity)、「当て推量」(guess) といった水準を被験者に選択させるといった場合が例として挙げられる (e.g., Dienes, 2008; Scott & Dienes, 2008)。認知心理学における再認記憶研究も、同様の実験的手続きを取ることがある (e.g., Yonelinas, 2002)。

この方法論において、規則反応や記憶反応は、意識的な構造的知識 (structural knowledge) かつ意識的な判断知識 (judgment knowledge) の表れであり、直観反応や親密さ反応は、無意識的な構造的知識および意識的な判断知識の表れ、さらに当て推量反応は、無意識

的な構造的知識および無意識的な判断知識の表れであるとされる (e.g., Dienes, 2008; Scott & Dienes, 2008)。ここでいう構造的知識とは、ある文法構造の規則に関する命題のようなものであり、判断知識とは、判断に関する知識、つまり一種の高階思考を反映するものとされる (Scott & Dienes, 2008)。この分野では、規則反応、記憶反応が、明示的知識を反映し、それ以外の反応は暗示的知識を反映するものとして受け入れられている。繰り返すことになるが、ここでいう明示的知識および暗示的知識は、外国語教育研究や第二言語習得研究でいわれる明示的知識および暗示的知識とは、必ずしも同一なものではない。

また、当該の研究分野では、主観的変数と課題成績の関係も重要な観点である。そもそも、知識がある、または知識がない、ということは、ある課題、たとえば二択問題において、チャンスレート以上の成績を示すかどうかといった基準によって判定すべきである。よって、ある被験者が、ある刺激に対して暗示的知識をもつとは、「当て推量反応を示した試行における課題成績がチャンスレート以上である」といった条件によって判断されることになる。これは、当て推量基準 (guessing criterion) と呼ばれる解析法のひとつである (e.g., Dienes, 2008; Scott & Dienes, 2008)。

また、同様に、確信度ないし自信 (confidence rating) と成績が、観測において無関係であることが、被験者が暗示的知識を使用した証拠であるとみなすこともある。これを無相関基準 (zero-correlation criterion) と呼び、当て推量基準やこれらの亜種と発展形とを合わせ、総称してメタ知識基準 (meta-knowledge criteria) と呼ぶ (Dienes, 2008)。逆にこれらの相関が強い値を示すとき、それは明示的知識の使用を示す証拠ともみなされる。このような方法論を採用した第二言語習得研究は、少数であるが、近年比較的増えてきているようである (e.g., Rebuschat, 2013; Rebuschat, Harmick, Riestenberg, Sachs, & Ziegler, 2015)。認知心理学に由来するこの観点は、R. Ellis の分類における、意識の程度、確信度、注意、そしてメタ言語などに関する特性にも関係している。本来であれば、このような特性の、第二言語習得研究に対する適用は十分に検討されるべきものであったが、R. Ellis の分類では、明示的知識および暗示的知識という概念が、上記の心的状態に関する特性を部分的に踏襲するものになっている。

2.6 文法知識の二分法に関する知識の獲得過程

明示的知識および暗示的知識の獲得過程は、第二言語習得研究における主要な問題点のひとつである。特にこの問題は、インターフェースの問題として知られる議論に深く関連し、さまざまな研究者が議論に参加しているが、これといった定説がない状態である (R. Ellis et al., 2009)。また、インターフェースの問題は、R. Ellis (2005) などが述べるように、この問題に関する議論が盛んであった 1990 年代においては、単純に解析技術が未発達であったため、実証性が著しく乏しかった。それ以前に、そもそも実証可能かどうか不明である主張も数多く含まれており、細部に目を配ると、検討しきれないほ

どの多様性があることも事実である。ここでは、上記のような理由から、知識の獲得過程に関する主な見方を、以下の3つに大きく分けて紹介することとする¹⁶。

(a) **独立過程論 (the independent process perspective)** : 明示的知識は明示的学習によって獲得され、暗示的知識は暗示的学習のみによって獲得される

(b) **連結過程論 (the conjunct process perspective)** : 明示的知識は明示的学習によって獲得されるが、何らかの行動によって、それが暗示的知識に変化する

(c) **交差過程論 (the cross process perspective)** : 明示的知識は明示的学習によって獲得され、暗示的知識は暗示的学習のみによって獲得されるが、明示的知識については、何らかの行動によってそれが暗示的知識に変化する

独立過程論

これは、一般的にノン・インターフェースと呼ばれる立場に相当し、Krashen のモニター・モデルがその元型である (Krashen, 1981)。独立過程論は、明示的学習および暗示的学習という区分を設けており、さらに知識と学習において完全な対応関係を仮定している。明示的学習と暗示的学習を定義することは非常に困難であるが、ここでは、認知心理学において意味するところの定義を使用する。すなわち、明示的学習は、意識ないしメタ認知を伴う学習行動の一種であり、暗示的学習は、意識ないしメタ認知を伴わない。

より厳密に、独立過程論の含意を表すとするならば、以下のように考えることもできる。つまり、構成概念であるところの明示的知識については、基本的に明示的学習と呼ばれる行動の事前および事後間において値が変化し、暗示的学習と呼ばれる行動によっては変化しない。それと同時に、暗示的知識については、基本的に暗示的学習と呼ばれる行動の事前および事後間において値が変化し、明示的学習と呼ばれる行動によっては変化しない。

連結過程論

これは、一般的に強いインターフェースと呼ばれる立場に相当する。この立場の含意は、明示的知識と暗示的知識がともに独立した構成概念として存在し、継続的な明示的学習や反復などの行動によって、明示的知識が暗示的知識に変容するということである。この立場は、技能習得理論、自動化理論 (Segalowicz & Segalowicz, 1993) や Anderson (1983)

¹⁶ これらの用語は、著者独自のものであり、まったく一般的なものではない。大局的に見て、ここでの独立過程論は、より一般的な第二言語習得の専門用語であるところのノン・インターフェースの立場、連結過程論は、強いインターフェースの立場、交差過程論は、弱いインターフェースの立場にそれぞれ相当する。ただし、本論において、これらの用語が指し示すものは、過去の第二言語習得研究における含意と全て一致しているわけではない。特に、インターフェースの議論と呼ばれる用語は、観測の共変性などについての含意をまったくもたないため、従来の用語の使用を避けた。

などの立場¹⁷と相同すると考えられる。ただし、この立場では、暗示的学習、特に無意識的学習の存在や効果については、あまり積極的に論じられることがなく、むしろ明示的知識の使用が反復に伴い、その手続きが自動化され、暗示的知識に変化する（Anderson, 1983）といったような説明を好む。この主張では、構成概念であるところの明示的知識と暗示的知識が、ある種の学習の事前および事後において共変する、といった観測を予測するところになる（e.g., Ziori & Dienes, 2012）。日本などの EFL 環境における英語教育関係者は、このような主張をもつことが相対的に多いと考えられる。

交差過程論

これは、一般に弱いインターフェースと呼ばれる立場の亜種に相当する。この見方は、端的に言って、上記 2 つの見方の混合型である。そのもっとも顕著な点は、暗示的知識の獲得過程として、暗示的学習と、反復などによる自動化といった 2 種類の過程の存在を認める点である。

このような見方によると、無意識学習や反復といった行動の両方は、その事前および事後間において構成概念である暗示的知識の値に影響を及ぼすことが期待される。また、少なくとも交差過程論の上では、構成概念である暗示的知識の値は、その獲得過程から弁別可能なものではない。

一方、研究者の中には、暗示的知識とされているものの獲得過程を差別化するために、自動化された明示的知識（automatized explicit knowledge）といった概念を設けるものがある（Suzuki, 2017）。Suzuki（2017）が述べるように、自動化された明示的知識という概念は、これまで十分に検討されているとはいえ、今後の研究が期待されるものである。しかし、少なくとも、自動化された明示的知識という概念を使用する研究者は、無意識的学習といった行動と、反復といった行動の帰結は、任意の観測から弁別可能である、という前提をもっていると推測できる。また、この主張によると、文法知識は（a）明示的知識、（b）暗示的知識、そして（c）自動化された明示的知識、の三種類に分けられているはずである。その上、暗示的知識と自動化された明示的知識については、上記の連結過程論の含意とは異なり、独立した過程によって得られると仮定しているであろう。

上記の 3 つのような観点は、後ろ向きデータ、つまり言語適性、学習開始年齢といったデモグラフィックデータとの関連によって論じられる場合も多く、R. Ellis による分類では、学習可能性などに相当する。

¹⁷ ここでの Anderson などの立場とは、具体的に ACT-R 理論や ACT* という枠組みを参照している。これらは、記憶、言語理解や学習に関する包括的な認知モデルであり、宣言的知識と手続き知識の導入をそのもっとも顕著な特徴としてもつ。これらのモデルは人工知能研究などにも応用されており、計算論上の観点とも親和性が高い。

2.7 運用と心的状態の関係

ここまでは、文法知識の二分法に関して、(a) 運用、(b) 心的状態、そして (c) 獲得過程という3つの異なる観点があることを紹介した。しかしながら、以前述べたように、これらの観点は、R. Ellis が主張するように容易に結びつけられるものではない。その問題点は、これら3つの観点における関係性である。ここからは、R. Ellis (2005) が仮定していると考えられる、運用、心的状態、そして獲得過程間の関係性について批判的に見ていく。

まずは、運用と心的状態の関係性である。R. Ellis (2005) や第二言語習得研究における一般的な知見によると、明示的知識は、学習者に典型的なように、低いレベルの運用を示す観測の原因になり、暗示的知識は、母語話者に典型的なように、高いレベルの運用を示す観測の原因になる。ここでいう明示的知識および暗示的知識を、認知心理学における定義と結びつけると、意識性を示す心的状態は、運用を示す観測において低い値を、無意識性を示す心的状態は、運用を示す観測において高い値を示すという予測が立てられるだろう。この仮定は、実証性が乏しいばかりか、合理的に考えても真であるようには考えられない。

ここで、話を具体化するための例を出す。認知心理学において典型的なように、運用を GJT における正答率に読み替えるとしよう。そうすると、意識性を示す心的状態、たとえば規則反応や高い確信度は、概して低い正答率を予測するはずである。そして、無意識性を示す心的状態、たとえば直観反応や低い確信度は、高い正答率を示すはずである。さらに、言い換えるならば、意識的知識の方が総じて低い水準の運用における原因となり、無意識的知識の方が高い水準の運用における原因となるはずである。

同様に、GJT における反応時間に対しても、同様の予測を立てることができる。すなわち、意識性を示す心的状態、たとえば規則反応や高い確信度は、長い反応時間を予測するはずである。そして、無意識性を示す心的状態、たとえば直観反応や低い確信度は、短い反応時間を示すはずである。

このことに関して、R. Ellis (2004) は、明示的知識および暗示的知識とその作動時間に着目して、独自のモデルを提案している。このモデルは、第二言語における文法性の判断には、順次作動的な3つの認知的過程が含まれていると仮定する。一つ目は、意味処理 (semantic processing) であり、この過程では、文の意味に焦点を当て、およそ自動的に文の意味についての理解をすると仮定されている。この際、学習者は暗示的知識のみへアクセスしているとされている。次の過程は、気づき (noticing) である。この過程では、文中における非文法的な要素を探索するとされている。この過程についての詳細は、他の過程に比べて詳細に述べられていない。最後の過程は、内省 (reflecting) である。この最後の過程では、文中のどの要素が非文法的であるか、またはそれはなぜかといったことについて、明示的知識へアクセスをしながら同定するとされている。このモデル

がもつ特徴は、大きく分けて二つである。一つ目は、これらの過程における強い順次性である。時間軸の上で、意味処理は、気づきに対して必ず先行し、気づきは内省に対して必ず先行する。さらに、二つ目の点は、時間制限といった課題条件から、特に内省の過程を阻害することができる点と期待されている点である。時間制限を設けることによって、内省、つまり明示的知識へのアクセスを防止することが仮にできるとなると、時間制限という課題条件は暗示的知識を引き出す方策といえる。このように、明示的知識は遅いパフォーマンス、暗示的知識は早いパフォーマンスを表す、ないし時間制限を設けると暗示的知識のみへのアクセスを、設けないと明示的知識へのアクセスを外発的に促す、というように仮定された。

しかしながら、このような仮定は、認知心理学におけるこれまでの実証研究と徹底的に矛盾している (e.g., Mealor & Dienes, 2012, 2013; Ziori & Dienes, 2012)。人工言語学習研究では、無意識的な判断知識の使用を示す観測の反応時間は長く、さらに意識的な判断知識の使用を示す観測の反応時間はより短いことを報告している (e.g., Ziori & Dienes, 2012)。もちろん、同様に、無意識的な判断知識の使用を示す観測が正答である確率は低く、意識的な判断知識の使用を示す観測が正答である確率は高い、という結果も示されている (Ziori & Dienes, 2012)。

また、認知心理学における再認記憶研究では、R/K 手続き (Remember/Know Procedure; e.g., Yonelinas, 2002; Tulving, 1985) という主観的変数に似た手法を使用するが、同様に、Remember 反応 (学習の記憶を思い出すことができる) の方が、一般的に高い正答率と短い反応時間を示し、Know 反応 (思い出すことができないが、親密度が高い) の方が低い正答率と長い反応時間を示すという報告がある (e.g., Dewhurst & Conway, 1994)。

また、自信と反応時間の関係について取り上げるならば、これは数理心理学 (mathematical psychology) における古典的な常識であり、自信があるという反応は、より短い反応時間から予測され、自信がないという反応は、より長い反応時間から予測されることがわかっている (e.g., Ratcliff & Starns, 2009; Volkman, 1934)。より明確に、初期の数理心理学では、択一課題における自信は、その試行の反応時間を項にもつ関数であるとされた。この自信と反応時間の関係についての数理モデルは、現在においても主要な研究テーマである (e.g., Pleskac & Busemeyer, 2010; Ratcliff & Starns, 2009, 2011; Voskuilen & Ratcliff, 2016)。

このように、他分野における観測的事実は、R. Ellis が述べるような明示的知識および暗示的知識の特性とは矛盾している。このことについては、R. Ellis の主張を好意的に解釈して、構造的仮定と構造の優先によるものであると考えているが、外国語教育研究において、明示的知識および暗示的知識という枠組みに関する運用と心的状態の関係性は、未だ十分に検証されているとはいえない。これは本論が実証することのひとつである。

2.8 運用と獲得過程の関係

次に、明示的知識および暗示的知識における運用と獲得過程の関係性について考察する。2.4節で述べたように、第二言語習得研究において、明示的知識は、学習者に典型的な運用、つまり、母語話者に及ばない正答率や、母語話者よりも長い反応時間といった観測に関連付けられており、一方の暗示的知識は、母語話者らしい運用に関連付けられてきた。この二項関係は、獲得過程を複雑に媒介して成り立っている。

まさにナチュラル・アプローチと Krashen が呼ぶように、母語話者らしい知識の獲得過程、すなわち習得は、母語話者がもつ暗示的知識を伸長させ、排他的に明示的知識を伸長させず、そして専ら批判の対象であるところの学習者らしい知識の獲得過程、すなわち学習は、学習者がもつ明示的知識を伸長させ、排他的に暗示的知識を伸長させない、と主張されている。それと同時に、母語話者らしい過程を経て獲得された暗示的知識は、母語話者らしい運用の原因であり、学習者らしい過程を経て獲得された明示的知識は、学習者らしい運用の原因となる。

しかしながら、本論は、厳密な意味でこのような論に立脚することはない。むしろ、上記の論は、母語話者と学習者に関連付けられる、さまざまな事象を二分化した短絡的な類推に過ぎず、少なくとも不確かな前提をもっている、と評価している。ここで、先述のように、外国語の運用を、具体的に GJT の正答率や反応時間に置き換えて考えてみる。たとえば、正答率のような尺度において、ある値が得られたとき、その値の大小のみから、その値の原因であるような過去の行動が完全に弁別可能であると考えることには、無理があるといわざるを得ない。より具体的にいい換えると、正答率が 95%であるという観測を根拠に、この正答率を生み出す原因は、母語話者らしい獲得過程である、と断定することは、相当に問題点を含んでいる。かなり冗長になってしまうが、以下のように論理を考えれば明確である。命題 P をある母語話者らしい運用、命題 Q を母語話者らしい獲得過程であるとしたとき、この主張は P と Q が同値¹⁸であると形式化することができる。

$$P \Leftrightarrow Q \quad (2.1)$$

この論理が成立し、そして観測からその獲得過程を弁別するためには、少なくとも P が真、 Q が偽のときに、この論理は偽でなければならない。これは、母語話者らしい運用を示し、かつ母語話者らしい獲得過程をもたないものの存在についての真偽であるが、これをおしなべて偽であると判断することは経験的に難しい。その理由のひとつは、母語話者と学習者の判別が不可能な観測は、合理的にいくらかでも存在しうるからである。たとえば、学習者のほとんどが天井効果を示すような測定具の上で、この情報のみから

¹⁸ 論理学上の同値は、 \Leftrightarrow と記す。同値とは、二つの命題がともに真および偽であるときに真となる。

学習者と母語話者を弁別することは不可能である。第二に、この論理が、その前提である母語話者らしい運用や母語話者らしい獲得過程という命題自体の定義に深く関わり、単純に論が循環するからである。たとえば、母語話者らしい運用とは、「学習者の運用と弁別ができるもの」と定義すれば、この論理は循環し、常に真であることになる。実際に、第二言語習得研究では、正答率データにおいて学習者と母語話者の差は観察されないが、反応時間データでは差が観察される、または反応時間データでは差が観察されないが、視線計測データでは差が観察される、といったように、学習者と母語話者における観測の差を示す実験方法や変数を探求してきている。たとえば、近年の研究では、より新しく、より課題の負荷が大きく、より時間に対して敏感 (time sensitive) な測定法が次々と新しい暗示的知識の尺度として一意的に提案されている (e.g., Suzuki, 2017)。このような状況では、この論理を信憑することが有益な方法にはなりえない。

より一般的な意味、たとえば認知心理学の再認記憶研究などにおいても、ある知識の存在を示すであろう観測、たとえば判断課題の正答率のみによって、過去の学習方法や学習行動を完全に弁別することが可能であるとはみなされない。個人の身長や体重のみから、過去の食生活を予測することが、現代科学において健全な因果推論とみなされないこととまったく同じである。

よって、本論では、少なくとも観測であるところの運用のみから、観測を経ない獲得過程を弁別することは不可能である、という立場を取る。この立場によって、本論では、独立過程論、連結過程論、そして交差過程論などに立ち寄ることはない。

本論は同時に、前節でも少し触れたように、もっとも一般的な獲得過程を表す証拠は、いうまでもなく時系列データや因果推定であると考えている。反実仮想的な (counterfactual) 因果モデル¹⁹にもとづくとする、ある学習過程が成果に影響を及ぼすことを推論するためには、それに先んじて、単一の成果が定められていなければならない。その上で、たとえば関心が学習と習得であったならば、仮想上、ある個人が学習をした世界とある個人が習得をした世界の二通りにおける、その個人の成果の値の差を検討しなければならない。

2.9 心的状態と獲得過程の関係

心的状態と獲得過程の関係も複雑なものである。認知心理学における人工言語学習研究では、一般的なレベルに限れば、意識を伴う学習は、意識的な判断知識に帰着し、意識を伴わない学習は、無意識的な判断知識に帰着すると考えられている。しかし、構造的知識はどちらの種の学習によっても獲得されると考え、この点において、構造的知識の有無から獲得過程は弁別可能なものではない。

ただし、上記は一般論であり、その詳細については決定的な説明がなく、諸説併存し

¹⁹ 一般にルービンの因果モデルとして広く知られる態度を示している。

ている状態である。Anderson (1983) は、学習の初期段階において明示的ないし宣言的知識が得られ、それが暗示的ないし手続き的知識に変化する、または宣言的知識を有することが、手続き的知識を獲得する条件であると考えている。Cleeremans and Jiménez (2002) の主張はその逆であり、暗示的知識は学習の初期段階のみに表れる、いわば、弱い知識表象の形態であり、知識表象の質が向上するにつれ、段階的に、より強い知識表象である明示的知識へ変化すると仮定している。このような仮説は、無意識的学習の成果と、学習以前の知識の有無との関係によって検証されている。Ziori and Dienes (2012) の実験では、暗示的知識のみを有するように統制した状態で、その後、無意識的学習を続けた被験者は、学習量が増えるに伴い、段階的に明示的知識の使用を示す反応を見せた。この結果から、Ziori and Dienes は、Cleeremans and Jiménez の主張を支持している。このように、無意識的学習が必ずしも暗示的知識の獲得のみへ影響するわけではない、ということを見ると、心的状態と獲得過程の関係は、それほど強いもの、たとえば同値であるなどとは容易にみなすことはできない。よって、本論では心的状態と獲得過程についても、獲得過程の弁別不可能性を仮定する。

2.9 第二章のまとめ

本章では、外国語における文法知識研究、その中でも明示的知識および暗示的知識に関する先行研究を、第二言語習得研究や認知心理学の観点から幅広く論じた。本章の要点は以下のようにまとめられる。

- (a) 文法知識は、集団、個人、そして状況に帰属するものとみなされると同時に、文法に関する運用の原因となる潜在変数である
- (b) 第二言語習得研究において、1980年代から1990年代に亘って流行した中間言語可変性研究では、当時さまざまな運用に関する測定法が開発されたという歴史的背景も相まって、従来、集団および個人に帰属するものとみなされてきた文法知識を、状況に帰属するものと捉える新しい見方が生まれた
- (c) 文法知識を状況に帰属するものと捉える見方より、Krashen のモニター・モデルや R. Ellis らによる心理測定的研究に見られる、明示的知識および暗示的知識の構造的仮定が導出された
- (d) R. Ellis らによる心理測定的研究では、明示的知識および暗示的知識の特性について、構造の優先によって、実証が十分でない主張が統合されることになった
- (e) 明示的知識および暗示的知識の特性については、その運用、心的状態、そして獲得過程という3つの観点から理解することができる
- (f) 運用と心的状態の関係は強いものであろうが、R. Ellis が主張する関係とはまったく異なる観測が得られる可能性がある

- (g) 観測であるところの運用から、明示的知識および暗示的知識の獲得過程を弁別することは不可能である
- (h) 心的状態と獲得過程の関係については諸説あるが、明示的知識および暗示的知識の獲得過程は独立および背反ではない

次節からは、このような知識の二分法とそれがもつ含意や問題点に関して、数理的アプローチの観点から、より批判的に検証する。

第三章 外国語における文法知識の数理

3.1 外国語教育研究における数理的アプローチ

この章では、明示的知識および暗示的知識の二分法に関して、特にその数理的な観点から、さらに批判的な考察を加えていく。しかしその前に、本論が全面的に依拠している、数理的アプローチ (mathematical approach) と筆者が呼ぶ、ひとつの研究姿勢について紹介しなければならない。

数理的アプローチは、国内外の外国語教育研究において、比較的新しい研究姿勢のひとつである。しかしながら、それと同時に、人文社会学のみならず、工学や自然科学をも含む、およそ全ての学術分野において、もっともありふれた、そしてもっとも基礎的な研究姿勢であるともいえる。現在では、数理的アプローチという用語ではなく、生物学、経済学、そして心理学を中心に、統計モデリング (statistical modeling) や数理モデリング (mathematical modeling) という用語を使うことによって、同種の研究姿勢を表すことも多い。国内外におけるベイズ統計の台頭や、データサイエンスの流行といった事柄にもその発露が見られるように、近年における解析技術の向上や解析環境の急激な普及を受けた、ある意味において流行の兆しを見せているアプローチでもある。

外国語教育研究における数理的アプローチは、技術至高主義に類するものではなく、むしろ、このアプローチは、技術的側面ではなく、ある種の顕著な思想的特徴によって形成されている。数理的アプローチは、具体的に以下のような思想的特徴をもっている：(a) 帰結主義・効用主義、(b) 経験主義・行動主義、(c) 形式主義・構造主義 (草薙, 2017b)¹。ただし、特に (a) に関しては、通常、統計モデリングといわれるアプローチがもつ思想的特徴とはやや異なる場合がある。

最初に、外国語教育研究の目的論に関する思想的特徴についてである。数理的アプローチは、外国語教育研究の究極的目標を、その研究の帰結、つまり、社会ないし個人の効用を高めることに据えている。同時に、あらゆる研究の価値は、研究過程のいかなる特徴からではなく、その研究がもたらしうる帰結から評価されるべきである、と考

¹ 本論における帰結主義、効用主義、経験主義、行動主義、形式主義、構造主義といった用語の含意は、科学思想や哲学に関わる用語としてというよりも、より一般的な用語としてのものである。たとえば、帰結主義および効用主義は、研究や教育実践についての価値判断を行なうとしたとき、その過程自体ではなく、その行為に由来する帰結を重視する考え方であり、功利主義とも呼ばれる。次に、経験主義および行動主義は、唯物論や機械論の系譜にあり、一般的に合理主義との対比によって論じられる性質である。本論は中でも、心理学における行動主義のアプローチないし、行動計量学のアプローチを基盤としている。この姿勢は、第二言語習得研究の文脈においては、データ駆動型アプローチ (data-driven approach) と呼ばれる考えに近い。形式主義および構造主義は、数学ないし社会学における厳密な含意ではなく、研究をするに当たって、自然言語による記述よりもなんらかの記号や構造的記法を重用する傾向を表す。

いる。これは、実用主義の一種であるとしても差し支えなく、本質主義²といった特徴をもつ従来の研究、たとえば第二言語習得研究に見られる研究姿勢とは大きく異なる。このような数理的アプローチの特徴は、外国語教育という営為の意義と目的を所与として、その意義と目的の背景に潜むあらゆる種類の秩序を、研究の実践に先んじて受け入れることに由来している。具体的な方策として、さまざまなレベルにおける現実の意思決定を媒介することを前提とし、帰結主義によって、その帰結に関わらない要素を積極的に捨象するという強い傾向をもつ。この研究姿勢は、経済学や工学などに広く見られる研究姿勢と共通しており、ほとんどの応用的学際分野において一般的なものである。今に続く外国語教育研究の歴史的成立過程から見ても、一種の帰結主義は、もっとも自然かつ順当な態度であると考えられる。

次に、研究の背景に関する思想的特徴である。数理的アプローチは、経験主義、その中でも厳密な行動主義を標榜している。数理的アプローチは、特に外国語教育に関して、研究者の合理性や内省に対して過大な信頼を置かず、ほとんどの場合、観測やそこから極めて自然に導かれ得る事象に対してのみ、優先的に信頼を置くこととしている。すなわち、ある認知過程といった、直接的観測が根本的に不可能な機構の存在について、観測と独立したなにかによって殊更確信をもつことはない。同時に、著者の研究姿勢に限れば、外国語教育研究や第二言語習得研究に関する諸概念の实在を容易に受け入れることはせず、むしろ、研究概念の導出について、導出ということばがまさにその姿勢を示すように、構成主義的 (constructionism) な認識をもっている³。他分野の研究者と同様に、観測が不可能な事象については、我々が、高々その振る舞いについての数理的近似を与えることができるのみであり、そのように我々が用意した数理的近似が、そのまま真理として実在するなどというような楽観的な思考をもたない。本論においても、その特徴が顕著であるが、数理的アプローチは、観測より得られる数理的構造が、理論上の概念や機構に先立つものであると考えている。つまり、所与の観測間における数理的構造それ自体と、それを生成するであろう、不可知のメカニズムの存在、ないし实在については否定しようのないものとして存在を受け入れる。しかしながら、そのメカニズムに対

² ここでの本質主義とは、自然現象、行動、社会のあり方などに関わるさまざまな事物には、客観的で固定的な、ないし斉一的な特性があるとする考え方を示す。数理的アプローチでは、ある社会的な関心となりうる観測下において、その観測に近似する数理モデルから効用を得られるということ自体を重要視している。本質、つまり事物の本質やこの世界の斉一的な規則を解明することは、このアプローチにおける研究の一義的な目的ではない。また、人間が自然言語を使用し、何かを理解したと思うことによって、個人の知的欲求を満たし、心的に充実することを殊更重要視するものでもない。これは、外国語教育研究が、元来、現実世界におけるある種の価値観をそのディシプリンに含む応用分野であるとの認識による。

³ 一般的に、構成主義ないし構築主義と呼ばれる姿勢を示しているが、社会学における社会構築主義的態度に近く、研究における諸概念は、研究者を含む人間の認識の仕方に依存しながら存在し、その事物が、その認識のまま独在することはない、と考える。つまり、明示的知識および暗示的知識といった概念は、研究者の認識によってのみ存在すると考えている。

して近似された数理的モデルや、その数理モデルを説明する理論上の概念や機構自体は、あくまでも人為的な、または人工的な構成物であると捉えている。その上で、理論上の概念や機構の存在については、本質的にはその真偽を問うことができず、数理モデルが現実に対してもつ近似度と、その構成された理論がもたらす帰結によって評価を与えることができるのみである、と考える。これは上記の帰結主義にも強く関連する。つまり、ほとんどの場合、数理的アプローチは、観測が不可能な事象について、科学的真偽や自然主義的実在論に積極的に関与することはない。

最後は、研究方法に関する思想的特徴である。数理的アプローチは、当該分野における共約不可能性（incommensurability）の問題⁴を非常に重く受け止めている。そして、当該分野における共約不可能性の問題は、形式主義と構造主義を採用することによって、もっとも効率的に回避することができると考えている。形式主義的なアプローチは、自然言語ではなく、むしろ、簡素な数式、論理、その他の記号的記法、または、ある種のコンピュータ言語や、データ記述言語といった人工言語の使用を好む。一方の構造主義とは、研究実践における概念と観測間の構造、概念間の構造、そして観測間の構造といったように、研究実践をメタ的に捉えるために、概念や観測が自己定義的に表す内容的含意、すなわち、理論⁵ではなく、概念や観測の構造自体に注目する考え方である。たとえば、ある観測間の構造を等しく予測する二つの理論的枠組みがあったとすれば、いかに自然言語によって異なる解釈や定義が与えられていたとしても、構造同型⁶であるという理由をもって、それら二つの理論を等価であるとみなす。それだけではなく、それらの理論がそれぞれもつ含意の差異を積極的に捨象することもある。これは俗にいうオッカムの剃刀⁷であるとか、モーガンの公準⁸といった姿勢にも通じる。

数理的アプローチの具体的な研究方法論について述べるとするならば、外国語教育に関する観測に対して、一次関数や非線形関数といった各種の関数、ポアソン分布やガンマ分布と言った確率分布、または、拡散過程やマルコフ過程といった非決定的な確率過程の類をフィットさせ、得られたモデルにもとづいて意思決定を行なうことによって効

⁴ 共約不可能性とは、用語、方法論、目的論などが一致しないさまざまなパラダイム間において、研究に関する概念の対応付けがされなくなる状態を示す。

⁵ 自然科学では、理論は数式やある種の記号体系によって与えられる。たとえば、熱力学の諸法則がそうである。しかしながら、外国語教育研究では、理論とされるものは、実際、自然言語であるところの用語による定義を示す場合が多い。ここでの理論とは、後者の例に類するものを示している。

⁶ ここでの構造とは、数理モデルを含む。

⁷ オッカムの剃刀は、思考節約の原理などとも呼ばれ、ある事象を同程度に説明するものであるならば、より少ない数の仮定をもつ説明を正当化する方針を示す。

⁸ モーガンの公準とは、より簡潔な心的能力の仮定のみによって説明できる現象に対して、より複雑で高次な心的能力による解釈を与えない、という比較心理学における方針である。本論の実証部分における主張は、この方針に沿っている。

用化 (utilize) する, という手続きが, その特徴をもっとも適切に示している (草薙, 2017b, 2017c)。また, ベイズ統計を積極的に受け入れており, 従来の主流であった統計的帰無仮説検定の類を研究実践において使用することは少ない⁹。

3.2 構造的仮定と因子分析モデルの適用に対する批判

さて, ここからは, 上記のような思想的特徴をもつ数理的アプローチに則って考察を進めていくこととする。まずは, 前章において述べた, 明示的知識および暗示的知識の二分法に関する構造的仮定についてである。R. Ellis らによる構造的仮定とは, 明示的知識および暗示的知識をそれぞれ潜在変数とみなし, それらの潜在変数に対応する観測変数として, 数種の課題成績を置くものである。前章でも述べたように, このような数理的な条件は, 因子数を 2 とした因子分析モデルに該当する。これは, 反映モデルでもある。

本論では, まず, 数理的アプローチの観点の上で, 上記の構造的仮定, 因子分析モデルの適用, そしてその上に成り立つ研究実践が, 以下の 6 つの問題点を抱えていることを指摘する。これらの問題点を考慮すると, 構造的仮定と因子分析モデルの適用による, 明示的知識および暗示的知識の二分法には, まったく信憑性がないということを主張する。以降, それぞれの問題点について個別に検討する。

- (a) 因子分析モデルと課題の分類
- (b) モデル構築の自由度
- (c) 推定因子の解釈
- (d) 母集団間における測定不変性
- (e) 対象項目間における測定不変性
- (f) 反映モデルと形成モデル

(a) 因子分析モデルと課題の分類

因子分析モデル¹⁰は, 任意の集団, 任意の観測変数間における分散共分散構造, そして任意の潜在変数の数, といった種々の条件を前提とした分析手法である。R. Ellis (2005) が報告するように, この研究における母集団内において, この研究で用いられた観測変数 (時間制限を設ける GJT, 時間制限を設けない GJT, メタ言語テスト, 口頭模倣テスト, 口頭物語テスト) の構造が, 二つの潜在変数によって説明されることは, 少なくとも事実であろう。本論では, このように, 文法運用に関する任意の観測変数のセットが, 因子分析モデル上において二因子構造をもつ場合が存在すること自体は否定しない。

しかしながら, R. Ellis (2005) などの研究において繰り返し述べられた, ある観測変

⁹ 本論に関わるところの確率過程やベイズ統計については, 次章を参照のこと。

¹⁰ 本論では, 因子分析を手法のひとつではなく, あくまでも一種の数理モデルであると捉え, 因子分析モデルと呼ぶ。

数が、一意的に、そして相互排他的に、明示的知識および暗示的知識の測定具である、といった見方は、端的に言って誤りである。たとえば、R. Ellis and Loewen (2007) が述べる以下の内容は、少なくとも現在の観点からみると厳密なものではない。

Ellis (2005) and his coresearchers developed a number of tests with a view to providing relatively separate measures of explicit and implicit knowledge.

R. Ellis and Loewen (2007, p. 119)

また、R. Ellis (2005) の以下の表現も同様である。

The imitation test and the oral narrative test were predicted to measure implicit knowledge because the participants would rely predominantly on feel, they would be under pressure to perform in real time, they would be focused primarily on meaning, and they would have no reason to access their metalanguage. In contrast, the metalinguistic knowledge test was predicted to measure explicit knowledge because it involved a high degree of awareness, was unpressured, focused attention on form, and, obviously, required the use of metalinguistic knowledge.

R. Ellis (2005, p. 157)

ここでの問題点は、上記の引用に見られるように、ある程度の合理的予測を伴って、ある任意の潜在変数に対して高い因子負荷量をもつであろう観測変数の値が、潜在変数を直接的に測定する、というような奇妙な解釈を与えている点である。または、因子分析モデルが、ある観測変数の測定概念を決定ないし分類しているかのように説明されている点である。つまり、課題 A は明示的知識を測定し、課題 B は暗示的知識を測定する、といった表現、それ自体である。これは因子分析モデルの適用における根本的な誤りである。

基本的に、任意の測定モデルによって推定される潜在変数の値は、その測定モデルがもつ因子構造と、すべての観測変数の実測値より推定されるものに過ぎないのであり、潜在変数の推定値とひとつの観測変数の値は、せいぜいが共分散をもつにすぎない関係である。やや極端かつ冗長な説明になるが、たとえば、一般的な因子モデルの数式は、(3.1) のように記述することができる。ここでの Y は観測変数、 a は因子負荷量、 f は共通因子、 e は独自因子である。潜在変数の数は 2、観測変数の数を 4 と仮定している。

$$Y_1 = a_{11}f_1 + a_{12}f_2 + e_1 \quad (3.1.a)$$

$$Y_2 = a_{21}f_1 + a_{22}f_2 + e_2 \quad (3.1.b)$$

$$Y_3 = a_{31}f_1 + a_{32}f_2 + e_3 \quad (3.1.c)$$

$$Y_4 = a_{41}f_1 + a_{42}f_2 + e_4 \quad (3.1.d)$$

そこで、(3.1.a) について、ある任意の観測変数の値 (Y_i) が、ある共通因子 (f_j) と同じ値を取るというために必要な、もっとも整合性の高い数理的条件のひとつは、(3.2) の通りである¹¹。

$$[a_{12} = 0] \wedge [e_1 = 0] \quad (3.2)$$

共通因子の負荷量についての条件ならまだしも、独自因子の値が 0 であるという仮定は、著しく不可解なものになってしまう。このように、測定モデルにおいてある種の因子構造が得られたからといって、ひとつの観測変数と潜在変数の値が等しいと結論づけることは原理的にできない。因子分析モデルは、任意の一観測変数に対する、その測定概念の判別機ではなく、あくまでも、全ての観測変数の値と得られた測定モデルを使用して、潜在変数の値を推定するモデルである。同時に、所与の因子構造が、ある程度の単純構造を示したとしても、ある観測変数がひとつの潜在変数の影響のみを排他的に受けている、といった議論をすることは基本的に不可能である。これは、ある質問紙の一項目だけを任意に取り出して、その項目のみによって構成概念を直接的に測定することが不可能であることと同じである。

このような因子分析モデルに関する根本的な誤りの原因は、2000 年代の第二言語習得研究において、心理統計についての理解が未熟であったことだと考えられる (e.g., Isemonger, 2007)。事実、2000 年代当時、文法知識の測定を対象とするような古典的な第二言語習得研究において、因子分析モデルは手法上画期的なものであり、このような誤りが起きることは十分にもっともらしいことであった。さらに、当時の第二言語習得研究には、「構成概念の定義」といった用語の頻用に見られるように、やや過激な操作主義と認知主義的思想が、論壇上にて共存していたように見える。もちろん、このような表現は、現在の心理学においては、好ましくないものとされている (Lilienfeld et al., 2015)。このことを表すように、たとえば、Hulstijn (2005) は、当初より R. Ellis (2005) における構成概念の定義や操作化の手続きといったもの自体については、認知主義的な観点より、一定の批判的な姿勢を示している。しかし、同時に今後の研究の進展を促すものとしてその貢献を認めてもいる。

¹¹ 論理積を \wedge と記す。論理積は命題がともに真であるときのみ、真を返す演算子である。

(b) モデル構築の自由度

上記の点に関連するが、因子分析モデルによって得られる因子数や、その因子数の下での因子構造は、因子分析モデルに投入する観測変数の数によって大きく変化する。たとえば、R. Ellis らが報告した因子分析モデルに対して、観測変数を新たに投入、または除外するとしたときに、異なる因子構造、たとえば三因子構造が比較的高い適合度を示す可能性は非常に高い。もし、その観測において、三因子構造が得られるのであれば、明示的知識および暗示的知識の二分法は破綻してしまう。また、Isemonger (2007) が十分に指摘したように、因子数の決定方法にも研究者の自由度が残る。前章で述べたように、構造の優先、つまり、理論的に仮定された二因子構造から、それに見合う基準を幅広く導入した上で、事後的に観測変数やモデルを選択したという手続き自体に問題があったと考えられる。

仮に、明示的知識および暗示的知識なるものが実在し、そしてそれ以外の潜在変数が存在せず、排他的に、または混成的に、あらゆる種類の文法運用に影響を及ぼしているのだとしたら、どのような観測変数を足したところで、本質的には二因子構造が期待されるべきである。因子分析において、ある潜在変数の、ある観測変数に対する負荷量の値は、他の観測変数に対する負荷量と、理論的には独立である。しかしながら、そのような検証結果が実際に得られることは稀である。なぜならば、潜在変数の推定値自体が任意のセットの観測変数に由来するものだからである。繰り返しになるが、因子分析モデルは、ある観測変数のセットという条件下において、それらの観測変数間の構造を説明する数理モデルのひとつに過ぎないと考えるのが妥当である。

たとえば、十分な数の観測変数から、投入変数がある程度探索的に決定できるという条件下で、目標とする因子数、たとえば二因子構造を得ること自体は、比較的容易である。逆にこのことは、明示的知識および暗示的知識の二分法における実証的証拠としての価値付けを疑わせるものになる。本来であれば、運用におけるさまざまな観測変数を網羅的に取り入れたモデルか、または観測変数の選択自体によって高い公共性や代表性が満たされるモデルを構築すべきであったといえよう。

(c) 推定因子の解釈

因子分析モデルによって、ある観測変数のセットにおいて二因子構造が優れた適合度を示すことと、それら推定された潜在変数が、測定を目指す構成概念自体に相当するかどうかは、基本的にはまったく別の問題である。後者は、一般に妥当性として表現される性質の一部に近い。

このことは、Isemonger (2007) が指摘したことに通じるが、得られた潜在変数の解釈について、R. Ellis (2005) の研究には理論先取りの傾向が見て取れ、妥当性の手続きについての関心が低いように見受けられる。たとえば、彼が主張する明示的知識および暗

示的知識のモデルに対する対立的なモデルの実証的比較は一切行われておらず、理論的重要性に見合うほどの、実証的に厳格な手続きはなされていない。これは、一般的な心理統計の手続きと比較すると、非常に独創的である。通常であれば、対立仮説を検討した上で、それを反駁するという手続きを取る。

このこと具体例として、Isemonger (2007) は、明示的知識および暗示的知識のモデルに対する対立的なモデルとして、決定と産出という二因子モデルも理論的に想定できるということを主張している。ただし、このモデルについては、実際に観測に対する適合度が低いと R. Ellis and Loewen (2007) によって反論されている。また、Suzuki (2017) は、自動化された明示的知識と暗示的知識という二分法を主張しているが、R. Ellis (2005) において、暗示的知識の測定具とされた時間制限を設ける GJT が、自動化された明示的知識という潜在変数に対して高い負荷量を示したことを報告している。ただし、Suzuki の研究における観測変数は、R. Ellis (2005) における観測変数と全く異なるセットである。先に述べたように、因子分析は所与の観測変数のセットを条件としたモデルにすぎないことには注意が必要である。観測変数のセットが異なれば、いうまでもなく得られる因子構造が変化するため、本質的に、たとえば、時間制限を設ける GJT が、どのような潜在変数に高い因子負荷量をもつか、といったことは、Suzuki の研究からもわかるように、普遍的なものではない。これは、R. Ellis (2005) に見られる明示的知識および暗示的知識の二分法に関する基準とは独立した問題である。

さらに、メタ言語テスト、時間制限を設ける GJT、時間制限を設けない GJT や各種の口頭課題などがもつ課題特徴の対立は、非常に多岐に渡り、列挙しきれものではない。たとえば、時間制限を設ける GJT、時間制限を設けない GJT、メタ言語テストは、明らかに判断行動を伴い、その他の課題はそうではない。各種の口頭課題は、聴解やスピーキングに関するものであると同時に、GJT などは、読解や語彙に関わるものと捉えることもできる。このように、一般的に下位技能と呼ばれる他の構成概念との関連性も十分に考慮すべきである。

特に表象主義に基づく研究の場合、知識表象について推論するための手がかりとして、何かしらの判断課題における判断行動に着目することは、もっとも一般的な研究アプローチである。判断行動への注目は、人口言語学習、再認記憶研究、言語研究、そして従来の第二言語習得研究におけるスタンダードといってもよい。一方、口頭課題の成績については、通常、知識ではなく、技能や能力といった枠組みに関連付けられる場合が多い。このこと理由は、判断課題では、刺激文や実験条件を統制することによって、研究者が関心をもつ知識表象に関する特有の反応を観測できるからである。一方、口頭課題は、パフォーマンス評価などの枠組みに顕著なように、より現実に近い状況での行動を観察するものである。よって、R. Ellis (2005) によって暗示的知識の測定具であると主張された課題についても、一般的な知識表象の実験手続きから見てもかなり特殊であ

るため、測定する構成概念が、他の何かしらの技能ではなく、知識の一種であるという主張はそれほど整合性の高いものではない。このことの是非についての議論は置いておくとしても、推定因子の解釈は、ただ一つに求まるものではないという理解は重要である。いずれにせよ、R. Ellis (2005) における問題点は、潜在変数であるところの明示的知識および暗示的知識が、排他的ないし混成的¹²に、いかに観測変数であるところの課題成績に影響を及ぼしているか、その因果モデル自体が十分に論じられていない点であるといえよう。

たとえば、R. Ellis (2005) の一部を再掲する。

The imitation test and the oral narrative test were predicted to measure implicit knowledge because the participants would rely predominantly on feel, they would be under pressure to perform in real time, they would be focused primarily on meaning, and they would have no reason to access their meta language. In contrast, the metalinguistic knowledge test was predicted to measure explicit knowledge because it involved a high degree of awareness, was unpressured, focused attention on form, and, obviously, required the use of metalinguistic knowledge.

R. Ellis (2005, p. 157)

上記の引用では、口頭課題は直観に依拠するであるとか、またはメタ言語に対するアクセスを要しないとといった予測が合理的になされているが、これらの予測に関わる実証が十分でないだけでなく、合理的にもっともらしいとも判断しかねる。たとえば、前章で述べたように、認知心理学における暗示的知識は、一般的に運用面において低い正当率や長い反応時間などに特徴づけられている。また、即時的 (real-time) に高い成績を示すのは明示的知識であるとされている。前章でも述べたように、R. Ellis の説明は、少なくとも一般的に認知心理学に見られる予測と完全に矛盾している。もしも、この直観や意識といった用語が、認知心理学における意味ではなく、R. Ellis が意図するような特殊な意味で使われているのであったならば、その定義をより厳密に示すべきであったろう。また、認知心理学の用語を借りるとするならば、これらの課題が引き起こすであろう認知的処理は、課題に特有 (process-pure) ではない。

このように、明示的知識および暗示的知識、そして、さまざまな課題特性やそれらより引き出される運用の関係については、彼の主張の大きさに見合うほど実証性に優れるものではない。つまり、実際に推定された二因子が、彼が述べるような意味での明示的知識および暗示的知識であるかについては、疑念が残るといわざるをえない。

¹² 混成的 (heterogeneous) とは、異種であるところの複数のものが混ざり合う性質を示す。ここでは、明示的知識および暗示的知識が、ともに観測、つまりパフォーマンスに影響を及ぼすことをいう。

(d) 母集団間における測定不変性

測定不変性 (measurement invariance) と構造不変性 (structural invariance) は、測定論的特性 (measurement property) の中でも重要な要素である。たとえば、ある測定モデルにおいて、男性群と女性群の間において因子構造が大きく異なれば、男性群において推定される潜在変数の値と、女性群において推定される潜在変数の値は、これらの群間において比較可能なものにはなりえない。測定不変性の分類にもさまざまなものがあるが、Vandenberg and Lance (2000) の分類によると、(a) configural invariance, (b) metric invariance, (c) scalar invariance, (d) residual invariance がある。これらはそれぞれ、因子数、因子負荷量、切片、誤差分散の等値制約に結びついている。また同様に、Vandenberg and Lance によると、構造不変性には、(a) factor variance invariance, (b) factor covariance invariance, (c) mean invariance などがある¹³。

ここでの関心は、当該の構造的仮定における測定不変性についてである。R. Ellis (2005) は、母語話者と学習者を同じ被験者群としてプールしているが、母語話者と学習者間の測定不変性については、現在のところ未検証のままである。もちろん、明らかに異質の母集団が混在する標本による尺度構成は好ましいものではない。

また、完全な測定不変性が満たされるべきなのか、そうでなければ、上記のどの水準の測定不変性が満たされるべきなのか、といった点は理論的にも不明瞭である。統計的な観点からいえば、母語話者の群内において、各種の課題が示す分散および共分散の値が学習者群のそれよりも遥かに小さいであろうという予測からいって、配置不変性 (configural model) が満たされるかさえ不明であり、それ以上の水準の測定不変性や構造不変性が満たされることは期待できない。

母語話者群と学習者群の間において、測定不変性が満たされないのであれば、これは第二言語習得研究における従来のさまざまな知見と矛盾することになる。たとえば、測定不変性が成り立たないという条件下において、母語話者がもつ暗示的知識と学習者がもつ暗示的知識は、そもそも比較不可能である。これでは、明示的知識と暗示的知識、そして学習者と母語話者という類推関係は破綻してしまう。すなわち、この構造的仮定では、学習者の個人間分散をある程度説明できたとしても、学習者と母語話者に見られる運用の差を説明することが、そもそもできないのである。この点については、次節にてさらに詳しく述べる。

学習者と母語話者に見られる運用の差を、この構造的仮定によって説明するためには、測定不変性が厳しい水準をもって満たされ、さらに構造不変性のみが満たされない、という条件を期待しなければならない。つまり、学習者群と母語話者群の間において、構

¹³ 心理統計において、測定不変性とは、ある種のカテゴリー変数間において、因子分析モデルといった数理モデル上の係数が一致している状態を示す。つまり、群間において測定は等値であるということである。構造不変性とは、それに加え、因子平均といった潜在変数の値自体が等値であることを示す。これらの具体例が本文のとおりである。

造的仮定において因子構造が等しく、そして因子分散、因子共分散、そして因子平均のみが自由母数¹⁴でなければならない。具体的には、これまでの第二言語習得研究に見られる主張のように、明示的知識に関しては、母語話者群が示す因子平均が著しく低い値を取り、学習者群が示す値が著しく高い値を取るであるとか、または暗示的知識に関して、母語話者群が示す値が著しく高い値を取り、学習者群が示す値が著しく低い値を取る、といった具合である。因子分散についても同様に、一般に、学習者の運用は母語話者のそれよりも高い散布度を示すことが知られているのだから、学習者群の分散は母語話者群のそれよりも高い値を示すべきである。もちろん、このような観測が得られる見込みは限りなく低い。たとえば、Ellis らが明示的知識の測定具として示した時間制限を設ける GJT の成績のみを比較したとしても、母語話者群が学習者群の成績を圧倒する確率が極めて高いことを考えれば、この仮定に関する期待は非常に薄いものであることがわかる。このように、従来の第二言語習得研究に見られた学習者と母語話者における運用の差を、R. Ellis などが主張するところの明示的知識および暗示的知識および当該の構造的仮定によって説明することは、不可能なものになってしまう。少なくとも、学習者を対象として検証された測定モデルにおける暗示的知識の推定値は、母語話者を母語話者たらしめるようなものではないこと、またはその明示的知識の推定値が学習者を学習者たらしめるようなものでないこと、この二つは確実であろう。

(d) 対象項目間における測定不変性

母集団のみならず、文法知識が参照するドメイン、つまり、さまざまな言語機能や言語規則、さらにいうなら、それぞれの知識表象に関する測定不変性も重要な観点である。たとえば、R. Ellis (2005) は、17 の文法項目を対象としてテストバッテリーを構築したが、ここでの主張は、当該の 17 の文法項目全体において、構造的仮定がフィットしていることを示しているのにすぎない。よって、厳密に言えば、異種の文法項目を対象とした研究の場合、R. Ellis (2005) のモデルにおける測定不変性が満たされるかどうかは、基本的に不明である。また、より微細な観点に着目するならば、それぞれの文法項目の関係についてであっても、明示的知識および暗示的知識という構造の測定不変性が成立しているかどうか未検証のままである。

ここで、Kusanagi and Yamashita (2013), Tamura and Kusanagi (2015a, 2015b) などが示している課題構造交互作用 (task-structure interaction) という現象について考えたい。課題構造交互作用は、その名の通り、ある運用を示す成績における、課題の種類と刺激の構造がもつ統計的交互作用である。そのもっとも典型的な研究である Kusanagi and Yamashita は、GJT における時間制限という課題要因と、二種の構造、つまり、刺激がもつ言語的要因 (数の一致の種類、形態的規則性) がもつ交互作用を検証した。その結果、

¹⁴ ある数理モデルにおいて、独自の値を取ることができる母数を示す。ここでは、等値制約から独立した推定母数である。

時間制限という課題要因は、形態的規則性のみと交互作用をもつということを示している。

このような課題構造交互作用は、対象項目間における測定不変性が成り立たないという予測の根拠になる。GJT に対する時間制限によって、相対的に成績の変動が発生しない項目を対象とする場合、時間制限を設ける GJT と時間制限を設けない GJT の成績が示す共分散の値は高くなるため、二因子構造が再現される可能性は低くなるからである。このような現象を考えると、当該の測定モデルが、実に複雑で不安定であることがわかる。さらに、草薙 (2013) や Kusanagi (2014) が指摘するように、より細かな課題条件、たとえば時間制限の設定方法によっても、時間制限を設ける GJT と時間制限を設けない GJT の共分散が劇的に変化するという点にも注意が必要である。

(e) 反映モデルと形成モデル

R. Ellis (2005) といった 10 年以上前の研究実践に見られる統計上の技術的問題点を逐一指摘することは、必ずしも建設的な研究態度ではないことは筆者も理解している。ただし、そのような理解の上でも、敢えて指摘しておかなければならない技術的問題点がまだ残っている。

それは、R. Ellis (2005) や一連の研究において、因子分析モデルと主成分分析モデルを混同しているような表現が多数見受けられることである。R. Ellis (2005) において、実際に行われたのは、因子分析ではなく、主成分分析であると予測されるが、"*principal component factor analysis*" (R. Ellis, 2005, p. 162) というような奇妙な表現や、論文中のセクションによって、因子 (factor) や成分 (component) という用語が代替可能なように使用されている様子が見受けられ、相当に誤解を招きかねないことが見て取れる。これは、技術的問題点のみに限らず、理論的にも大きな問題にもなりかねない。

繰り返しになるが、潜在変数モデルにおいて、因子分析モデルは一種の反映モデルに相当する。反映モデルにおいて、潜在変数は観測変数の因果である。一方、主成分分析モデルは形成モデルに相当する。形成モデルにおいて、潜在変数は観測変数の結果である。潜在変数モデルに関する一般的な枠組みからみて、明示的知識や暗示的知識というような、ある認知的機構の存在を根拠とするような構成概念については、反映モデルを適用することが妥当である。これは、たとえば明示的知識であれば、本論で繰り返し述べているように、明示的知識を、観測に影響を及ぼす原因であると考えるか、そうではなく、課題成績から人工的に構成した概念であると考えるかといった違いとなってしまう。

しかしながら、本論では、これまでも、そしてこれからも、R. Ellis とその共同研究者は、あくまでも反映モデルを想定して分析を進めたものとして解釈している。悪くいえば理論先取の、よくいえば検証的性格が見られる当該の研究において、反映モデルの方

が、比較的その研究姿勢の上で適切である、または研究姿勢において一貫した選択であると総合的に判断した。

また、R. Ellis (2005) について、Isemonger (2007) は、おそらく好意的に主成分分析を因子分析と読み替えた上でのことだと推察されるが、探索的因子分析ではなく、検証的因子分析を使うべきであったと主張した。この Isemonger の主張を、本論は全面的に支持するが、いずれにせよ、これを受けて R. Ellis and Loewen (2007) は、検証的因子分析を使用して、当初の結論と同様の結論が得られることを主張している。ただし、この研究において、再度主成分分析モデルを使用した分析も同時に報告されているが、その結果を表すパス図は、明らかに主成分分析モデルのものではなく、因子分析モデルのものである。さらに、同論文内で報告された検証的因子分析は、主成分分析モデルではなく因子分析モデルを適用していた。それに加え、主成分分析モデルを使用するほうが適切だという見方の上で、主成分分析と因子分析の両方を行ったところ、因子構造に差がないという報告も注に見られる (p. 126)。実際のところ、少なくとも当時の多変量解析の技術では、主成分分析モデルと因子分析モデルの選択は、非常に難しいものであったと推察できる。

ただし、仮に主成分分析モデル、すなわち形成モデルの使用を想定していたのであれば、明示的知識および暗示的知識は、あくまでも研究者が人工的に作成した概念であると考えざるを得なくなり、理論のあり方が根本的に変容し、重大な誤解を招きかねない。この点が本質的には不明だという点も、問題点のひとつとして挙げることができる。

以上の問題点からわかるように、R. Ellis らによる構造的仮定と、その因子分析による検討は、総合的に判断して、R. Ellis らの主張を満たすものだとはいえない。なお、繰り返しになるが、R. Ellis (2005) によって使用された観測変数のセットが、二因子構造を示すこと、そして測定モデルによって推定された潜在変数を、人工的に作成した概念として明示的および暗示的知識と呼ぶことについて反論するつもりはない。また、最後の件に関連するが、一連の研究が、形成モデルによって、研究者間の合意形成を図り、Hulstijn (2005) が述べるように構成概念を課題によって定義したり、課題によって操作化したりするといった活動を目的とするものであったのであれば、それは過激な操作主義であるという評価は免れないものの、本論の批判の多くには当たらない。過激な操作主義は確かに問題であるが、本論が批判しているのは、上記の構造的仮定とその検証方法、さらにそこに置かれた理論的仮定と実際の観測における数理的構造が整合性に欠くこと自体である。

3.3 明示的知識および暗示的知識と観測の数理的構造

このように、構造的仮定とその検証方法に多くの問題点が見られるということの主たる原因は、R. Ellis を始めとする第二言語習得研究者が、さまざまな現象に関する数理的構造を極端に簡素化し、その数理的構造を説明する方法を、実証性の高い方法に依拠するわけではなく、主に合理的推論のみによって求めたことであろう。これは、Hulstijn (2005) が述べるところの、「極端な合理主義」(extreme rationalism) といった態度に近いと本論は見ている (e.g., Isemonger, 2007)。一方、数理的アプローチを標榜する本論は、「極端な経験主義」(extreme empiricism) といった態度をもつと認識されても、差し支えない。

さて、この節では、前章から本章まで述べてきた研究史上の概念や、それに対する批判点を総合的に振り返りながら、明示的知識および暗示的知識の二分法がもつ問題点を、数理的アプローチの観点から、再度総括することとする。数理的アプローチでは、先述の通り、理論の構築に先立って、観測間の構造に着目する。

前章からの繰り返しになるが、第二言語習得研究は、分野固有の関心のひとつとして、学習者と母語話者という集団間における運用の差に長い間着目してきた (Hulstijn, 2005)。これは、NS を母語話者、NNS を学習者としたとき、ある運用の期待値 (E) において、

$$E[NS] > E[NNS] \quad (3.3)$$

また、その散布度 (V) に着目して、

$$V[NS] < V[NNS] \quad (3.4)$$

というように表すことができよう。これは、母語話者における運用の期待値は学習者よりも高く、学習者における運用の分散は、母語話者のそれよりも大きいということである。このような数理的条件は、やがて理論的に、学習者がもつ中間言語と、母語話者の文法知識という二項対立によって具体的な説明が与えられた。帰属理論における共変モデルの見方を援用するならば、集団における一貫性が高く、集団に帰属するものとして認識されているわけである。

つまり、上記の条件は、それぞれの集団に対応する、中間言語と母語話者の文法知識がもつ特性による結果であり、運用における期待値の差、分散の相対的大小関係は、学習者 vs. 母語話者という二項対立の具体的特徴でもあった。このような二項対立は、知識の獲得過程に関する知見を媒介して、明示的知識および暗示的知識の二分法と同一視されることもあった (e.g., Krashen, 1981)。すなわち、以下のような関係である；暗示的知識は明示的知識よりも高い成績を示す、明示的知識は暗示的知識よりも成績において大きい散布度を示す。ここでの暗示的知識は、あくまでも母語話者という集団に対して特

異的な行動の因子，といったような含意である。

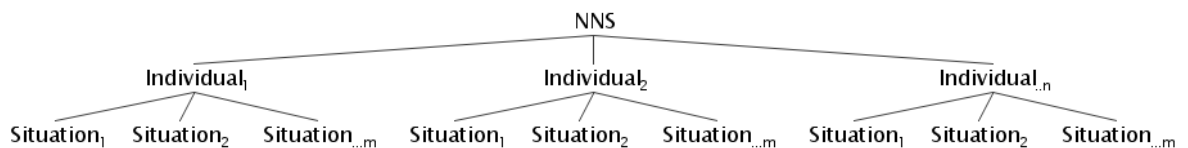
$$E[Implicit] > E[Explicit] \quad (3.5)$$

$$V[Implicit] < V[Explicit] \quad (3.6)$$

しかしながら，新たな言語的課題が複数開発され，個人内における課題や条件間の差が頻繁に観測されるようになった1970年代からは，集団間の期待値や分散ではなく，個人内分散に対して研究者の焦点が移っていった (e.g., Larsen-Freeman, 1975)。帰属理論によって解釈を与えるならば，まずは集団レベルを捨象し，個人内分散の大きさ，つまり，一貫性の低さによって，個人から状況へ原因を帰属するように認識をシフトしたわけである。これは，中間言語可変性研究において顕著であったように，個人内分散の値が個人間分散を圧倒するというような見方にほかならない。

ここで，個人に対してさまざまな条件や課題における成績がネストする階層構造 (3.7) を考えてみる。

(3.7)



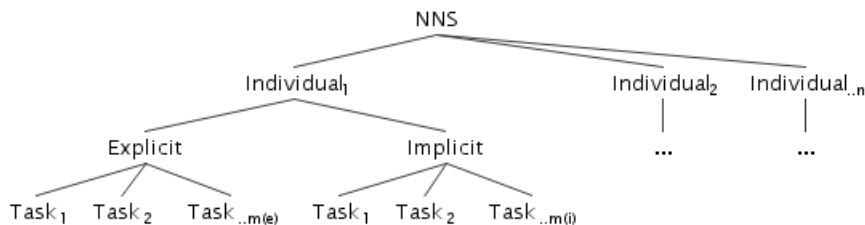
このような階層構造の下で，個人内分散 σ_{within}^2 と，個人間分散 $\sigma_{between}^2$ がもつ条件は，以下のように表すことができる。これは分散成分の比でもある。

$$\sigma_{between}^2 < \sigma_{within}^2 \quad (3.8)$$

このような数理的条件に対して関心が移った際，個人間分散は発達段階ないし学習者要因の影響としてみなされ，個人内分散については，二つの種類の知識，たとえば明示的知識と暗示的知識といった機構の存在と，それらに対してネストする課題によって説明されるようになった。つまり，個人内分散を体系的に説明するために，さまざまな課題がネストする二つの変数を導入したということである (e.g., Han & R. Ellis, 1998; R. Ellis, 2005)。これは，階層構造 (3.9) のように表すことができ，構造的仮定と本論が呼ぶものと相似している。繰り返しになるが，この構造の上で，明示的知識および暗示的知識は，それぞれ異なる課題によってネストされている。このことは，前節で詳しく述べたように，当該の一連の研究 (R. Ellis et al., 2009) が，因子分析を，課題が測定する構成概念を

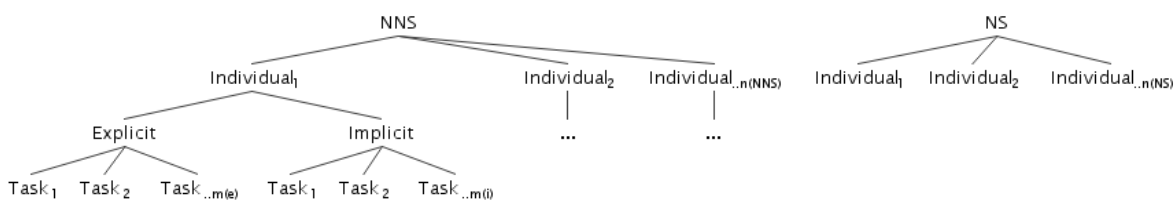
判定できるようにみなしていることから明白である。つまり、ここでは、明示的知識と暗示的知識は、さまざまな課題の親カテゴリーとして考えられると同時に、個人から見て子カテゴリーでもあり、個人がもつ特性でもある¹⁵。

(3.9)



構造的仮定、つまり任意の観測変数の条件下で、二因子構造が優れた適合度を見せること自体は、十分にもっともらしい。ただし、このとき、本論がこれまで繰り返し述べたように、明示的知識や暗示的知識と呼ばれるものが、R. Ellis (2005) が主張するように、認知心理学における明示的知識と暗示的知識と等しいという根拠はまったくない。これは本論が再三に渡って構造の優越と呼ぶ問題である。同時に、ここで仮定されている暗示的知識が、認知的機能において、母語話者の運用を支える機構、つまり母語話者という集団に対する特異性を与えるものであるとの根拠はない。階層構造 (3.10) が示すように、母語話者群が示す期待値や分散と、暗示的知識という級内における課題の期待値や分散は、本質的にまったく異なるスケールのものである。

(3.10)

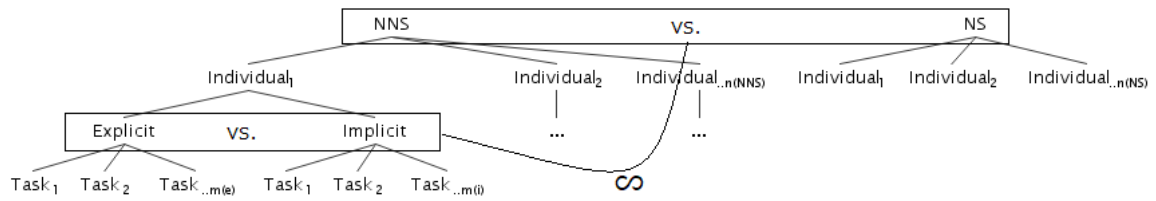


本論は、R. Ellis らによる明示的知識および暗示的知識という二分法の、そのもっとも顕著な特徴をここに見る。すなわち、集団間における期待値や分散の差と、個人内における期待値や分散の差を、同一の理論的概念で説明しようとしていることである。または、集団に対して特異的な行動の因子を、個人内の現象に対して構造的に投影している点である。このことは、(3.11) が示すような、構造の相似¹⁶を考えれば理解しやすい。

¹⁵ 純粋に状況に依存する行動は、帰属理論からみて、一致性と弁別性が高く、一貫性が低いという状況に由来する。

¹⁶ 相似とは、形、構造などが似ているさまを示し、のなどと記す。まず、母語話者および学習者間に見られる観測の差を説明

(3.11)



つまり、その観測の構造に着目する限り、明示的知識および暗示的知識は、学習者と母語話者という集団間の差を説明するものであると同時に、学習者における個人内分散をも説明するものでもあるということになる。このような見方の類例には、男性の中における女性性であるアニマ、そして女性の中にある男性性であるアニムスといった有名な概念がある。少なくともその数理的構造の上で、男性と女性という集団間の特徴の差と、それぞれの性をもつ個人における状況間での行動の差（男性という集団に対して特異的である行動 vs. 女性という集団に対して特異的である行動）を相似させている点において、明示的知識および暗示的知識の二分法に対するある程度の類似性を指摘することができる。集団に対して特異的な行動や、その因子の対立を、個人内に投影する見方は上記のように類例が豊富であるものの、少なくとも明示的知識および暗示的知識の二分法に関しては、構造的仮定の上で明らかに矛盾している。二つの構成概念、つまり明示的知識および暗示的知識を個人に帰属させ、それらに対して課題をネストさせる構造は、仮に個人内分散を適切に捉えることができるとしても、学習者と母語話者という集団間の差についても同時に説明するものにはなりえない。

さらに、前節にて詳しく述べたように、以下のような問題点が残されている。第一に、少なくとも因子分析モデルによっては、課題が単純に構成概念にネストしているとはみなせないことである。第二に、課題間分散と集団間の差という全く異なるスケール、そして全く異なる水準の事象を、直接的に比較することができない点である。第三に、母語話者と学習者間において、測定不変性が成り立たない、または測定不変性を仮定しないという点である。第四に、次節にて述べるように、能力と課題の困難度を混同していることである。

するものであった明示的および暗示的知識は、集団の期待値ならびに個人間の分散に着目したものであった。ここでは、学習者の運用を明示的知識によるものとし、母語話者の運用を暗示的知識によるものと帰属させている。一方、中間言語可変性研究以後、特に R. Ellis らによる研究では、まさに中間言語可変性という用語が示しているように、個人であるところの学習者内において、状況によって変化する運用に着目している。これは、集団や個人ではなく、状況に観測の原因を帰属させているということになる。ここでの明示的知識は、ある状況（e.g., 時間制限のない状況）に付随する観測の原因であり、暗示的知識は、別の状況（e.g., 時間制限のある状況）に付随する観測の原因である。しかし、これら二つのまったく異なる説明の背景にある数理的構造は、(3.11) が示すように相似している。

3.4 課題の困難度および速さと正確さの二律背反

ここでは、最後の点、つまり、能力と課題の困難度を混同しているという点について詳しく掘り下げたい。本論では、この問題点の解決が、明示的知識および暗示的知識の二分法に対する代替モデルの提案に対する鍵であると考えている。

さて、集団間の差、すなわち、暗示的知識に特徴づけられている母語話者群の運用が、学習者群の運用を圧倒するという事実によって、同時に学習者における個人内分散を説明することを可能にするには、暗示的知識にネストしている課題の成績が、明示的知識にネストしている課題の成績を同様に圧倒するものでなければならない。すなわち、個人内においても、暗示的知識は、明示的知識よりも高いレベルの運用の原因になるべきである¹⁷。もちろん、明示的知識および暗示的知識に対して、それぞれネストしている課題は異なるのであるから、単純な変数の値の比較は、事実上不可能である。しかしながら、ここでの値の比較は、意味を成さないため、暗示的知識の方が高いレベルの運用を示すということを当該の構造上にて表現するように、暗示的知識を測定すると主張されている課題は、どの課題を取り上げても、課題の困難度が非常に高いものとなっている。

もっとも顕著な例は、時間制限を設ける GJT と時間制限を設けない GJT の対比である。あらゆる種類の判断課題において、時間制限は課題成績の値を下げる、つまり課題の困難度を上げる要因になることが知られている。このことは速さと正確さの二律背反 (Speed-Accuracy Tradeoff, SAT; e.g., Heitz, 2014; Luce, 1986; Wickelgren, 1977) としてよく知られている現象で説明できる。SAT とは、判断課題の成績において、反応速度と正確さに二律背反関係が見られることを示す。

その分かりやすい例は、課題指示の効果についてである。ある実験協力者が判断課題に従事するとして、実験者が実験協力者に対して、できるだけ早く反応するように指示を与えると、できるだけ正確に回答するように指示を与えた条件に比べ、相対的に不正確だが、短い反応時間が得られる。このような反応速度と正確さの共変関係から、正確さを反応速度の関数であると捉える。これを SAT 関数 (SATF) という (e.g., Heitz, 2014; Wickelgren, 1977)。SAT に関する研究は分野横断的に行われており、SAT 現象を引き出すための実験方法として、上記の課題指示のみならず、報酬法 (payoff)、時間制限法 (deadline)、反応信号法 (response signal) などが知られている。時間制限法は、R. Ellis (2005) における時間制限を設ける GJT に見られる手法でもある。図 3.1 に、SAT 現象を簡略的に示すイメージを示す。

¹⁷ 繰り返し述べてきたように、暗示的知識とは、母語話者に特徴的な言語運用の原因であるならば、暗示的知識による運用は、より高いパフォーマンスを示すべきである。

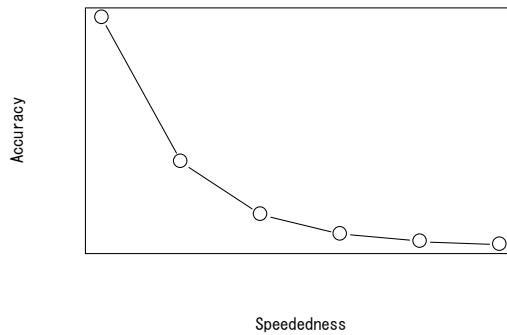


図 3.1 SAT 現象を示すイメージ

個人内における SAT 現象は、本質的には個人がもつ能力とは独立した概念である。たとえば、母語話者を対象にした GJT の判断成績において、時間制限を適切に設けた条件の成績が、時間制限を設けない条件の成績よりも低くなるのは当然である。さらに一般的な例を出すと、野球のバッティングにおいて、球速が上がれば上がるほど、バットがボールに当たる確率は必然的に低くなる。ある球速、たとえば 100km/h のボールを安定的に打つことができる個人が、160km/h のボールに関して同様のパフォーマンスを発揮できないとする。この観測をもって、あるひとりの個人について、高速の球を打つ能力と、比較的低速の球を打つ能力の二つが実在していると考えるのは早計である。そしてそれは、プロ野球選手が高速の球を打つことができ、アマチュアがそれができないという集団の差が見られるということとも、基本的に独立である。

むしろ、課題成績の値は、能力を項にもつある関数によって与えられている、といった理解の方が適切である。たとえば、潜在変数モデルの一種である項目反応理論 (item response theory) がその本質をよく表すように、同一の能力をもつもの同士において、課題成績の差が示すものは、課題の困難度である。そして同等の困難度をもつ課題の上で、個人間の成績差は能力を表す。このように、課題の困難度と個人の能力は、基本的には相補的なものであるものの、概念上は異なるものである。すなわち、一般的に時間制限を設ける GJT の成績は、時間制限を設けない GJT の成績よりも低い値を示すが、これは、時間制限という操作自体が課題の困難度を上げているからであり、個人内における異種の能力を反映していると考えerことは自然な推論ではない。また、困難度の高さが、母語話者の高い能力を連想させることはいうまでもないことである。つまり、構造的仮定は、母語話者という集団に帰属する高い能力を、暗示的知識というカテゴリーに帰属される課題の困難度に差し替えて投影しているわけである。第二言語習得研究では、このように能力と課題の困難度を混同する例がしばしば見られる。

時間制限の有無という対立ほど明確ではないが、SAT 現象は、時間的制約などに関わる課題には普遍的に現れるため、暗示的知識を測定するとされる各種の口頭課題における低成績が、直接的に学習者の貧しい暗示的知識を反映している、というように結論づ

けることもできない。つまり、教育現場などで頻繁に語られる「知っているけど、コミュニケーションの場では使えない」といったある種の印象の大部分は、時間的制約などによる課題の困難度の違いによって説明できる可能性が高い。そしてそれは、運転速度が上がればハンドルの精密さを欠くであるとか、急ぐと仕事に手落ちが増えるといったことと基本的には同じ現象であり、あらゆる運動技能 (motor skills) や認知課題の多くにみられる、非常にありふれたものであり (Heitz, 2014)、これらの現象のすべてが、明示的知識および暗示的知識などといった二種の構成概念を立てることによって説明されるわけではない。

このように、個人内分散は SAT によって説明されるとしても、一部の第二言語習得研究、特に明示的知識および暗示的知識の二分法に関する研究が当初目指したように、集団間の差と個人内分散を同時に説明する枠組みが求められることには変わらない。しかしながら、このことはいうまでもなく自明であり、そして同時に皮肉なことではあるが、逆に、学習者と母語話者の運用を群間において測定不変性をもつ同一尺度上にて比較を行なうことのみによって可能になる。つまり、従来の明示的知識および暗示的知識の二分法が仮定していた、二分法およびその構造的仮定を廃する、ということである¹⁸。構造的仮定を廃し、文法知識に対して一元性を仮定することによって、これまでは無視されていた母語話者間の個人差、そして個人内における分散を、学習者に対するものとまったく同じ枠組みによって同時に説明できるようになる。

たとえば、GJT の成績に着目するならば、母語話者は集団レベルにおいて、学習者と比べて比較的高い正答率と短い反応時間を示すが、前節で述べたように、個人内において SAT 現象が見られること自体は、学習者と異なるわけではない。つまり、時間制限を設けるといった課題操作によって正答率が下がること、または、やや抽象的にいって、困難度が高い課題において、個人内で相対的に低い成績を示すといった現象自体については、集団間においてまったく同等なのである。

ここで、個人内における SAT 現象と、その関数の個人差について考えてみる。図 3.2 (a) が示すように、ある判断課題における個人の正答率は、課題操作によって、ある関数に従って低減する。一方 (b) は、複数の個人における関数を示している。一般的に、能力が高くなれば、課題操作の影響から独立していくため、速さに関する同じ条件下においても、より高い正答率を示すようになる。つまり、個人の能力は、SAT を表す関数全体を規定しているのである (e.g., Wickelgren, 1977)。このような SAT の枠組みは、従来の外国語教育研究および第二言語習得研究において、あまり十分に検討されたものではなかった。しかしながら、個人内分散と個人間分散、そして学習者や母語話者といった集団間の差を同時に説明することができる点において優れているといえよう。このよう

¹⁸ ここでは、明示的知識および暗示的知識が実在し、それぞれが観測変数の値の変動における原因となっている、という考え方を示す。

な SAT の枠組みを踏襲することによって、新しい文法知識に関する枠組みが提案できるようになると本論は考えている。

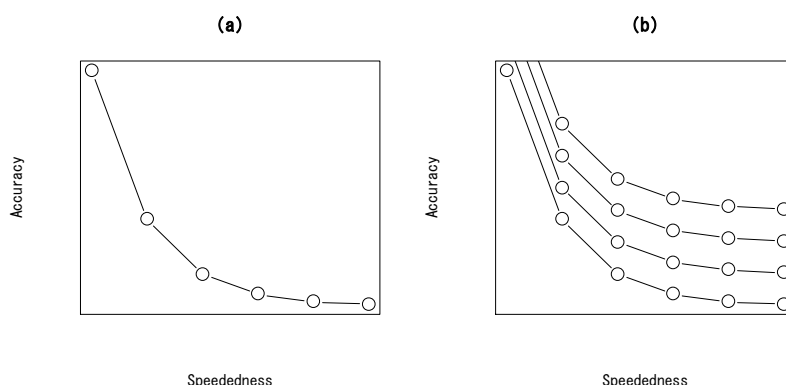


図 3.2 SAT 現象と個人差

3.5 外国語の文法知識に関する一元観

本節では、明示的知識および暗示的知識の二分法に対する代替案として、外国語の文法知識に関する一元観を提案する。この枠組みは、あくまでも外国語教育における実務的応用を念頭に置いた、従来の明示的知識および暗示的知識の二分法に対抗する認識論的な枠組みであり、認知主義的な機構や、具体的な数理モデルに必ずしも対応するわけではないことに注意されたい。一元観は以降で説明するように、認知心理学や数理心理学における基本的なモデリング方針の一部と大筋において相違ないが、それらに属するモデルと区別するため、独自の名称をつけている。

さて、一元観は、第一に、外国語の文法知識に関して表象主義を採用しており、個人がもつ知識表象を、モデルの中核的単位とする。その根幹をなす概念である証跡 (evidence) とは、学習者、母語話者を問わず、個人がある有限的な知識表象ないし項目に対してもつ、表象の質 (quality of representation; cf. Farah, 1994; Cleeremans & Jiménez, 2002) ないし強度 (e.g., strength, trace) である。証跡は、単一の、そして一次元の連続量であり、知識表象と個人の組み合わせに対して固有である¹⁹。この証跡についての仮定を、外国語における文法知識の一元性と呼ぶ。つまり、ある個人 i は、知識表象 p について、単一であるところの連続量 e_{ip} のみをもつ。よって、明示的知識および暗示的知識の二分法に仮定されるような、ある知識表象 p について、個人が二つの排他的表象、つまり複表象 (plural representations) をもつであるとか、その上で課題や状況によって使用する表象が異なる、といった仮定は不要になる²⁰。同時に、証跡は個人と項目の組み合わせのみにおいて固有

¹⁹ 個人 x は、文法構造 a, b, c に対してそれぞれ固有の証跡をもち、個人 y が文法構造 a, b, c に対してもつ証跡とは異なる値を取ると考える。

²⁰ たとえば、複数形態素に関する表象 P があるとして、明示的知識であるところの P 、暗示的知識であるところの P といっ

の値であるのだから、学習者や母語話者といった集団間の差は、集団における期待値の差でしかなく、判断行動の機構自体が母集団間によって異なるものとも仮定しない²¹。また、証跡の値自体はいかなる課題条件などとも基本的に独立である。

次に、あくまでも理論的、または概念的にはあるが、証跡は、ある個人が当該の知識表象を使用する際に、運用に関わる SAT 関数自体を規定する。すなわち、証跡における高い値は、高い正答率、短い反応時間の原因となり、証跡における低い値は、低い正答率、長い反応時間の原因となり、高い正答率、長い反応時間といった観測の原因になることはない。つまり、反応時間と正答率は基本的に共変関係にあると仮定している。課題成績における個人内分散は、このように理論上の SAT 関数によって説明され、同時に、個人間や集団間における運用の差は、個人間および集団間の期待値などによって直接的に比較できるものになる。このように、証跡の値を、個人と項目の組み合わせにおいて固有であるとみなすことは、文法項目間の観測の違いを知見に取り入れることを可能にする。つまり、一般に習得困難であるとされる項目について、個人は弱い証跡をもち、それによって、低い正答率や長い反応時間を示す、といった具合である。ここでの項目とは、抽象的な単位であって、判断課題における具体的な項目や、ある種の構造に関する知識表象としてみなすこともある。

このような仮定の根本は、信号検出理論 (signal detection theory; 外国語教育研究の文脈では, Kusanagi, 2014; 草薙・後藤, 2016)²²や再認記憶研究における古典的な痕跡強度 (trace strength) といった枠組み、具体的には、強度理論 (Strength Theory; Banks, 1970; Norman & Wickerlgren, 1969; Mardock & Dufty, 1972) による。強度理論では、課題特性に関わらず、同一の記憶システムが使用され、項目がもつ痕跡強度によって、判断行動がなされると仮定する。また、痕跡強度は反応時間を決定する要因であると捉える。現在の数理心理学や認知心理学において、古典的な信号検出理論や強度理論は必ずしも主流であるとはいえないが (e.g., Wixted, 2007; Yonelinas, 2002)、項目が、単一の連続量であるところの痕跡強度をもち、判断成績、反応時間、そして確信度やさまざまな主観的変数に影響を及ぼすという見方自体は概ね受け入れられているものである。さらに、主観的変数に関して、人工言語学習研究において、Cleeremans and Jiménez (2002) や Ziroi and Dienes (2012) も、これと同様の捉え方を示しており、単一であるところの表象の質が高くなるにつれ、その知識表象は意識化可能なものに変化すると主張している。

たものの存在を仮定しない。

²¹ ここでは、あらゆる言語運用を仮定しているわけではなく、主に文法性判断といった判断行動のみを念頭に置いている。よって、あらゆる種類の言語運用が学習者と母語話者に共通であるという含意はなく、少なくとも判断行動は、集団間においても等質な過程によるものだけということである。

²² 信号検出理論は、判断行動の分析に関する古典的かつ基本的な手法のひとつであり、判断行動のデータから、弁別力、反応バイアスといった指標を返す数理モデルである。

数理心理学は、特に判断行動や判断課題に関する数理的なモデルを扱う分野であり、そのもっとも代表的な数理モデルは、Ratcliff による拡散過程モデル (Diffusion Model; Ratcliff, 1978; Ratcliff & Rouder, 1998, 2000; Ratcliff & McKoon, 2008; Ratcliff & Smith, 2004) であり、本論は次章において、このモデルに依拠した実験を行っている。このモデルは、判断課題における正答率、正答の反応時間分布、誤答の反応時間分布を同時に予測することを目的としている²³。拡散過程モデルは、物理学におけるブラウン運動ないしウィーナー過程といったランダムウォーク現象をモデルの根本原理にするものであるが、さまざまな分野において顕著な応用例が見られる。このモデルにおいても、同様に個人ないし項目や、その組み合わせにおいて、単一の痕跡強度を理論的に仮定している。これらのように、知識表象に対して、ある種の一次元的な連続量を仮定することは、認知心理学や数理心理学におけるもっとも古典的な、そして基本的な方略のひとつである。そして、実際に次章以降にて示すように、単一の痕跡強度のみを仮定し、複数の知識表象を仮定しない拡散過程モデルは、外国語における GJT から発生するデータの優れた数理的近似になる。すなわち、本論が提案する一元観は、このような認知心理学および数理心理学における古典的かつ汎用的なモデリング方針が、外国語に関する判断行動のモデリングにおいても有用であり、外国語における判断行動は、他の判断行動に対して特異的ではない、と主張する²⁴。

一元観は、運用に対してのみではなく、心的状態に対しても一元性を仮定する。すなわち、強い証跡は、ある種の心的状態、たとえば強い確信度であるとか、規則、想起、説明可能といった印象の原因になり、一方の弱い証跡は、弱い確信度、直感、推量といった印象の原因となると仮定する²⁵。

ここまでの仮定をまとめると、証跡は個人と知識表象の組み合わせにおいて固有の潜在変数であり、判断成績、反応時間、そして主観的変数²⁶といった観測変数に影響を及ぼしていることになる。この関係を、簡略的に図 3.3 に表す。

²³ 拡散過程モデルやその亜種のモデルは、判断結果のみを対象とする信号検出理論と比較して、判断結果および反応時間分布に対する数理モデルであるといわれることが多い。

²⁴ 一元観と具体的な数理モデルであるところの拡散過程モデルは同一のものではない。一元観の主張に沿う他種のモデルは存在し得るし、同時に拡散過程モデルは一元観に先んじて導出されている。一元観とは、明示的知識や暗示的知識といった二分法に対する比較の上でのモデリング方針、ないしモデリング・コンセプトのひとつである。

²⁵ 認知心理学における明示的知識および暗示的知識の二分法に従えば、強い証跡をもつ表象は明示的知識、弱い証跡をもつ表象は暗示的知識であると考えられることになる。

²⁶ 心的状態に対応する観測変数のひとつとして主観的変数を想定している。

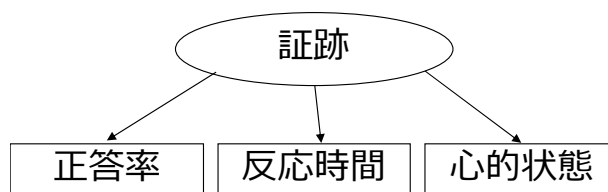


図 3.3 証跡と観測の関係

ここまでの議論において、証跡の獲得過程、つまり証跡の時系列間における変化についてなにも言及していないことに注意されたい。一元観の下では、たとえば判断課題の成績といった観測のみから、その知識表象に関する過去の獲得過程、ないし証跡の変化が弁別可能であるとは考えない。むしろ、時系列間における実測データより、証跡の変化を経験的に推定する方法論がより適切であると考え。すなわち、一元観は、知識表象の獲得過程、ないし証跡の変化に関して厳密な経験主義的態度をもつ。個人内における証跡の変化は、時点 t_1 と時点 t_2 における証跡の差、

$$e_{t_2} - e_{t_1} \quad (3.1)$$

として表すべきであると捉え、ある期間内における発達や、ある学習法の効果は、このような証跡の差によって説明することを念頭に置く。このように、任意の時点間における証跡の変化を、便宜的に証跡の遷移 (transition) と呼ぶ。

ある学習行動の帰結として、観測より得られるであろう証跡の遷移パターンに、ある種の一貫性や法則性が見出されることは、十分にもっともらしいことである。たとえば、Ziori and Dienes (2012) が示すように、無意識的学習は、表象の質に対して漸進的な変化をもたらすであろうし、そしてメタ言語を伴うような意識的学習は、表象の質に対して離散的な変化をもたらすということは、十分に予測可能である。このように、ある種の学習行動と、それが影響を及ぼしうる遷移パターンの組み合わせを、遷移制約 (transition constraints) として、形式的な手法によって記述していくことは、さまざまな場面における外国語教育の実践において、十分な効用をもたらさだろう²⁷。

さらに、一元観は、これまで説明が困難であった、以下の二つのような観測的事実について一貫した説明をもたらさう。

²⁷ ただし、このような言明に関する実証的研究を本論は行っていない。遷移パターンに関する研究は、今後の研究の優先的課題である。

(a) GJTにおける反応時間分布

GJTにおいて、明示的知識および暗示的知識の利用という二種類の認知過程があるのであるとすれば、そして時間制限の有無によって、それぞれの認知過程が弁別可能になるのであれば、観測の反応時間分布は、混合分布 (mixture distribution)²⁸を示さなければならない(草薙, 2013)。しかしながら、一貫してこの予測の証拠は得られず、指数正規合成分布に従うような観測が得られることが報告されている(e.g., Kusanagi, 2014; 草薙・川口, 2015)²⁹。また、混合分布モデルは多くの母数をもつため、外国語教育研究における小標本実験では、その解析が技術的にやや困難である。しかし、一元観を採用することにより、そもそも観測に対して混合分布を期待する必要はなくなる。

(b) 運用と心的状態の関係

R. Ellis (2005) では、暗示的知識が、高い判断課題の成績、そして短い反応時間の原因であり、その上、暗示的知識の使用は直感といった心的状態を伴うと予測されていた。しかしながら、小標本の実験であるが、草薙・川口 (2015) によると、認知心理学における一般的な結果と同様に、観測上、短い反応時間を示す反応は、直感反応であるよりも、規則反応である確率が高く、さらに正答は、直観反応よりも、規則反応である確率が高かった。一元観では、この研究結果も踏まえ、短い反応時間は、高い正答率と関連し、そして規則反応である確率が高いことを予測する。

外国語教育の実践において、一元観は、さまざまな面において高い有用性をもつ。そのもっとも重要な点は、教育評価の効率性についてである。明示的知識および暗示的知識の二分法は、その測定モデルにおいて測定不変性をもたないため、異なる母集団間、そして異なる項目間において、その値が比較不可能であり、教育的な応用可能性が相対的に低い。そもそも教育現場において、複数の測定具からなるテストバッテリーの運用は、困難である。また、明示的知識および暗示的知識というように、二つの成果を想定する場合、成果が分離することにより、シラバスの策定や教材開発が一層複雑化してしまう。

一方、一元観が主張するように、文法知識を一元的に捉えることで、測定不変性の問

²⁸ 複数の確率分布の混合からなる分布を混合分布と呼ぶ。もっとも典型的な例は、男女の身長分布である。男女の身長は、それぞれ異なる母平均、母標準偏差をもつ。それぞれの分布に正規分布を仮定するならば、男女問わずにサンプルを得た身長分布は、混合比とそれぞれの母平均および母標準偏差をもつひとつの確率分布としてみなすことができる。もしも仮に、明示的知識の使用、暗示的知識の使用という二つの認知過程が混在するならば、その帰結であるところの反応時間が混合分布を示すと予測することは自然な推論である。

²⁹ 指数正規合成分布は、反応時間や読解時間を記述するためのもっとも頻繁に使用されている確率分布である。この分布は正規分布と指数分布の合成からなり、 μ , σ , τ という3母数をもつ。

題を回避し、測定における実施上の負担を軽減し、さらに成果となる変数を一元化することによって、評価と他の教育的行動における一貫性が一層高まることが予測される。また、結果的に、評価行動を全面的に効率化することにつながり、将来的には、かつては不十分であった時系列データの取得やそれにもとづく効果測定が容易になる。そのようにして得られた時系列データは、証跡の遷移制約を示す理論的材料にもなると考えられる。このように、一元観は、明示的知識および暗示的知識の二分法に比べ、高い社会的効用をもたらすと期待される。

3.6 第三章のまとめ

本章の内容は、以下のようにまとめられる。

- (a) 外国語教育研究における数理的アプローチはその研究姿勢に、帰結主義・効用主義、経験主義・行動主義、形式主義・構造主義的な要素をもつ
- (b) 外国語教育研究における数理的アプローチは、観測に関する数理的構造を所与のものとして、その構造を説明する理論的体系を、構成物として捉える研究観をもっている
- (c) 明示的知識および暗示的知識の二分法に関する構造的仮定と、その因子分析モデルによる検証法には、さまざまな問題点が見られ、信憑性が低いと判断できる
- (d) 明示的知識および暗示的知識の二分法に対する理論的動機は、学習者と母語話者という集団間の差を説明する概念と、学習者における個人内分散を説明する概念を統合するということであった
- (e) これは、数理構造の上では、集団に対して特異的な行動を、異なる集団に属する個人内の行動に投影する見方である
- (f) しかし、明示的知識および暗示的知識の二分法に関する構造的仮定では、このような構造を適切に説明することができない
- (g) 一方、文法知識に関わる観測変数の構造を一元的であるとみなすことにより、学習者と母語話者の差、そして個人内の差は、SATの観点から統一的に説明することができる
- (h) 一元観は、個人と知識表象ないし項目の組み合わせに固有な証跡を仮定する
- (i) 証跡は、理論的に一次元的な量であり、個人のSATを規定する
- (j) 一元観は、人工言語学習研究、再認記憶研究、数理心理学の基本的な概念に対して整合的である
- (k) 証跡における値の時系列変化を遷移と呼び、合理的推論からではなく、実際の観測から遷移を経験的に記述する方法論を推奨する
- (l) 一元観は、従来の枠組みでは説明できない観測的事実を説明できる
- (m) 一元観は、明示的知識および暗示的知識の二分法や、その因子分析モデルに比べ、教育実践上、高い社会的効用をもたらすと期待される

次章では、本論が提唱した一元観に関する、もっとも基礎的な実証的根拠を得るための実験について紹介する。

第四章 本実験

4.1 数理モデルと研究仮説

前章にて提案した一元観は、外国語教育研究全般に関わる主張を含んでいるが、本論では、一元観がもつもっとも基礎的な前提についてのみ検証の対象とする。その前提とは、文法性判断課題の試行から得られる3つの観測変数（正答率、反応時間、主観的変数）の関係性についての予測から導出されたものであり、これら3つの観測変数の関係性が、一元的であるところの証跡ないし表象の質のみを仮定した数理モデルの予測にフィットするということである。第一に、正答率、正答の反応時間分布、そして誤答の反応時間分布が、単一的な表象の質を仮定する代表的なモデルのひとつである、拡散過程モデルによってある程度適切にモデリングできることは、上記を示すひとつの証拠になり得る。

拡散過程モデル (e.g., Ratcliff, 1978; Ratcliff & McKoon, 2008; Wagenmakers, 2009) とは、前章でも触れたように、比較的単純な判断行動から発生するデータを説明するための数理モデルのひとつである。判断行動とそれに至るまでの認知過程を、決定を下すまでに十分な情報を集積する過程として、そしてその情報の集積を、ノイズを含んだ確率過程の一種として捉えることに、このモデルの顕著な特色がある。具体的に、このモデルは、観測における正答率、正答の反応時間分布、そして誤答の反応時間分布を9つの母数によって説明する。これを簡略的に表すと、図4.1のとおりになる (e.g., Ratcliff, 1978)。

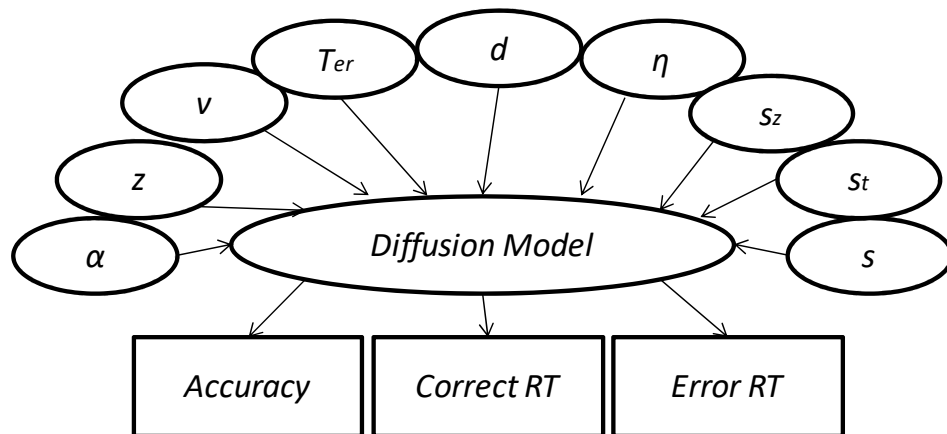


図 4.1 拡散過程モデルの概要 (Ratcliff & McKoon, 2008 などを元に著者が作成)

それぞれの母数についての概要は、以下の通りである (e.g., Ratcliff & McKoon, 2008)。

- (a) α : 境界母数ないし閾値母数 (boundary/threshold separation)¹。判断行動が実行されるまでに集積されるべき情報量の大きさとされる。この値が大きければ、より慎重な判断を要し、小さければスピードを優先するスタイルを表す。このことから、さまざまな認知的課題において SAT を仲介する母数のひとつとして解釈される
- (b) z : 開始地点 (starting point)。判断行動のバイアスを示すものとして解釈される。バイアス²がない場合、 $\frac{\alpha}{2}$ である。本稿ではバイアスが一切ないものとして扱う
- (c) v : 漂流母数ないし漂流比率 (drift rate)。判断行動に至るまでの情報の集積率として解釈される。この母数が高ければ全体的に優れたパフォーマンスを示す。認知的な文脈では、表象の質、証跡の強さ、または記憶の痕跡になぞらえられる。一元観が仮定する証跡という概念は抽象的なものであり、必ずしも数理モデル上のこの母数と相同ではないことに注意されたい
- (d) T_{er} : 非決定時間 (non-decision time) : 情報の符号化や身体的反応行動などに要する時間をあらわす
- (e) d : 反応 A と反応 B における身体的反応行動などにかかる時間の差をあらわす。本論ではこの値を扱わない
- (f) η : 項目間における漂流比率のばらつき
- (g) S_z : 項目間における開始地点のばらつき
- (h) S_t : 項目間における非決定時間のばらつき
- (i) s : 漂流定数 (diffusion constant)。本論ではこの値を扱わない

これらの母数がどのように正答率、正答の反応時間分布、そして誤答の反応時間分布を説明するかについては、図 4.2 のような確率過程が理解に役立つ。まず、拡散過程モデルでは、反応時間を、非決定時間と決定時間に分離する。非決定時間は、先述の通り、情報の符号化や身体的反応行動に要する時間である。決定時間とは、決定を下すまで行われる情報の集積に要する時間であるとされる。非決定時間は、 T_{er} とそのばらつきである s_t に従う。

次に、決定時間は、開始地点 z から、情報の集積率を表す漂流比率 v に従う酔歩 (random walk) によって、閾値母数 α または 0 まで到達する時間とされる。この過程は、時間的一様性 (time homogeneity) ないし定常増分性をもつ加法過程であるという意味では、レヴィ過程にも属する³。その性質から漂流母数 v によって、閾値母数 α まで到着する分布

¹ このモデルでは、以下に示すように、判断行動をある閾値ないし境界に達するまでの情報の集積過程と捉えており、その集積過程を確率過程とみなしている。このようにこのモデルにおける境界ないし閾値、そして漂流といった用語は、確率過程であるところのブラウン運動、ないしウィーナー過程に由来している。なお、それぞれの母数の解釈は本文の通りである。

² ここでのバイアスとは、反応バイアス (response bias) のことである。バイアスがある場合、正答誤答とは関係なく、どちらかの反応に多く回答が偏る。

³ そもそも確率過程とは、時間によって変化する確率変数を表す。拡散過程モデルは、文法的または非文法的であるという決定

は、逆ガウス分布⁴に従う。GJT を例にすると、閾値母数 α に到達すると「文法的」という判断を、逆方向である 0 に到達すると「非文法的」という決定を下す⁵。このとき、漂流母数における項目間のばらつきは η である。

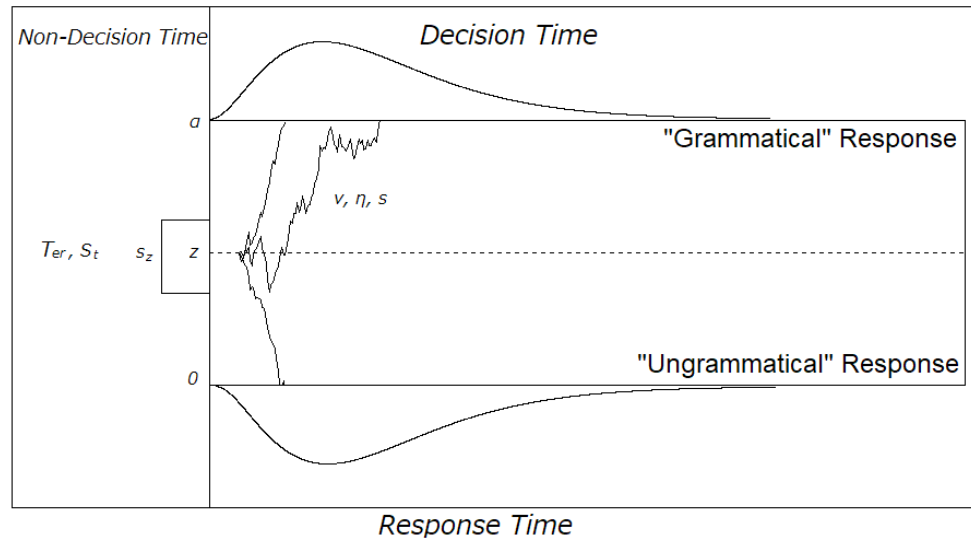


図 4.2 拡散過程を示す概略図 (Ratcliff, et al. 2004 などを参考に著者が作成)

拡散過程モデルは、正答率と反応時間分布を同時に説明することができることから、SAT を説明するための優れたモデルのひとつとされている (Ratcliff & McKoon, 2008)。それは、拡散過程モデルの主要な母数である、漂流比率 v と閾値母数 α によって、それぞれ判断者がもつ情報の集積率と SAT を、同時に解釈することができるためである。

この数理モデルの振る舞いを理解するために、それぞれの母数が正答率、正答の反応時間分布、そして誤答の反応時間分布にどのように影響を及ぼすかについて、図 4.3 にまとめる。閾値母数 α が高くなるにつれ、正答および誤答の反応時間分布の形状は変化し、それに伴い、分布のスケールも大きくなる。一方、漂流比率 v が高くなるにつれ、正答率は上昇し、さらに反応時間分布の中心傾向と形状も変化する。ここで、単一の漂流比

を下すための情報量が、時間によって変化する確率変数であるとみなしている。酔歩とは、時系列におけるその確率変数の変動がランダムである性質を示す。開始時点は、その確率変数の初期値であり、漂流母数は、時系列における変動の平均的な傾向として理解することもできる。閾値母数は確率変数の終点であり、確率変数が閾値母数と等しくなった時点において、決定が下ると想定される。確率過程における時間一様性ないし定常増分性とは、時系列に対して確率変数の値が独立であるということであり、拡散過程モデルにおいては、時系列上で漂流母数の値が変化しないという制約に相当する。時間一様性ないし定常増分性をもつ確率過程の総称が、レヴィ過程である。

⁴ 逆ガウス分布も反応時間分布に適合する連続型確率分布として知られている。

⁵ 理論的には本文の通りであるが、本論では、バイアスがないものとして扱い、正答と誤答として分析している。

率 v が、正答率と両方の反応時間分布に対して同時に影響を及ぼすことに注意されたい。すなわち、漂流比率が高くなれば、正答率が高くなり、反応時間は短くなり、この点によって SAT 現象を説明することができる。拡散過程モデルは、この点において一元観が仮定する条件と一致する。

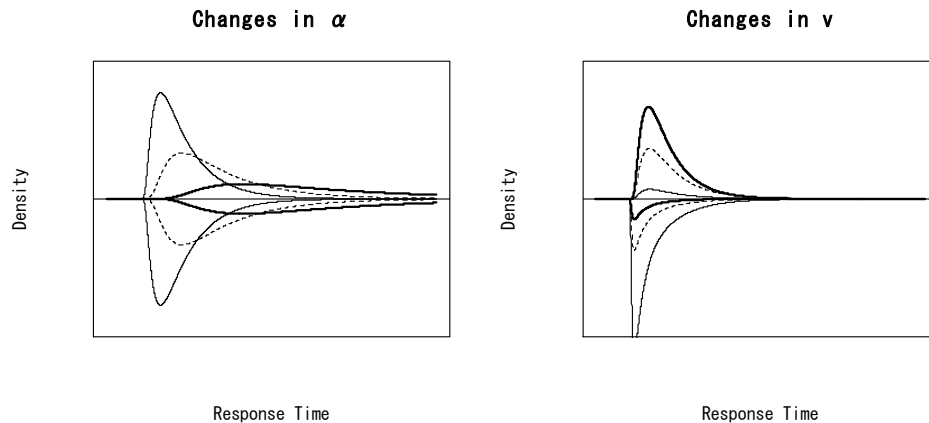


図 4.3 閾値母数と漂流比率の変化が正答率と反応時間分布に及ぼす影響の模式図

もしも、この拡散過程モデルが、外国語における GJT の正答率、正答の反応時間分布、誤答の反応時間分布に対して優れたフィットを見せるならば、明示的知識および暗示的知識といった複表象を仮定する必要がないことを支持する証拠のひとつとなる⁶。繰り返しになるが、拡散過程モデルが示すように、正答率と反応時間は、負の共変関係を示すということである。よって、本論では、以下の命題を研究仮説のひとつとする。

研究仮説 1: 拡散過程モデルは外国語における GJT の正答率、正答の反応時間分布、誤答の反応時間分布に対して優れた数理的近似となる

しかしながら、現在のところ、拡散過程モデルが外国語における GJT に適用された例はない。Ratcliff and McKoon (2008) が述べるように、拡散過程モデルは比較的単純で、短い反応時間 (< 1,500ms) を示す判断課題のみに適用されるべきであるとされている。一方、外国語における GJT は、これまで拡散過程モデルが適用されたあらゆる判断課題よりも、圧倒的に大きい測定スケールの反応時間分布 (e.g., 2,000ms < RT < 40,000ms) をもつことが明白である。そのため、外国語における GJT に対して、このモデルを適用す

⁶ 明示的知識および暗示的知識の枠組みは、現実の正答率および反応時間分布を予測することができない上、モデルがより複雑であるためである。少なくとも数理的アプローチの下では、現実への適合度が低いモデルがもつ性質のうちのいくつかは捨象するべきであると考えられる。

ることについては、慎重であるべきであろう。

ただし、本論は、その測定スケールの大きさに顕著な違いこそあれ、Ratcliff et al. (2004) が述べるような、多段階的な認知過程を要するような課題、すなわち推論課題などと比べ、拡散過程モデルの適用が原理的に不可能であるような特異性を、外国語における GJT がもつとは考えていない。実際に次章で示すように、拡散過程モデルは外国語における GJT から発生するデータのよい数理的近似となる。また、これまでの研究例において、正答率、正答の反応時間分布、そして誤答の反応時間分布を同時に説明する対立的モデルがない状況において、優れた数理的近似を示すということ自体がもつ効用についても十分に考慮するべきである。

次に、第三章で述べたように、心的状態を示す主観的変数に対して、正答率と反応時間が影響を及ぼすことも期待される。本論では、GJT における各試行の判断について「規則にもとづく判断である」という心象を伴う反応、つまり規則反応 ("Rule" response)⁷と、「直感にもとづく判断である」という心象を伴う反応、つまり直観反応 ("Intuition" response) を取り扱う。一元観が仮定する、正答率、反応時間、そしてこの主観的変数の連関関係にもとづき、応答変数⁸を主観的変数、説明変数を判断結果と、対数変換をした反応時間とした場合のロジスティック回帰モデル (4.1)⁹を想定するとき、観測から得られる推定結果は (4.2) の条件に沿うはずである。

$$\ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{1,i} + \beta_2 \log(x_{2,i}) \quad (4.1)$$

$$(\beta_1 > 0) \wedge (\beta_2 < 0) \quad (4.2)$$

ただし、このときの P は、試行が規則反応である確率、 x_1 を判断結果、すなわち正答 (1) ないし誤答 (0)、そして x_2 を反応時間とする。しかしながら、GJT の各試行のデータは階層構造を成すため、本論では、階層ベイズモデルによって、切片 β_0 , β_1 , β_2 についての個人間分散をモデルに取り入れている。具体的なモデルと推定方法については、本章、4.8 節を参照されたい。よって、(4.1) 式は、あくまでも簡略化した例であって、実際の解析モデルとは異なることに注意されたい。

本論では、(4.2) の条件を、二つ目の研究仮説とする。すなわち、

⁷ より具体的に、規則反応とは、主観的変数の測定において、「規則にもとづいて答えた」と回答した反応である。

⁸ 応答変数は、近年のモデリングにおける用語のひとつであり、より一般的には従属変数とも呼ばれる。

⁹ この式において、左辺はリンク関数と呼ばれる数理的処置を施している。記号 \ln は自然対数を底とした対数変換である。

研究仮説 2 : 試行が規則反応である確率は、判断結果および反応時間の影響を受ける¹⁰

本論は、以上の二つの研究仮説が正しいと判断されることを示す。

4.2 実験参加者

実験参加者が属する想定上の母集団は、英語を外国語として使用し、日本語を母語とする英語学習者である。日本の大学に在籍し、英語を学習する 212 名の大学生を標本とした。分析に含められた標本の全員が、上記の条件に適合することを、実験参加者によるアンケートを使用した自己報告によって確認した。また、裸眼または矯正視力において、標本の各人が実験参加に際して、視力に関連する実験上の不都合がなかったことを実験者に報告した。

標本サイズは、拡散過程モデルの推定が可能なデータのサイズを満たすように、データ収集の事前において 210 人と設定され、その後、標本サイズの調整は行わなかった。実験として使用する GJT における試行数 ($K = 48$) と掛け合わせ、およそ 10,000 試行程度必要であると判断した。この標本サイズの値は、GJT を研究方法として使用した過去の外国語教育研究と比べ、比較的大きい。しかし、本論は一切検定の類を行わないため、標本サイズが大きすぎることによる推論への弊害はないと考えられる。

実験参加者は、2016 年度および 2017 年度において、著者が指導する総合的英語学習科目を受講しており、授業内学習の一環として実験に参加した。所属大学における年次は一年次である。標本内における性別比は、男性 55% (116 人)、女性 45% (96 人) であり、年齢は 18 歳から 21 歳の範囲である。この標本の熟達度は概ね、TOEIC スコアにおいて 500 点台を中心に分布をしており、この熟達度は、CEFR における A2 から B1 程度に位置する。

なお、特に研究仮説 1 に関して、GJT における個人のデータに対して拡散過程モデルを近似させることは、推定の材料となる試行数の少なさから見て適切ではない。そのため、研究仮説 1 に関しては、10 人および 11 人のデータを無作為にプール化し、この 10 から 11 人のデータをひとまとまりにして、大被験者 (supersubject) という単位として扱う。本研究での大被験者数は 20 である。大被験者に関わる統計的処遇については、Ratcliff, Thapar, Gomez, and McKoon (2004) に詳しい。

¹⁰ ここでのより具体的な予測は、(4.2) の通りである。すなわち、試行が正答であれば、規則反応である確率が上がり、同様に、反応時間が短い場合、同確率が上がるということである。(4.2) は、回帰係数の符号によってこの予測を示しているものである。

4.3 実験具

実験具は、コンピュータ・ベースの GJT であり、この GJT を実施することにより、文法性の判断における判断結果、反応時間、主観的変数を観測した。当該の実験プログラムは、著者がプログラム開発ソフトウェアのひとつである、HSP (Hot Soup Processor) を使用して作成したものであり、いくつかの研究において、同種のプログラムが使用されている (e.g., 草薙・川口, 2015; Tamura et al., 2016)。このプログラムは、Windows OS 上のみで動作する。

すべての実験参加者 ($N = 212$) に対して同等の機械的環境を使用して実験を実施した。PC モニターは 20 インチ、解像度は SXGA (1,280×1,024)、リフレッシュレートは 60Hz であった。その他の PC に関するハードウェア上の特性も、実験参加者間において相違がなかった。

実験プログラムは、以下のような手順を繰り返す。まず、画面上左側に注視点 (+) が 1 秒間表示される。その 500 ミリ秒後、注視点から文字送り方向に刺激文が提示される。刺激文の提示から反応時間の計測を始める。刺激文の提示後に、キーボードの左矢印キーが押された場合、「文法文」反応として、右矢印キーが押された場合、「非文法文」反応として反応を記録する。同時に、反応が得られるまでの時間を反応時間として記録する。反応が得られた 500 ミリ秒後、画面左側に「規則」、右側に「直観」と表示される。文法性の判断と同様に、キーボードの左矢印キーが押された場合に「規則」反応を、右矢印キーが押された場合には「直観」反応を記録する。これを 48 試行繰り返すとプログラムは終了し、実験結果を出力する。

実験プログラムの概要を図 4.4 に提示する。

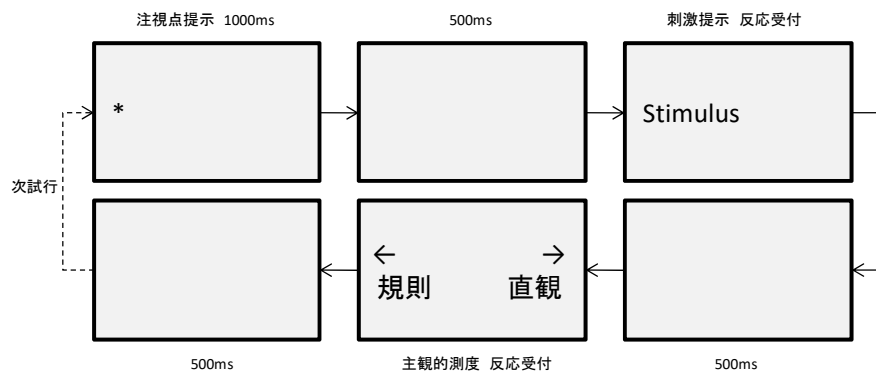


図 4.4 実験プログラムの手続きを示す模式図

4.4 刺激

文法文と非文法文からなる刺激文を使用した。刺激文の総数は48文であり、そのうちの半数となる24文が文法文であり、残りの半数が非文法文 ($k=24$) である。

この刺激のセットは、Kusanagi (2014) や草薙・川口 (2015) といった研究において、既に使用されているものと同じであり、R. Ellis (2006) などを参考とし、日本語を母語とする大学生レベルの学習者の文法知識を測定するために開発された。測定ドメインとなる言語規則において、ある程度の代表性を確保するために、8種類の言語構造を選定し、それぞれ6文ずつ刺激文を割り当てている。その8種類の言語構造とは、(a) 副詞の位置、(b) 助動詞、(c) 所有格、(d) 動詞の補語、(e) 埋め込み疑問文、(f) 仮定法、(g) 不定冠詞、そして (h) 三人称単数現在形である。それぞれの構造における文法文、非文法文の例は以下のとおりである。太字は、対比となる点である。なお、本論では、これらの言語構造を説明変数としてもちいることはない。

- (a.1) *The old man **always** likes always watching television.*
- (a.2) * *The old man likes **always** watching television.*
- (b.1) *The boys must not **leave** this building now.*
- (b.2) * *The boys must not **to leave** this building now.*
- (c.1) *Joseph is going to meet the local **leader's** son in two days.*
- (c.2) * *Joseph is going to meet the local **leader** son in two days.*
- (d.1) *The young man decided **to buy** a car next week.*
- (d.2) * *The young man decided **buying** a car next week.*
- (e.1) *She has no idea about what **she should** do next.*
- (e.2) * *She has no idea about what **should she** do next.*
- (f.1) *If she had another bag, she **could** pack all the books.*
- (f.2) * *If she had another bag, she **can** pack all the books.*
- (g.1) *The woman caught **a cold** because of the bad weather.*
- (g.2) * *The woman caught **cold** because of the bad weather.*
- (h.1) *The woman sometimes **plays** the piano in the room.*
- (h.2) * *The woman sometimes **play** the piano in the room.*

全48文における語数の分布は図4.5および表4.1の通りである。語数の平均値は、9.41語である。

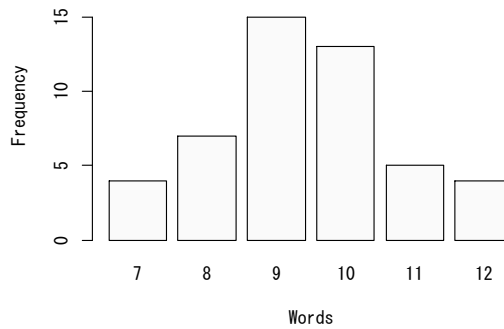


図 4.5 実験に使用された刺激文における語数の分布 (K = 48)

表 4.1

実験に使用された刺激文における語数の分布とその比率 (K = 48)

7 語	8 語	9 語	10 語	11 語	12 語
4 (.08)	7 (.15)	15 (.31)	13 (.27)	5 (.10)	5 (.08)

4.5 刺激の測定論的質

続いて、刺激文の質について検証した。まずは、 α 係数について、ブートストラップ法 ($B = 1,000$) をもちいて区間推定を行ったところ、95%水準において、[.65, .78]であり、点推定値は $\alpha = .71$ であった。このことから、古典的テスト理論の下で、当該の刺激は、測定の内部一貫性をもっていると評価した。

つぎに、ラッシュ・モデルおよび二母数モデルを利用し、テスト情報量関数の推定を試みた。まず、ラッシュ・モデルと二母数モデルを、判断結果に対して、最尤法によってフィットさせたところ、より優れたフィットを示したのは、二母数モデルであり、その尤度比検定の検定統計量は、330.89 ($df = 47$) であった。ラッシュ・モデルと二母数モデルのフィットについて表 4.2 に示す。

表 4.2

ラッシュ・モデルと二母数モデルのフィット ($N = 212, K = 48$)

モデル	対数尤度	赤池情報量基準	ベイズ情報量基準
ラッシュモデル	-4,997.82	10,093.64	10,255.50
二母数モデル	-4,832.38	9,856.76	10,173.87

二母数モデルが比較的優れたフィットを示したため、このモデルを使用してテスト情報量関数をもとめたところ、図 4.6 のような結果となった。テスト情報量は 35.39 であり、能力が負の値を取る場合におけるテスト情報量関数下の面積は、87.55% であった。また、

能力が正の値を取る場合におけるテスト情報量関数下の面積は、14.45%であった。このことから、刺激は、比較的好成績を示す実験参加者の能力推定に対してやや鈍感であることがわかった。ただし、この特性は、個人のテスト得点を使用しない本実験の分析において、結果の解釈に大きな影響を及ぼさないと考えられる。以上より、本論の目的に対して本実験の刺激は、その測定において十分な質を有すると総合的に判断した。

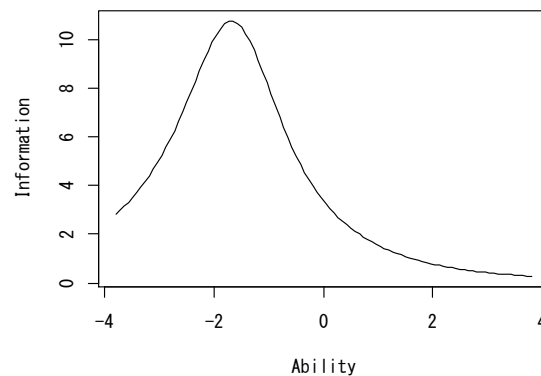


図 4.6 刺激のセットにおけるテスト情報量曲線 (K=48)

4.6 実験手続き

著者が担当する 4 つの総合的英語授業科目における、授業内学習活動の一環として、実験参加者に対して、それぞれ一斉に実験を行った。これらの授業科目では、文法や語法に関する知識の伸長を授業目標の一部として据えている。授業内において、実験手続きと個人情報の管理、研究倫理事項について説明を実験事前に行った。実験環境は、60 席からなる CALL 教室であり、実験具の節にて述べたように、機械的条件は実験参加者間においてまったく同一である。紙面にて、デモグラフィック情報および研究参加への是非を記入させたのち、口頭で実験の概要について、実験参加者に対して説明を行い、その後、5 試行からなる練習セッションを実施し、本実験である GJT (K=48) を実施した。実験に関する説明、デモグラフィック情報の記入、練習セッション、そして本実験に要した時間は、およそ 20 分程度であった。本実験終了後、それぞれの授業にて、本実験の刺激文についての文法的解説を著者が実施した。この文法的解説はおよそ 45 分の時間を要した。

本実験である GJT の試行に関する具体的な指示は以下のとおりである。まず、文法性の判断について、「提示された文章が、文法的か、非文法的か判断してください」という指示を与えた。次に、主観的変数について、「文法的か、非文法的か判断したのち、その判断が<規則にもとづく判断である>という印象をもったか、または、<直観にもとづく判断である>という印象をもったか教えてください」という旨の指示を与えた。また、

反応時間を記録していることを、実験参加者に伝えたが、「できるだけ早く回答をしてください」といった指示を与えることはしなかった。これは、時間的制約を設けることによって、信号検出理論が扱うところの反応バイアスが生じ、反応時間分布に強く影響を及ぼすからである (speeded effect; Kusanagi, 2014)。

4.7 変数とデータの構造

本実験が扱う変数は、三つである。(a) GJT の各試行における正答ないし誤答といった判断結果、(b) 反応時間、そして (c) 主観的変数である。研究仮説 1 を検証する際、判断結果と、正答の反応時間、誤答の反応時間という 3 つの変数が観測変数であり、拡散過程モデルの各母数は、ある意味において潜在変数である。一方、研究仮説 2 を検証する際には、判断結果と反応時間は説明変数であり、主観的変数が応答変数である。判断結果と主観的変数は二値データであり、反応時間は連続変数である。

ただし、文法性判断の各試行は、実験参加者と項目に対して交差的にネストしている。本論では、項目レベルについての変量効果を捨象する¹¹こととした。

4.8 解析モデルと解析環境

本論における解析には、全て統計解析環境である R (R Core Team, 2016) を使用した。記述統計および刺激文の質に関しては、psych パッケージ (Revelle, 2016) と ltm パッケージ (Rizopoulos, 2006) を、拡散過程モデルに関しては、rtdists パッケージ (Singmann, Brown, Gretton, & Heathcote, 2017) を、階層ベイズモデルには MCMCpack パッケージ (Martin, Quinn, & Park, 2011) を使用した。なお、信号検出理論関係の指標の計算は、草薙・後藤 (2016) の関数を、さまざまな場面における最尤推定には、stats パッケージ (R Core Team, 2016) の optim 関数、または nlminb 関数を使用している。

研究課題 1 に関して、拡散過程モデルの推定は、rtdists パッケージの ddiffusion 関数によって尤度関数を計算し、PORT ルーチンを使用する nlminb 関数によって、それを数値解析的に最小化することによって母数の推定値をもとめている。本論では、全母数 (full parameter) の拡散過程モデルではなく、母数 z , s , そして d について推定を行わず、定数として扱っている。rtdists パッケージのデフォルト設定に従い、それぞれの母数は以下の値をもつ。

¹¹ このこと理由は、解析および結果の解釈が著しく困難になるためである。ただし、項目間における変動が理論的な対象外になるわけではなく、以後の研究において当該の変動を検証することが望まれないわけではない。

$$z = \frac{1}{\alpha} \quad (4.3)$$

$$s = 1 \quad (4.4)$$

$$d = 0 \quad (4.5)$$

なお、尤度の最小化に関する数値解析には、常に以下のセットの初期値を使用している。

$$a = 5.5 \quad (4.6)$$

$$v = 0.2 \quad (4.7)$$

$$T_{er} = 2.0 \quad (4.8)$$

$$s_z = 1.0 \quad (4.9)$$

$$s_{ter} = 1.0 \quad (4.10)$$

$$s_v = 1.0 \quad (4.11)$$

次に、研究仮説 2 に関しては、全面的にベイズ統計を採用した。ベイズ統計とは、一般的な頻度主義による統計とは異なる確率観に基づく統計手法のひとつである。頻度主義下においては、母数は定数であり、観測が確率変数であると捉えるが、ベイズ統計においては、観測を定数とし、それ以外の全てを確率変数とみなす。また、ベイズ統計は、主観的確率を受け入れており、これを事前確率、または分布をなすものと考えて事前分布とよぶ。ベイズ統計の目的は、定数であるところの母数の値を推定することではなく、事前分布とデータのもっともらしさである尤度に比例する事後分布を構築することである。階層ベイズモデルとは、ある母数の分布を特徴づける母数が、別の事前分布に従うモデルのことである。ここでは、一般化線形混合効果モデル (GLMM) を、階層ベイズモデルの一種として表現している。本論はマルコフ連鎖モンテカルロ法 (MCMC)¹²によ

¹² マルコフ連鎖モンテカルロ法は、主にベイズ推定において使用される計算アルゴリズムのひとつであり、事後分布に従う確率分布から生成される乱数をサンプリングすることができる。この事後分布から生成される乱数の経験分布を事後分布とみなす。ベイズ統計では、解析的に事後分布が求められない場合、このような数値計算を行なうことが一般的である。つまり、マルコフ連鎖モンテカルロ法は、方程式で解けない問題を解くシミュレーション・ベースの方法のひとつである。具体的なアルゴリズムには、メトロポリス・ヘイスティングズ法、ギブス・サンプリング、ハミルトニアン・モンテカルロ法がある。本論では、前述のうち、最初からふたつのアルゴリズムを使用している。正答率、規則反応率、反応時間の記述目的において、MCMCを使用する場合は、メトロポリス・ヘイスティングズ法を、研究課題 2 に関するモデルでは、ギブス・サンプリングを使用している。前者の場合、チェーン数を 1、バーンイン区間を 500、MCMC サンプル数を 20,000 としている。ここで、すべての母数についての事前分布は無情報事前分布である。後者の場合、チェーン数を 1、バーンイン区間を 5,000、間引き区間 (thinning interval) を 10、MCMC サンプル数を 2,000 としている。母数についての事前分布

る階層ベイズモデルを使用している。

研究仮説 2 を検証する実際の解析モデルは、

$$\ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{1,i} + \beta_2 \log(x_{2,i}) \quad (4.12) = (4.1)$$

に対して、個人レベルの変量効果を β_0 , β_1 , β_2 の値に対して仮定するものであり、概念的には、

$$\ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{1,i} + \beta_2 \log(x_{2,i}) + r_j \quad (4.13)$$

として表すことができる。

本論では、上記の式の上で、 β_0 , β_1 , β_2 の個人間分散と共分散について報告する。ここで、それぞれの β ($q=3$) における事前分布について、多変量正規分布 (multivariate normal prior) を仮定したが、平均については (4.14) の行列を、分散共分散の事前分布は、(4.15) のようにした。なお、個人レベルにおける変量効果の分散共分散行列について、逆ウィッシュャート分布 (Inversed Wishert Prior) を仮定したが、その形状母数については $r=3$ とした。尺度行列 \mathbf{R} については (4.16) のように設定した。実際の推定結果およびそこから派生する解釈に事前分布の影響はほぼないものと考えられる¹³。

$$\boldsymbol{\mu}_\beta = [0,0,0] \quad (4.14)$$

$$\boldsymbol{\sigma}_\beta = \begin{bmatrix} 100,000 & 0 & 0 \\ 0 & 100,000 & 0 \\ 0 & 0 & 100,000 \end{bmatrix} \quad (4.15)$$

$$\mathbf{R} = \begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.01 \end{bmatrix} \quad (4.16)$$

は以下のとおりである。事前分布や事後分布といったベイズ統計における基本的な用語、およびベイズ統計全般の説明については付録 B を参照のこと。

¹³ この段落では、当該の数理モデルにおける事前分布の設定を述べており、ベイズ統計はその性質上、事前分布の設定について詳細を必ず示さなければならない。ただし、この設定は、本文で述べているように事実上、ほぼ無情報であり、つまりデータ取得事前における研究者の直観などを反映するものではなく、さらにこの設定がもつ結果に対する影響は小さいと考えられる。なお、これらの分布の設定に関する数学的な詳細は、(4.15) から (4.16) に示されているように行列演算を含んでおり、やや難解なものであるが、ベイズ統計において一般的にみられる設定である。

4.9 外れ値の除外

外れ値の除去のために、各実験参加者 48 試行におけるパフォーマンスのデータから、11 人の参加者のデータを外れ値として除外した。次の段落はその詳細についてである。これは全データに対して、およそ 5% に相当する。このことにより、次章にて以後報告されるデータは、 $n = 201$ によるものである。

まず、信号検出理論を使用し、実験参加者全員について、48 試行における弁別力 d' および判断基準 c を算出した¹⁴。表 4.3 は、その記述統計であり、(4.17) は分散共分散行列である。

表 4.3

全実験参加者 ($N = 212$) の弁別力 (d') と判断基準 (c) に関する記述統計

	平均	標準偏差	尖度	歪度
d'	1.27	0.67	-1.35	3.40
c	0.13	0.30	0.35	0.29

$$\text{Cov}(d', c) = \begin{bmatrix} 0.45 & 0.03 \\ 0.03 & 0.09 \end{bmatrix} \quad (4.17)$$

この二変数に対して、多変量正規分布を仮定し、それぞれのケースのマハラノビス距離¹⁵の自乗 D^2 をもとめた。二変量における D^2 は、 $df = 2$ とした χ^2 分布に従うことが知られている。観測における D^2 の度数分布と、 χ^2 分布 ($df = 2$) の確率密度曲線を描き足すと図 4.7 のようになる。また、この観測における上側 95% 点の値は、5.93 であり、この値を超えるケース ($n = 11$) を外れ値と判断し、以降、一切の分析から除外した。なお、この値は、理論上の累積分布確率において、 $p = .948$ に相当する¹⁶。

14 ここでの信号検出理論とは、標準的な等分散正規分布モデルを表す。弁別力 d' は、文法文と非文法文をどれだけ正確に弁別できているかを表す指標であり、弁別力が全くの偶然である場合は 0 を取り、正の値を取る場合は、文法文を文法文と、非文法文を非文法文と判断する程度を示し、負の値を取る場合は、文法文を非文法文と、非文法文を文法文と判断する傾向を示す。大きな値は、より高い弁別力を示す。ここで得られた弁別力 1.00 程度の平均値は、通常高い弁別力として解釈される。一方、判断基準 c は、実験参加者がもつ反応バイアスを示す指標であり、反応バイアスがない場合は 0 の値を取り、正の値は文法文へのバイアスを、負の値は非文法文へのバイアスを示す。

15 マハラノビス距離は、多変量の相関を考慮した距離の一種であり、多変量解析などに使用される。また、多変量における中心からの距離、と理解することもできる。本文で述べられているように、ここでは、弁別力と判断基準の相関関係の下で、二つの変数上での外れ値を検出するために使用している。

16 マハラノビス距離の自乗は、 χ^2 分布に従うため、その理論分布から外れ値の閾値を決定することもできる。ここでは観測値における 95% 点を閾値としたが、これは理論上の χ^2 分布における 95% 点とほぼ等しい値であり、理論上の 95% 点を閾値と

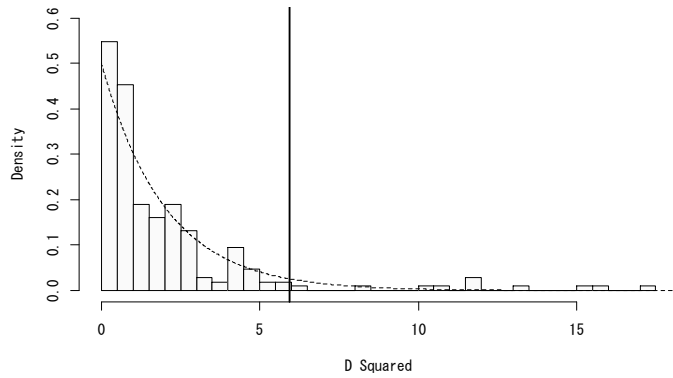


図 4.7 観測における D^2 の度数分布と、 χ^2 分布 ($df=2$) の確率密度曲線

これらの外れ値は、高い負の弁別力を示していたり、非常に高い判断基準を示していたりと、本来の実験結果からは想定できないデータであり、正答のキーと誤答のキーを誤るといった実験指示の誤解などがあったと考えられる。図 4.8 に、弁別力および判断基準の散布度を示している。この図において、三角で示されるケースは外れ値を、丸で示されるケースは分析に含められたケースを示している。楕円は 95%水準の相関楕円である。

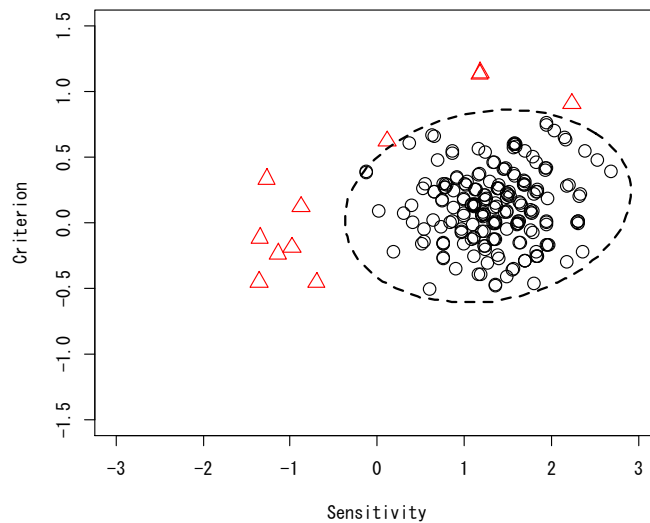


図 4.8 弁別力と判断基準の散布図と相関楕円

最後に、外れ値除去後における弁別力と判断基準の記述統計を表 4.4 に示す。この表からも分かるように、本実験における、統計上の平均的な実験者は、高い弁別力と、文法文に対するいくばくかのバイアスを持っていたことが推測される。

した場合は、結果的に変わらないことを示している。

表 4.4

分析対象となる実験参加者 ($n = 201$) の弁別力 (d') と判断基準 (c)

	平均	標準偏差	尖度	歪度
d'	1.36	0.51	-0.15	0.04
c	0.12	0.28	-0.04	0.02

第五章 結果

5.1 観測の記述とベイズ推定

この章の目的は、一元観に関する基礎的な前提についての実証的根拠を報告することである。前章で述べた 2 つの研究仮説について検証する前に、まずは観測の記述を行なう。最初に、個人レベルおよび項目レベルにおける主要三変数の記述統計を示す。表 5.1 は、個人レベル ($n = 201$) における平均正答率、平均規則反応率¹、平均反応時間の記述統計を示しており、表 5.2 は、項目レベル ($K = 48$) における同変数の記述統計である。図 5.1 は、観測の度数分布を示している。

観測の正答率において、その平均値が.71 程度であり、チャンスレートよりも高い値を示している。また、項目間における分散よりも、実験参加者間における分散の方がやや小さい傾向が見て取れる。この傾向は平均反応時間についても同様である。しかし、平均規則反応率については、項目間分散よりも実験参加者分散の方が大きい値を示している。

表 5.1

個人レベル ($n = 201$) における標本の記述統計

変数	平均	標準偏差	歪度	尖度	最小値	第三四分位数	中央値	第一四分位数	最大値
平均正答率	.71	.11	-2.00	5.34	.25	.63	.73	.79	.88
平均規則反応率	.63	.23	-0.44	-0.59	.02	.46	.67	.81	1.00
平均反応時間	8.40	0.68	0.34	0.55	6.68	7.93	8.38	8.84	11.03

注：反応時間の単位は秒である

表 5.2

項目レベル ($K = 48$) における標本の記述統計

変数	平均	標準偏差	歪度	尖度	最小値	第三四分位数	中央値	第一四分位数	最大値
平均正答率	.71	.17	-1.08	0.38	.22	.63	.78	.82	.94
平均規則反応率	.63	.11	-0.39	-0.21	.33	.55	.64	.69	.83
平均反応時間	8.40	0.96	0.46	-0.38	6.55	7.67	8.25	9.07	10.81

注：反応時間の単位は秒である

¹ この値は、個人が主観的変数において規則を選択した割合である。

次に、個人レベル ($n = 201$) および、項目レベル ($K = 48$) における平均正答率、平均規則反応率、平均反応時間の相関係数行列を表 5.3 に示す。これらの変数の散布図行列は、図 5.2 で示す通りである。

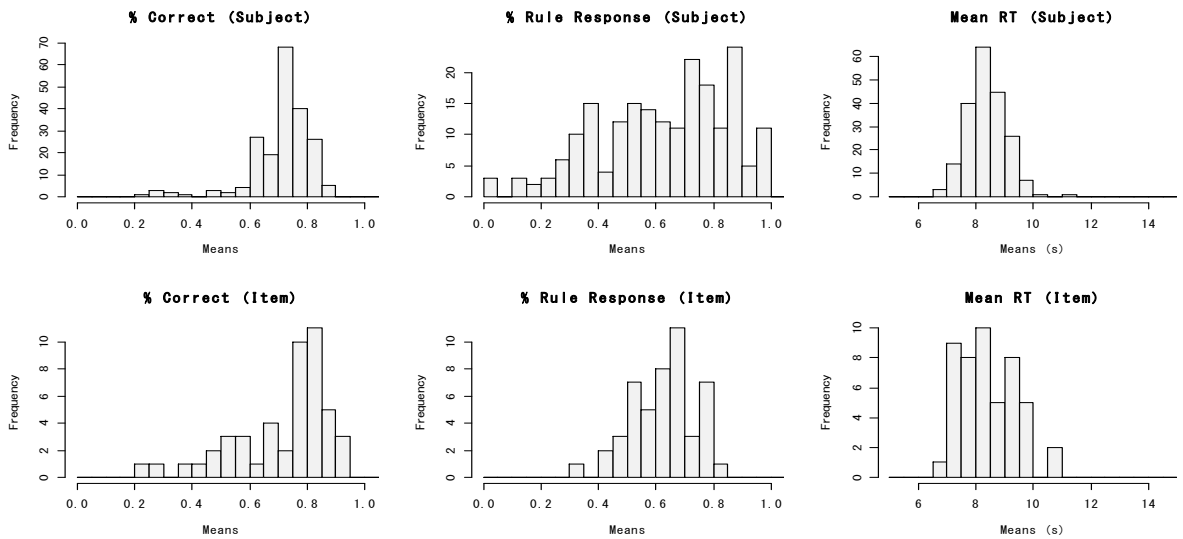


図 5.1 三変数の観測における度数分布

上段左の図は、個人レベルにおける平均正答率、上段中央の図は平均規則反応率、上段右の図は平均反応時間を秒単位で示している。下段はそれらに対応する項目レベルの値である。

表 5.3

個人レベルと項目レベルにおける三変数の相関係数行列

	個人レベル ($N = 201$)		項目レベル ($K = 48$)	
	平均規則反応率	平均反応時間	平均規則反応率	平均反応時間
平均正答率	.18	-.06	.53	-.25
平均規則反応率	1.00	.18	1.00	.53
平均反応時間	-.06	1.00	-.25	1.00

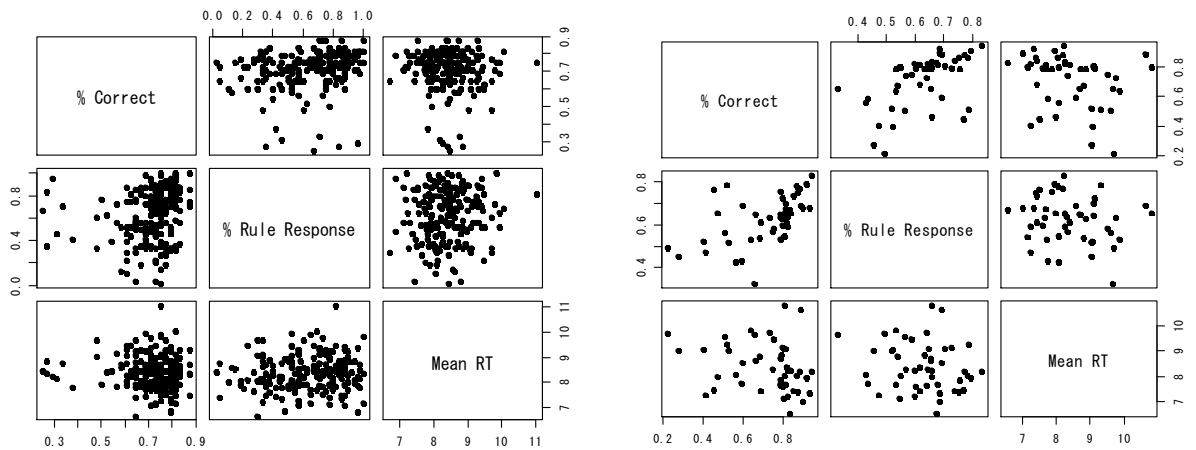


図 5.2 実験参加者レベル（左）と項目レベル（右）における三変数の散布図行列

次に、データの階層性を考慮し、一段階低いデータ階層である、試行レベルにおける観測の記述統計を検討した。

最初に、判断結果について、その正答数は 6,892 試行（71%）であり、誤答数は 2,756 試行（29%）であった。これを、切片 β_0 のみをもつロジスティック回帰モデルとして考え、切片、つまり β_0 の値についてベイズ推定することとした。推定方法は MCMC であり、ベイズ統計の方針にしたがって、各母数における事後分布のサンプルから、ベイズ信用区間 (credible interval)² を構築する。なお、信用区間の構築には、事後分布のサンプルにおいてパーセンタイル法を使用し、その水準を $\alpha = .05$ とした。以降も信用区間を構築する際は同様である。その結果、 $\beta_0 = .92 [.87, .96]$ となった。ここから、ロジスティック回帰モデルである、

$$P_i = \frac{1}{1 + e^{-\beta_0}} \quad (5.1)$$

に対して β_0 を代入すると、正答率の信用区間を得ることができる。その結果、 $P_i = .71 [.71, .72]$ となった。図 5.3 に、事後分布から得られたサンプルにおけるトレース図とカーネル密度推定³による経験分布を示す。ここでは、 P_i に変換している⁴。

² 信用区間という用語は、ベイズ統計における慣習である。これは、母数が分布をなすと考えることができない頻度主義との対比の中で使用される表現である。ベイズ信用区間の構築には、最高密度区間 (HDI) を使用する方法もあるが、ここでは、もっとも単純な方法であるパーセンタイル法を使用している。

³ カーネル密度推定とは、データの分布を推定するための方法のひとつである。あるデータが与えられたとき、そのデータの分布形状の概観を可視化するのに適している。MCMC を使用した推定によって、得られたサンプルをカーネル密度推定によって可視化し、事後分布の概観を掴むのはベイズ統計における手法のひとつである。

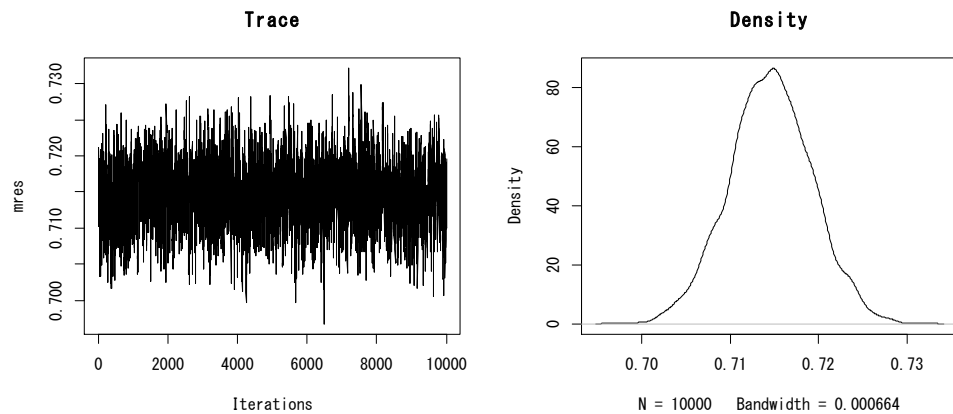


図 5.3 正答率に関するトレース図と事後分布から得られたサンプルの分布

次に、主観的変数について述べる。規則反応数は 6,036 試行 (63%) であり、一方の、直観反応数は 3,609 試行 (37%) であった。上記と同様に、規則反応率についてベイズ推定を行い、その信用区間を構築した。その結果、 $P_i = .63 [.62, .64]$ となった。図 5.4 に、事後分布から得られたサンプルにおけるトレース図とカーネル密度推定による分布を示す⁵。

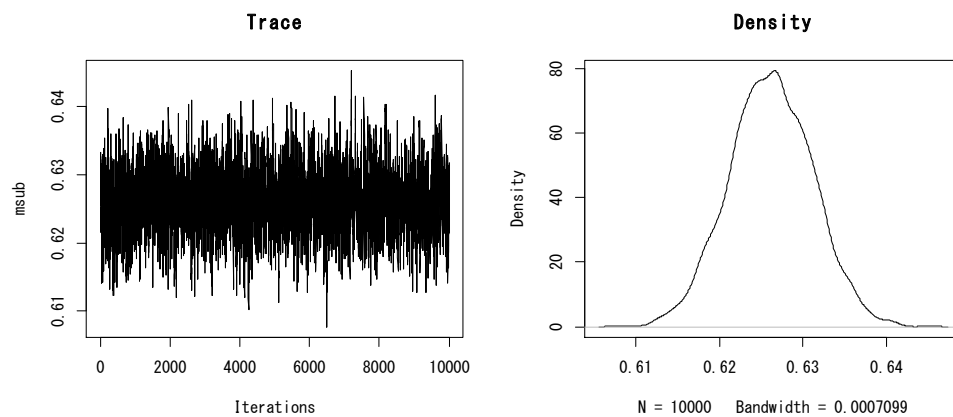


図 5.4 規則反応率に関するトレース図と事後分布から得られたサンプルの分布

⁴ ここでは、(5.1) によって、ロジスティック回帰モデルによって推定された切片の値を、解釈が容易な確率に変換している。
つまり、ここでの P は、正答率である。

⁵ ベイズ統計では、このように推定母数を分布としてもとめ、これを検証する。左側の図はトレース図と呼ばれるマルコフ連鎖モンテカルロ法によるシミュレーション結果を、右の図は、対象となる母数の事後分布とみて差し支えない。

最後に、試行レベルの反応時間分布について記述する。反応時間データは、正規分布に従わないことが経験的にも明白であるため、指数正規合成分布を観測の反応時間に対してフィットさせることとした。なお、推定はMCMCを使用した。その結果、 $\mu = 3.83$ [3.74, 3.93], $\sigma = 1.72$ [1.64, 1.79], $\tau = 4.56$ [4.43, 4.69]という推定結果が得られた⁶。MCMCによって得られたトレース図と事後分布から得られたサンプルの分布とを図 5.5 に示す。また、図 5.6 では、観測の度数分布と MCMC によって推定された母数による理論分布を比較している。この理論モデルは、観測の反応時間に対して、非常に優れた数理的近似となっている⁷ことがわかる。

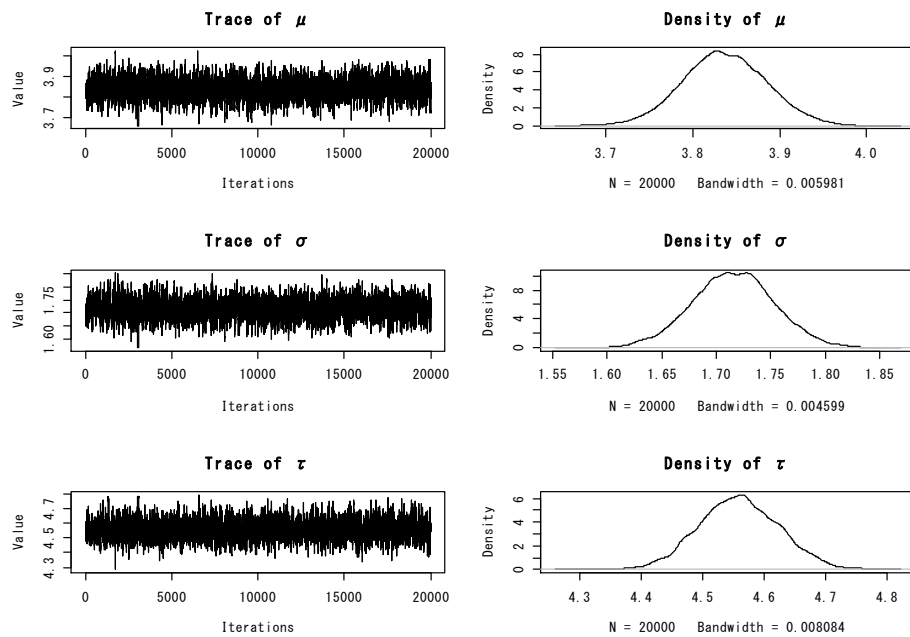


図 5.5 反応時間の母数に関するトレース図と事後分布から得られたサンプルの分布

⁶ 指数正規合成分布は、3 母数の確率分布であり、これらはそれらの母数である。一般に、 μ は母平均、 σ は母標準偏差、 τ は分布の歪みを形成する母数であると解釈される。

⁷ ここでは、図 5.6 のように、可視化によって判断を与えているが、たとえば、競合する分布のモデル (e.g., 対数正規分布、ガンマ分布) とその適合度を比較したとしても、指数正規合成分布がもっとも高い適合度を示すことを確認している。さらに指数正規合成分布は、三つの母数をもつ分布であるため、分布の記述により適している。

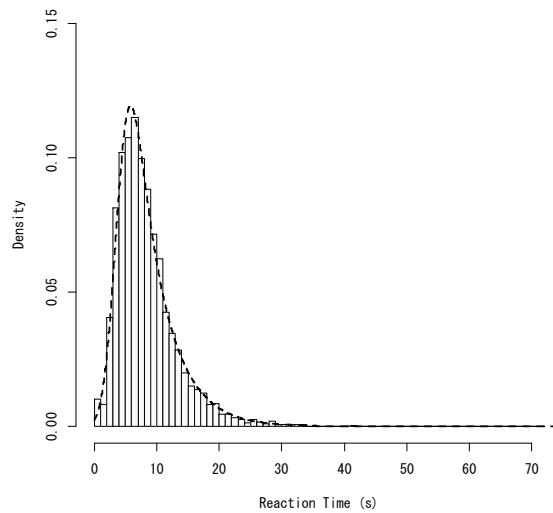


図 5.6 反応時間における度数分布と推定された母数による理論分布（全試行）

また、拡散過程モデルは、正答の反応時間分布と、誤答の反応時間分布を同時にモデリングするものであるため、正答の反応時間分布と誤答の反応時間分布を区別して、それぞれの母数を求めることも有益である。そこで、上記と同じ方法によって、それぞれの母数を推定した。その結果、正答の反応時間分布では、 $\mu = 3.75$ [3.64, 3.85], $\sigma = 1.57$ [1.49, 1.66], $\tau = 4.43$ [4.29, 4.59], 誤答の反応時間分布では、 $\mu = 4.10$ [3.91, 4.30], $\sigma = 2.06$ [1.91, 2.22], $\tau = 4.84$ [4.59, 5.10] という結果が得られた。それぞれの反応における反応時間の度数分布に、推定された母数の条件下での理論分布を描き足したものを図 5.7 に示す。また、それぞれの母数における事後分布から得られたサンプルの分布を図 5.8 に示す。図 5.8 における (a) と (b), (c) と (d), (e) と (f) の比較より、いずれの母数も、誤答における反応時間の方が大きな値を取ることがわかった。

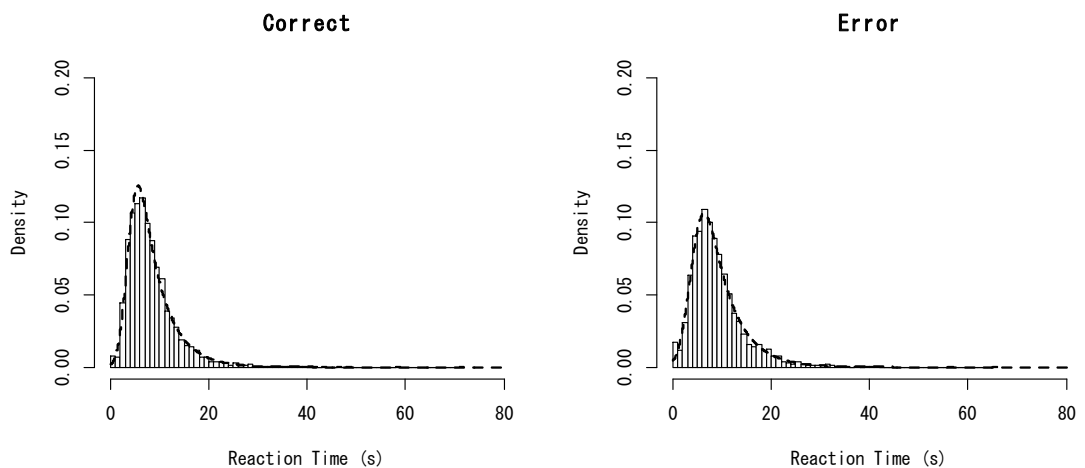


図 5.7 反応時間の度数分布と推定された理論分布（正答と誤答別）

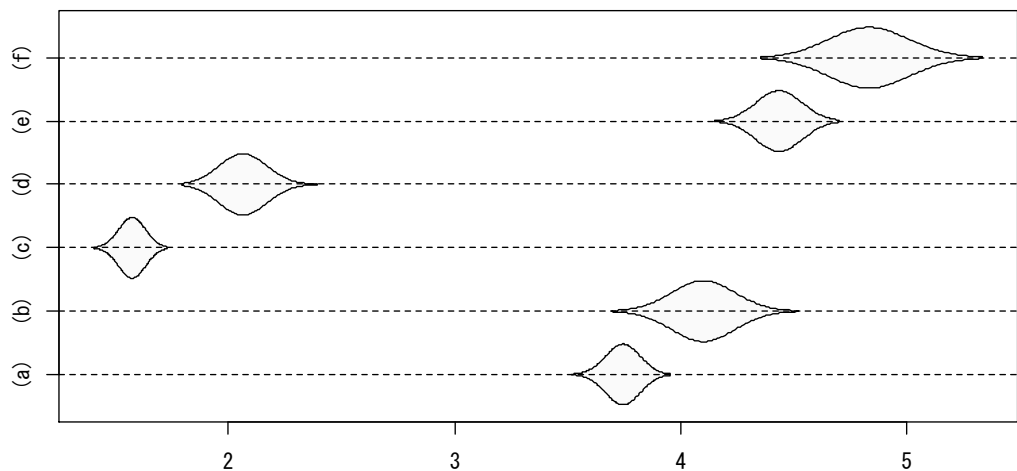


図 5.8 反応時間に関する母数における事後分布のサンプルから得られた分布

a = 正答反応の μ , b = 誤答反応の μ , c = 正答反応の σ ,
d = 誤答反応の σ , e = 正答反応の τ , f = 誤答反応の τ

さらに、反応時間は、研究仮説 2 の検証において説明変数となる。そのため、上記のような分布の歪みは、統計解析上望ましいものではない。そこで、研究仮説 2 について解析を行なう際には、この反応時間分布を対数変換することとした。対数変換によって得られた値は、おおよそではあるが、 $N(1.96, 0.63)$ の正規分布に従った。対数変換された反応時間分布と、その理論分布は図 5.9 のとおりである。なお、このときの信用区間は、対数変換後の平均において、 $[1.95, 1.97]$ 、標準偏差において、 $[0.62, 0.64]$ である。この推定結果において、変換後における観測の分布が、理論分布よりもやや尖っていることが見て取れる。しかし、中心傾向と散布度において大きなずれが見られないことから、ここでの対数変換は、統計的に妥当な処理であるとみなすこととした。

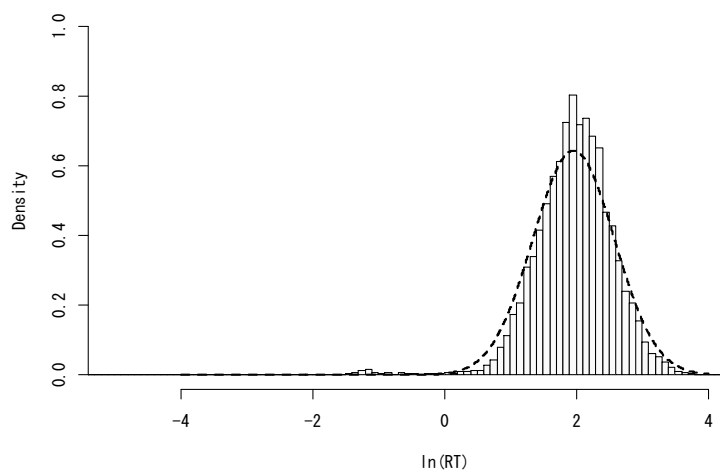


図 5.9 対数変換された反応時間の度数分布と推定された理論分布

5.2 研究仮説 1 : 拡散過程モデル

研究仮説 1 を検証するために、観測の正答率、正答の反応時間分布、そして誤答の反応時間に対して拡散過程モデルをフィットさせた。ここでは、ランダムに実験参加者 ($n = 201$) をグルーピングし、そのデータをプール化した上で、10 および 11 人からなる大被験者データを 20 個構築している。

それら 20 個の大被験者データに対し、それぞれ拡散過程モデルのフィットを試みた。その結果として、推定された拡散過程モデルの母数と対数尤度は表 5.4 の通りである。

表 5.4

拡散過程モデルにおける推定母数と対数尤度

	a	v	T_{er}	η	S_t	対数尤度
大被験者 1 ($n = 10$)	5.15	0.35	0.68	0.48	3.60	-1,511.91
大被験者 2 ($n = 10$)	5.57	0.31	0.66	0.32	4.71	-1,606.19
大被験者 3 ($n = 10$)	4.98	0.38	0.79	0.33	5.38	-1,529.19
大被験者 4 ($n = 10$)	4.96	0.31	0.74	0.27	5.66	-1,566.36
大被験者 5 ($n = 10$)	4.89	0.24	1.17	0.38	2.60	-1,512.14
大被験者 6 ($n = 10$)	5.69	0.33	0.62	0.37	5.52	-1,619.05
大被験者 7 ($n = 10$)	5.40	0.44	0.46	0.52	3.87	-1,498.67
大被験者 8 ($n = 10$)	5.64	0.47	0.75	0.52	4.71	-1,546.88
大被験者 9 ($n = 10$)	5.20	0.32	0.57	0.31	5.92	-1,596.69
大被験者 10 ($n = 10$)	5.40	0.20	0.81	0.18	2.44	-1,619.69
大被験者 11 ($n = 10$)	5.01	0.35	0.90	0.29	5.28	-1,538.61
大被験者 12 ($n = 10$)	5.41	0.36	0.72	0.24	6.16	-1,616.11
大被験者 13 ($n = 10$)	5.37	0.30	0.72	0.33	2.54	-1,593.04
大被験者 14 ($n = 10$)	5.20	0.36	0.91	0.36	4.64	-1,533.42
大被験者 15 ($n = 10$)	5.63	0.46	0.72	0.73	5.96	-1,567.62
大被験者 16 ($n = 10$)	5.33	0.22	0.75	0.21	1.62	-1,586.93
大被験者 17 ($n = 10$)	5.29	0.38	0.82	0.46	3.84	-1,511.70
大被験者 18 ($n = 10$)	5.73	0.31	0.59	0.25	3.71	-1,614.75
大被験者 19 ($n = 10$)	4.67	0.26	1.04	0.29	3.45	-1,501.88
大被験者 20 ($n = 11$)	5.01	0.41	1.21	0.55	0.00	-1,621.98

注 : $a =$ 境界母数, $v =$ 漂流比率, $T_{er} =$ 非決定時間, $\eta =$ 漂流比率のばらつき, $S_t =$ 非決定時間のばらつき。開始時点のばらつきである S_z は推定の結果、すべての標本で 0 の値を取ったため、この表では省略している。

また、これら、各大被験者における推定値の五数要約⁸については、表 5.5 が示している。大被験者群における中央値である、 $a = 5.31$, $v = 0.34$, $T_{er} = 0.75$, $\eta = 0.33$, $S_t = 1.62$, そして $S_z = 0.00$ は、標本全体における正答率、正答の反応時間、誤答の反応時間を代表するものとなるはずである。

表 5.5

拡散過程モデルにおける推定母数の五数要約

	最小値	第三四分位数	中央値	第一四分位数	最大値
a	4.67	5.01	5.31	5.45	5.73
v	0.20	0.31	0.34	0.38	0.47
T_{er}	0.46	0.68	0.75	0.84	1.21
η	0.18	0.29	0.33	0.47	0.73
S_t	1.62	3.57	4.68	5.45	6.16

このようにして得られた母数の値による拡散過程モデルが、観測に対して優れた数理的近似になっているかを確認するために、拡散過程モデルによる理論上の確率密度曲線と、指数正規合成分布による確率密度曲線を比較した (図 5.10)。この図からも分かるように、このモデルは、正答率、正答における反応時間分布、誤答の反応時間分布に非常によく近似していることが見て取れる。

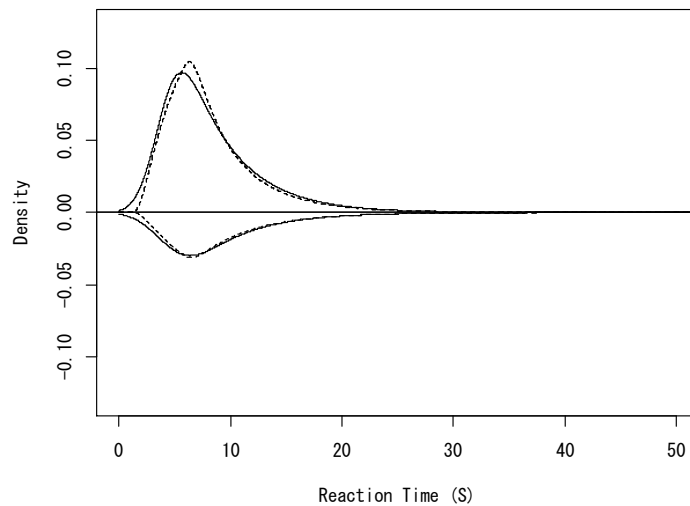


図 5.10 拡散過程モデルによる理論分布と指数正規合成分布による理論分布の比較

注：実線が指数正規合成分布、破線が拡散過程モデルによる確率密度曲線を表している

⁸ これらの推定値の全てが正規分布に従うとは考えにくいので、ここでは経験的な分位点によって要約している。五数要約は、最小値、第三四分位数、中央値、第一四分位数、最大値のことである。

次に、推定母数の条件下における拡散過程モデルから、実測値に等しい観測数の乱数を発生させ、実測値との比較を行った。表 5.6 は、それぞれのデータにおける、正答の反応時間における分位点と誤答の反応時間における分位点を示している。表 5.6 が示すように、それぞれの分位点において、実測値とシミュレーションは非常に近い値を示している。なお、実測値の正答率は.71、誤答率は.29、シミュレーションにおける正答率は.74、誤答率は.26 であり、その誤差は 3%以内であった。図 5.11 は、実測値とシミュレーションにおける分位点—分位点プロット (Q-Q plot) である。この図も実測値とシミュレーションにおける一致度の高さを示している。

表 5.6

実測値とシミュレーションにおける分位点の比較

		2.5%	25%	50%	75%	97.5%
実測値	正答	2.94	5.18	7.27	10.29	20.38
	誤答	3.06	5.72	8.01	11.22	22.70
シミュレーション	正答	2.76	5.31	7.17	9.91	20.00
	誤答	3.01	5.74	8.01	11.60	23.38

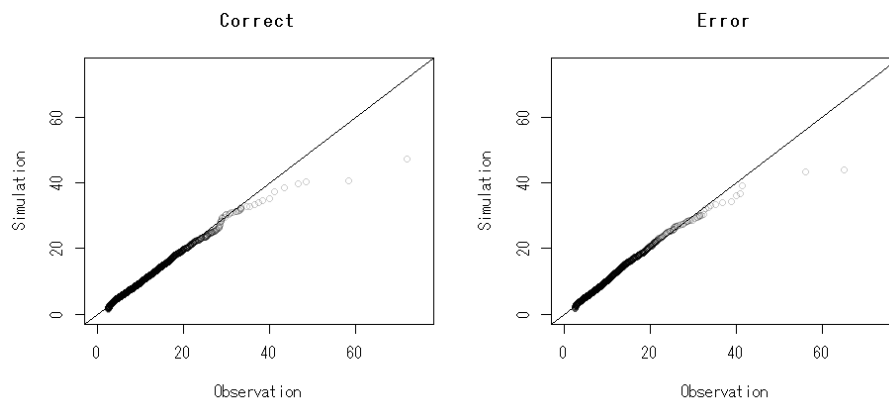


図 5.11 実測値とシミュレーションにおける分位点—分位点プロット

これらの観察は、外国語における GJT から生成される観測、すなわち、正答率、正答の反応時間分布、誤答の反応時間に対して、拡散過程モデルが非常に優れた近似を示すという証拠のひとつとなる。すなわち、明示的および暗示的知識が想定するような複雑な構造や、それに由来するいかなる合理的仮定を置かなくとも、一般的な判断行動を伴う他課題における莫大な例とまったく同様に、外国語における文法性判断は、拡散過程モデルが想定するように、単一的な証跡ないし表象の質といった、より簡潔な概念のみによって十分に説明できる可能性があるということである。

ここから、本論の研究仮説 1, すなわち,

研究仮説 1: 拡散過程モデルは外国語における GJT の正答率, 正答の反応時間分布, 誤答の反応時間分布に対して優れた数理的近似となる

は, 支持されたと判断することができよう。

5.3 研究仮説 2: 主観的変数に対する判断結果と反応時間の影響

研究仮説 2 を検証するために, MCMC を推定アルゴリズムとして使用した階層ベイズモデルの解析を行った。前章で述べたように, ここでの数理モデルは, (5.2) のとおりである。

$$\ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{1,i} + \beta_2 \log(x_{2,i}) + r_j \quad (5.2) = (4.13)$$

これは, 主観的変数における規則反応 P_i を応答変数とし, 説明変数を判断結果 x_1 , および反応時間 x_2 とする, ロジスティック回帰モデルである。ただし, 個人におけるそれぞれの説明変数の傾きと切片に対する変量効果も想定している。推定方法や, 各推定母数の事前分布については, 前章の通りである。本論では, 以下の母数の事後分布を総合的に検討することとした。

- (a) 試行レベルの切片: β_0
- (b) 試行レベルにおける判断結果の回帰係数: β_1
- (c) 試行レベルにおける対数変換された反応時間の回帰係数: β_2
- (d) 個人レベルにおける切片の分散: $V(\beta_0)$
- (e) 個人レベルにおける切片と判断結果の回帰係数をもつ共分散: $Cov(\beta_0, \beta_1)$
- (f) 個人レベルにおける切片と対数変換された反応時間の回帰係数をもつ共分散: $Cov(\beta_0, \beta_2)$
- (g) 個人レベルにおける判断結果の回帰係数をもつ分散: $V(\beta_1)$
- (h) 個人レベルにおける判断結果の回帰係数と対数変換された反応時間の回帰係数をもつ共分散: $Cov(\beta_1, \beta_2)$
- (i) 個人レベルにおける対数変換された反応時間の回帰係数をもつ分散: $V(\beta_2)$

なお、MCMCによる推定にあたって、Gewekeの収束診断とトレース図の検証を行い、各母数について定常分布⁹に収束したと判断した。表5.7は、それぞれの母数における事後分布から得られたサンプルの平均値、標準偏差と分位点¹⁰を示している。図5.12は、固定効果である三母数 $\beta_0, \beta_1, \beta_2$ についてのトレース図と事後分布から得られたサンプルの分布をカーネル密度曲線で表したものである。

表 5.7

階層ベイズモデルにおける各母数の事後分布から得られたサンプルの要約

レベル	母数	平均値	標準偏差	2.5%	25%	50%	75%	97.5%
試行（固定効果）	β_0	9.35	0.51	8.44	8.98	9.34	9.67	10.45
	β_1	0.70	0.05	0.61	0.66	0.71	0.74	0.80
	β_2	-1.02	0.06	-1.14	-1.05	-1.02	-0.98	-0.91
個人（変量効果）	$V(\beta_0)$	74.26	12.57	50.04	66.04	73.21	82.00	101.01
	$Cov(\beta_0, \beta_1)$	2.19	0.60	1.10	1.61	2.08	2.56	3.25
	$Cov(\beta_0, \beta_2)$	-7.73	1.29	-10.05	-8.56	-7.65	-6.88	-5.19
	$V(\beta_1)$	0.07	0.03	0.03	0.05	0.07	0.10	0.13
	$Cov(\beta_1, \beta_2)$	-0.22	0.06	-0.39	-0.27	-0.22	-0.17	-0.13
	$V(\beta_2)$	0.82	0.13	0.55	0.73	0.81	0.91	1.10

⁹ 定常分布とは、ある一定の範囲にシミュレーションが収束している度合いを示す。シミュレーションが収束していなければ、その推定は信用できないものである。その診断方法のひとつが Geweke の収束診断であり、トレース図は、その収束を視覚的に把握するための方法のひとつでもある。

¹⁰ ベイズ統計では、事後分布から得られたサンプルの平均値を事後平均値（expected a posterior）と呼ぶ。ここでの標準偏差は頻度主義下における標準誤差の概念と同様である。また、事後分布の値についての分位点を検討することも重要な方法である。たとえば、ここでの β_0 は 9.35 を中心に分布していると考えられることができ、97.5%程度の確率で、10.45 よりも小さい値を取るだろうと推論することができる。

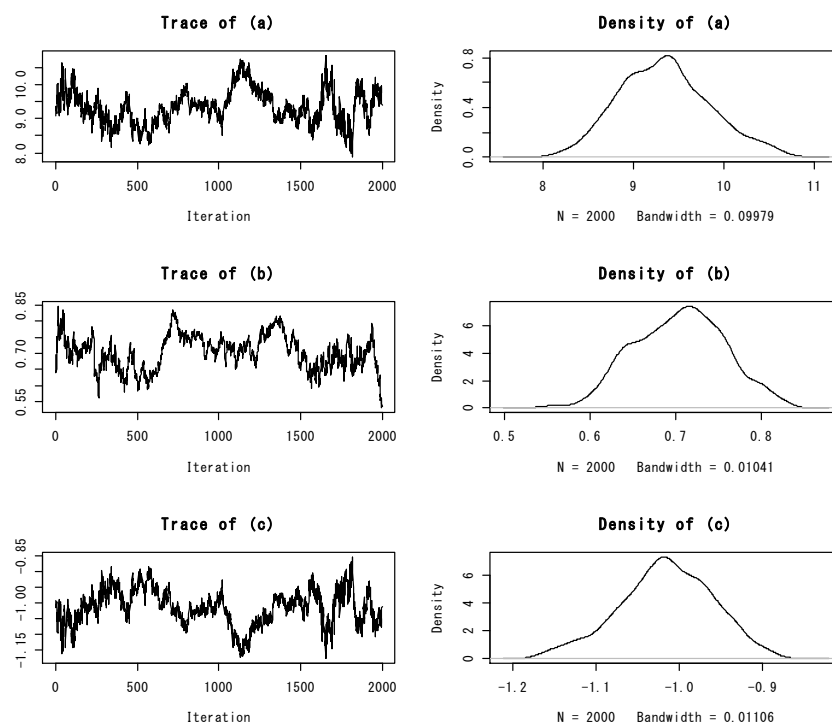


図 5.12 固定効果における母数のトレース図と事後分布から得られたサンプルの分布

ここから、 β_1 、すなわち、判断結果の回帰係数における信用区間が $[0.61, 0.80]$ であり、かつ、 β_2 、すなわち、反応時間の回帰係数における信用区間が $[-1.14, -0.91]$ であるといった、原点を超えない区間を取ることからも明白なように、研究仮説 2 の含意である、

$$(\beta_1 > 0) \wedge (\beta_2 < 0) \quad (5.3) = (4.2)$$

が真であると判断することができよう。このことは、判断結果が正答であれば、規則反応を示す確率が上昇し、その上、反応時間が短かければ、同確率が同様に上昇するというを示している。

次に、変量効果の母数について検討する。図 5.13 に、変量効果の母数についての事後分布から得られたサンプルの分布を示す。ここではトレース図を省略している。切片における変量効果の大きさの信用区間は、分散のスケールにおいて $[50.04, 101.01]$ である¹¹ のだから、標準偏差のスケールでの信用区間は、 $[7.07, 10.05]$ であり、その平均値は 8.62 である。個人レベルの切片に関する予測では、固定効果と変量効果の平均値から簡便に計算すると、その 95% 予測区間は、 $[-7.54, -26.25]$ 程度となる。次に、判断結果がもつ回帰係数における変量効果の大きさの信用区間は、分散のスケールにおいて $[0.03, 0.13]$ 、標準偏差のスケールにおいて、 $[0.17, 0.36]$ であり、その平均値は 0.26 である。ここから、

¹¹ 推定された値が分散であるため、標準偏差のスケールに変換している。

個人レベルにおける判断結果がもつ回帰係数について予測すると、その 95% 予測区間は、[0.19, 1.21] 程度となり、この区間は明らかに原点 0 を超えない。よって、判断結果の回帰係数が負の値を取る個人は、相対的にかなり少数であるということがわかる。同様に、対数変換された反応時間がもつ回帰係数における変量効果の大きさの信用区間は、分散のスケールにおいて [0.55, 1.10]、標準偏差のスケールにおいて、[0.74, 1.04] であり、その平均値は 0.91 である。個人レベルにおける判断結果がもつ回帰係数について、この情報から予測すると、その 95% 予測区間は、[-2.61, 0.60] 程度となり、この区間は明らかに原点を超えることから、判断結果がもつ回帰係数とは異なり、反応時間の回帰係数において、正の値を取る個人が一定数存在しうることが予測できる。つまり、個人によって、反応時間が長くなれば、規則反応を示す傾向をもつ場合があるということである。

なお、それぞれの変量効果の関係について観察すれば、高い切片をもつ個人は、判断結果の回帰係数において大きな値を示し、さらに反応時間の回帰係数においても負方向に大きな値を取る傾向が見て取れる。また、同様に、判断結果の回帰係数において高い値を示す個人は、反応時間の回帰係数において負方向に大きな値を示す傾向も把握できる。

この階層ベイズモデルの予測精度について検討することも重要である。まずは、当該のモデルにおける予測値を図 5.14 に示す。この図からも、視覚的に、 β_1 が正の値を取り、 β_2 が負の値を取ることがわかる。

次に、的中率と感度¹²を評価するために、実測値と予測値の混同行列を表 5.8 に示す。表 5.8 から見取れるように、このモデルの的中率は、67%¹³であり、感度は 79%¹⁴である。説明変数の量を考慮すると、十分に優れた予測であるといえる。

¹² 的中率とは、混同行列において、実測値とモデルによる予測値が一致している割合を示す。ここでの感度は、実測値が規則反応である試行のうち、モデルが規則反応であると予測する試行の割合である。

¹³ 表 5.8 の通り、全試行は、9648 試行であり、このうちの 6502 試行（5122 試行+1380 試行）が的中である。

¹⁴ 実測値において規則反応である 6498 試行（5122 試行+1376 試行）のうち、5122 試行が的中である。

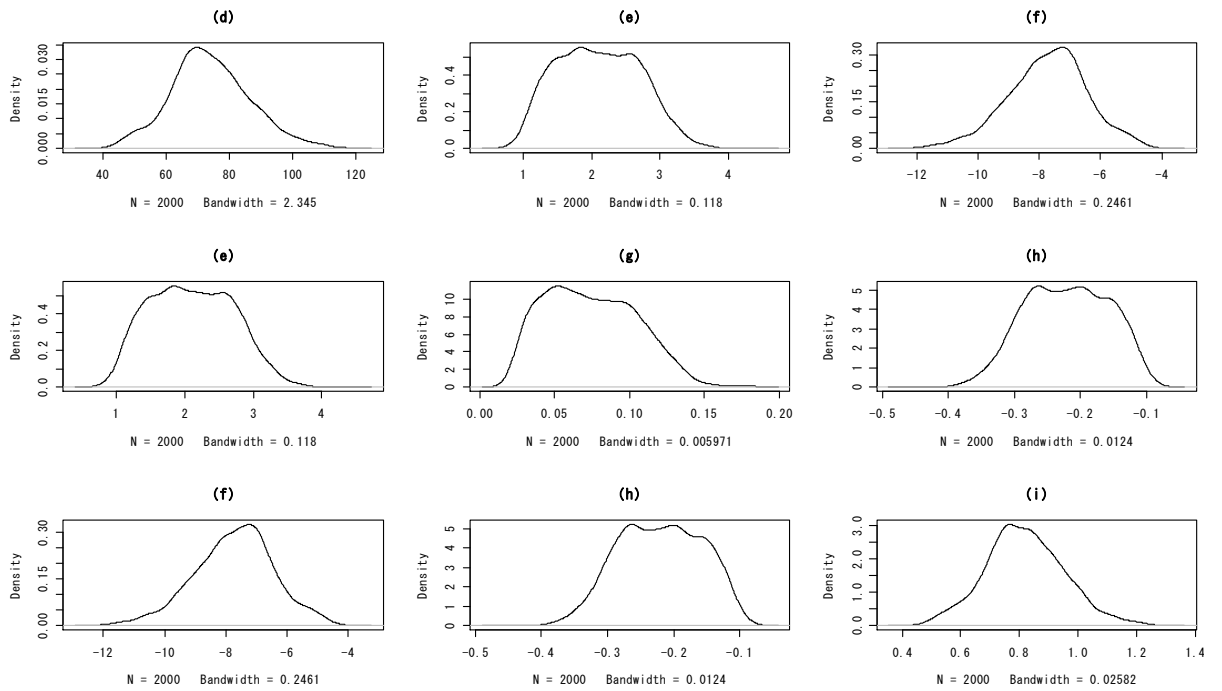


図 5.13 変量効果における母数の事後分布から得られたサンプルの分布

注：図 (d), (g), (i) はそれぞれ β_0 , β_1 , β_2 の分散の分布を示しており、

図 (e) は β_0 と β_1 がもつ共分散、図 (f) は β_0 と β_2 がもつ共分散、

図 (h) は β_1 と β_2 がもつ共分散の分布である。

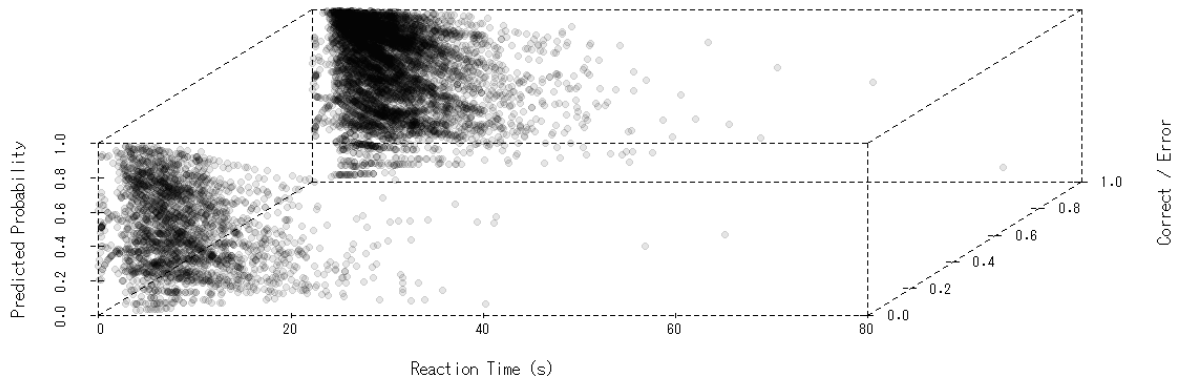


図 5.14 モデルによる予測値を示す三次元グラフ

注：x 軸は反応時間、z 軸は判断結果（1 = 正答, 0 = 誤答）を、y 軸がモデルによる予測値を示している。

正答の場合に高い規則反応率を示し、反応時間がより長くなれば規則反応率が低減していくことがわかる。

表 5.8

実測値とモデルによる予測の混同行列

		実測値 規則反応	実測値 直観反応
モデル 規則反応		5,122 (.53)	1,770 (.18)
モデル 直観反応		1,376 (.14)	1,380 (.14)

また、このモデルを使用し、実験参加者における平均規則反応率をもとめ、実測の平均規則反応率と比較した。その結果を図 5.15 に示す。モデルから予測した実験参加者における平均規則反応率と、実測値の平均規則反応率の相関係数は、 $r = .92$ であり、決定係数は、 $R^2 = .83$ であった。このことから、当該のモデルは、個人レベルにおいて非常に優れた予測結果を示すといえる。すなわち、判断結果と反応時間は、規則反応を十分に予測する変数といえる。

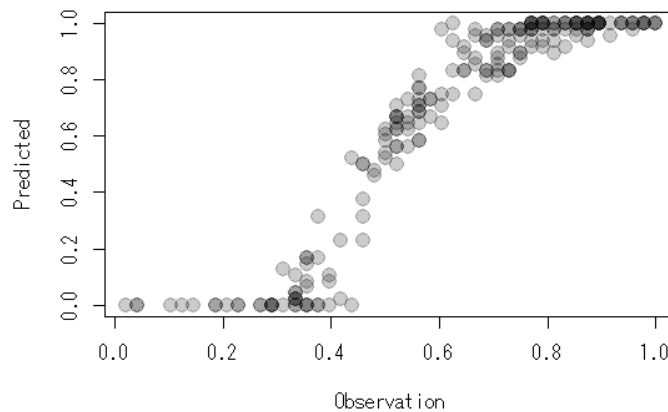


図 5.15 個人レベルにおける平均規則反応率の比較

5.4 第五章のまとめ

本章では、一元観についての実証的根拠を示す実験結果について報告した。その結果、以下の二つの研究仮説、すなわち、

研究仮説 1：拡散過程モデルは外国語における GJT の正答率、正答の反応時間分布、誤答の反応時間分布に対して優れた数理的近似となる

および、

研究仮説 2：試行が規則反応である確率は、判断結果および反応時間の影響を受ける

より具体的にいえば，ロジスティック回帰モデルの上で，

$$(\beta_1 > 0) \wedge (\beta_2 < 0) \quad (5.3)$$

この条件がもっともらしい仮説であることを確認した。すなわち，拡散過程モデルは，実測の文法性判断における正答率と反応時間を適切に予測し，さらに，判断結果が正答であれば，規則反応である確率が高く，反応時間が長くなれば，規則反応である確率が高いということである。

次章の総括では，これらの実証結果の解釈，結果について想定されうる反論と，その反駁，最後に，一元観がもたらすであろう効用について触れ，本論全体を総括する。

第六章 総括

6.1 結果の総括と解釈

本章では、第四章および第五章にて述べた実験結果を振り返り、本論全体の総括を行なう。第五章で述べたように、第三章にて提案した一元観に関する基礎的な証拠となる、二つの研究仮説は、実証データより支持されたものと判断した。

最初に、研究仮説 1, すなわち、

研究仮説 1: 拡散過程モデルは外国語における GJT の正答率, 正答の反応時間分布, 誤答の反応時間分布に対して優れた数理的近似となる

について、その解釈を述べる。拡散過程モデルは、本論において、これまでも繰り返し述べたきたように、一元観、すなわち一元的な証拠や表象の質が、さまざまな観測の原因になる、というコンセプトをもつ数理モデルの代表例である。よって、このモデルが、外国語における GJT より得られる観測に対して優れたフィットを示すということは、明示的知識および暗示的知識といった複表象、ないし多元的で複雑な判断行動に関する仮定を置かずとも、当該の現象が少なくとも数理的に説明可能である、ということのいくばくかの証拠になる。

たとえば、これは、R. Ellis (2004) が仮定するような、順次作動的な複数のプロセス、すなわち、(a) 意味処理、(b) 気づき、(c) 内省といった、直接的観測が不可能な認知的機構を採用する積極的な動機がない、ということになる。それに合わせ、意味処理を担うとされる暗示的知識、そして、内省を担うとされる明示的知識といった概念についての想定も、理論上不要なものとなる。

元来、この上記の R. Ellis による認知モデルは、正答率や反応時間分布などに対する具体的な予測をもたないものであるが、仮に、異種の認知機構による行動の結果が、観測において顕在化するのであれば、反応時間を変数とした混合分布モデルなどによって実証されるべきであった。しかしながら、これは、これまでに実証されていない。

むしろ、第五章において、本論が示した観測の反応時間分布も、これまでの先行研究と同様に、明らかな多峰性を示すわけではなかった。さらに、指数正規成分分布や拡散過程モデルが反応時間に対して優れたフィットを示すことは、混合分布を仮定せざるを得ないような、非常に複雑な認知モデルを廃する動機として十分である。

また、拡散過程モデルが、外国語における GJT より得られる観測の優れた数理的近似になるということは、正答率や正答および誤答の反応時間分布が、漂流母数または境界母数といった変数を仲介することで、以下に示すような共変関係にあるということを示

す。すなわち、観測の正答率が高ければ、必然的に反応時間分布の期待値も短くなる、といった関係である。このことから、あくまでも概念的にはあるものの、拡散過程モデルにおける漂流母数といった変数が、正答率や反応時間に対する因果関係の原因のひとつであると、モデル上において推測することは妥当であろう。

次に、研究仮説 2 についての解釈を示す。研究仮説 2 は、

研究仮説 2 : 試行が規則反応である確率は、判断結果および反応時間の影響を受ける

というものであった。具体的には、

$$\ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{1,i} + \beta_2 \log(x_{2,i}) + r_j \quad (6.1) = (4.13)$$

という数理モデルの上で、少なくとも、

$$(\beta_1 > 0) \wedge (\beta_2 < 0) \quad (6.2) = (4.2)$$

が真であるということであった。これも第五章で述べたように、それぞれの母数における事後分布を検討した結果、その信用区間が明らかに原点を超えないことから、(6.2) の命題は真であると判断されよう。

つまり、観測の判断結果が正答であれば、主観的変数が規則反応である確率が上昇し、そして、観測の反応時間が短ければ、同種の確率が上昇する、ということである。この結果は、従来の認知心理学や数理心理学における数多くの研究実践、そして外国語教育研究における草薙・川口 (2015) などの結果と一致している。繰り返しになるが、R. Ellis による明示的知識および暗示的知識の枠組みでは、直観反応は暗示的知識を示すものであるとされ、暗示的知識は母語話者のような短い反応時間の原因となるとされているのであるから、反応時間が短ければ、主観的変数が直観反応である確率が上昇する、という観測に至るはずである。しかし、この予測は、本論がこれまでにその反例を示したように、必ずしも真であるとはいえない。繰り返しになるが、R. Ellis (2005) にて示された予測は、あくまでも構造の優先によって、実証を経ぬまま、それぞれの概念に特性を割り当てられた所以であると考えられる。

研究仮説 2 の結果と研究仮説 1 の結果を組み合わせると、正答率や反応時間は、漂流母数などといったような一元的な母数の下において共変関係にあり、主観的変数は判断結果や反応時間の影響を受ける、と結論づけることができよう。やや間接的ではあるものの、これらの証拠の組み合わせによって、一元観が置くひとつの仮定、すなわち、潜

在変数であるところの証跡が、正答率、反応時間、そして主観的変数に影響を及ぼしているであろうことは、今回の実験結果から見て、もっともらしいことである。

このことから、以下のような一元観の含意が成り立つ。個人と項目の組み合わせに固有であるところの潜在変数、すなわち、証跡の値が高ければ、判断結果が正答である確率が高く、それと同時に反応時間が短く、そして、主観的変数が規則反応である確率が高くなる。同様に、証跡の値が低ければ、判断結果が誤答である確率が高く、反応時間が長く、そして主観的変数が直観反応である確率が高くなる。

これらの具体的な予測は、明示的知識および暗示的知識の二分法のような、それぞれ異なる課題や条件に対して、相対的に強い影響を及ぼすであろう異種の潜在変数が複数存在し、それらが、実証可能であるとも限らない、複雑な認知機構において特異的な役割を担う、と想定する従来のモデル (e.g., R. Ellis, 2005, 2006; R. Ellis et al., 2009) よりも、理論面においてかなり儉約的である。この点からも、本論が提案した一元観は、オッカムの剃刀やモーガンの公準といった一種の態度の下で、従来の明示的知識および暗示的知識の二分法よりも、より正当性があると評価できる。

もうひとつ、一元観において特筆すべきところは、証跡を個人と項目の組み合わせのみに固有であるとみなすことから、いかなる判断行動に関する課題条件ないし課題困難度、たとえば時間制限などに対して、証跡の値自体が独立であるとみなすことができる点である¹。これは、他種の、すなわち困難度の異なる課題を観測変数に置く構造的仮定とは、本質的に異なる姿勢である。一元観と、具体的な数理モデルであるところの拡散過程モデルは、必ずしも同一ではないが、拡散過程モデルにおいて、課題条件は、少なくとも理論的に、境界母数や開始時点に対してのみ影響を及ぼすと仮定されている。

これは、以下のような数値シミュレーションを考えればわかりやすい。拡散過程モデルでは、第五章で示したように、任意の母数を与えた擬似乱数を作成することも可能である。たとえば、第五章で推定した母数のセットにおいて、境界母数の値のみを変更した擬似乱数を作成することができる。ちなみに、実測の境界母数は、5.31 であったが、これを仮に 2.00 とした場合、シミュレーション上の正答率は 65%程度となり、本来の 77%程度から 12%程度減少する。同様に、反応時間は、当然より短いものになる。実際に、GJT に対して時間制限を設ける場合、時間制限を設けない場合と比べ、反応時間が短くなり、正答率が下がるということが報告されているが (e.g., Kusanagi, 2014; Loewen, 2009), このように、時間制限を設ける条件と時間制限を設けない条件の成績の差は、明示的知識および暗示的知識といった概念を殊更設けなくとも、十分に説明可能であるということである。

¹ 文法性判断における時間制限 (課題要因が正答率を引き下げる) は、課題に対応する知識が存在することを示すのではなくて、数理モデル上で規定される別の母数、たとえば SAT などによって規定される課題に由来する現象であるとみなすということである。

これらの解釈から、第二言語習得研究における根本的な研究テーマであった、学習者と母語話者間における文法運用の差、そして学習者の個人内における運用の変動は、明示的知識や暗示的知識といった概念を廃しても、そして、むしろそうすることによって、より明確なものになると考えられる。このことは、集団間における測定不変性や構造不変性についても同様であり、証跡の値を、課題条件と独立であるとみなすことによって、集団間の差についての検証もより容易になるだろう。

6.2 本論の限界

一元観は、上記のように優れた特性をもつものの、本論は以下のような限界点を持ち、さまざまな反論の余地がある。この節では、それぞれについて、本論に対する想定上の反論と、それに対する弁護を行う。

- (a) 数理的アプローチは認知的機構の存在や命題の科学的真偽を含意しないこと
- (b) 一般化可能性の限界と言語行動の限定性
- (c) 技術的問題点と潜在的交絡要因
- (d) 一元観の抽象性

(a) 数理的アプローチは科学的真偽や存在を含意しないこと

本論は、一元観という、外国語における文法知識に関する新しい認識論的基盤を提案することを目標とし、これは、特に第二言語習得研究において従来より広く受け入れられている明示的知識および暗示的知識の二分法と対立するものである。しかしながら、明示的知識および暗示的知識なる概念が、自然主義的に存在しないものであるとか、それに関連する認知機構におけるあらゆる存在の一切を否定するものでないことは、本論全体を通じて幾度か述べていることである。また、本論は、「外国語の文法知識が二つに分けられる」という命題が、科学的に偽であるという強い主張をなすものでもない。さらに、明示的知識および暗示的知識の二分法に依拠する、かつての莫大な先行研究の一切を、信憑性に欠くもの、として断じるつもりもない。それら、明示的知識および暗示的知識に関する先行研究の中には、いくつかの著者による研究も含まれている (e.g., 草薙, 2013; Kusanagi, 2014; Kusanagi & Yamashita, 2013; Kusanagi & Fukuta, 2015; 草薙・川口, 2015; Tamura & Kusanagi, 2015a, 2015b)。

そうではなく、本論が論じることは、数理的アプローチの上で、当該の二分法が、観測に対して著しく整合性を欠く、または、数理的近似を与えるという一種の目的の上で、単純に、理論的ないし概念的に殊更必要ではない、ということである。数理的アプローチは、あくまでも現象に対して数理的近似を与えるという目的の上で、ある概念が不要であることを示すことはできるが、その概念自体の存在を否定することや、なんらかの

命題について、その科学的真偽を直接的に判定することはできないし、そのようなことを研究目的としてもつわけではない。本論の実証部分が示したことは、いくつかの数理モデルが観測に対して優れた近似を示すということのみである。よって、明示的知識および暗示的知識という二分法が偽であり、一元観が真であると軽率に解釈されることは、本論の本意ではない。あくまでも、一元観は、実際に得られる文法運用の一部に関する観測に対して、優れた数理的近似を与えるための見方のひとつであり、従来のいかなる枠組みとも、少なくとも理論的には並立可能なものである。しかしながら、少なくとも帰結主義、経験主義、そして形式主義という研究姿勢に視野を限れば、一元観は、件の二分法に比べて優れた枠組みであると考えている。

ただし同時に、一元観は、ある種の認知機構の存在やそれを実装する神経的機構について、どのような想定も置かない。たとえば、拡散過程モデルは、判断行動の一部を、情報の集積過程に「喩えている」が、少なくとも本論では、このような過程が、認知主義的に、または神経科学的にもっともらしいであるとか、または証跡といった概念が、シナプスの連合パターンやその強度といった神経的ないし物理的基盤をもつ概念である、と主張することはない。むしろ、一元観は、繰り返しになるが、あくまでも、言語行動、そのなかでも刺激の文法性に関する判断行動から得られる観測を数理的に予測するための、より単純なモデリング・コンセプトのひとつである。

(b) 一般化可能性の限界と言語行動の限定性

本論の実験は、GJT を採用した過去の研究において、大標本の実験に分類されるものであるが、標本の大きさ自体は、標本のあらゆる種類の偏りをも否定するものではない。少なくとも、本論の標本は、本論が対象とみなすところの母集団に対して、比較的整合的であるものの、一般的な外国語教育研究ないし第二言語習得研究がこれまでに研究対象としてきた、あらゆる母集団に対して整合的であるわけではない。たとえば、(a) 年齢、(b) 言語環境、(c) 発達段階、(d) 母語、などといった要因を考慮した場合において、本論の実験結果が必ずしも再現されるわけではない。しかしながら、言語環境を除けば、殊更、再現可能性を妨げる特殊な要因が存在するというようには考えられない。というのも、数理心理学に関する先行研究にて述べたように、正答率、反応時間、そして主観的変数の関係については、さまざまな分野で既に報告されている既知のことだからである。ただし、いうまでもなく、今後の研究において、母集団がもつ性質の差が、本論の実証結果にどのような影響を及ぼすかは、関心の一部であることには変わらない。

次に、本論が使用した項目に関する一般化可能性についてである。本論が使用した項目は、8種類の言語構造に渡る48項目の刺激であったが、これらの項目は、あらゆる外国語の文法的側面を網羅しているわけではない。たとえば、意味論的、音韻的、語用論的側面についても、このような項目の性質によって、本論が示したような結果が動的に

変化することは、十分にもっともらしい。ただし、対象とする言語間において、本論が主張した変数の関係が再現できなくなるというような想定は、不適切であると考えている。本論の対象は英語であったが、本論の結果について、英語という個別言語が独自性をもつとは考えられない。

最後に、課題に関する一般化可能性である。一元観は、あくまでも判断行動をその視野に置いた枠組みである。文法知識を測定するとされる課題として、近年注目を浴びている自己ペース読み課題 (e.g., Jiang, 2007; Suzuki, 2017) などから発生する観測などは、まったく説明の対象外である。自己ペース読み課題などといった測定方法は、誤り検知や特定の読解時間の遅延効果といった、異なるパラダイムによるものであり、これらの言語行動とそれに付随する現象が、一元観が示すような構造を示すかどうかは不明である。ただし、第三章にて述べたように、ある種の知識表象を調査するためのもっとも基本的な方法は、判断行動から得られるデータの数理的解析であり、GJT を測定具として据えることは、外国語における文法知識研究において、もっとも一般的な態度であると考えている。同様に、R. Ellis が使用した口頭課題なども、説明の対象外である。すなわち、繰り返しになるが、一元観はあくまでも判断行動²を前提としたモデルである。

(c) 技術的問題点と潜在的交絡要因

さらに、本論が実施した実験や解析方法には、さまざまな問題点が残っている。

第一に、主観的変数の測定についてである。本論は、主観的変数について、二値反応として、規則反応および直観反応という水準を採用した。しかしながら、この水準の選択はやや恣意的であり、主観的変数としてこれまでに使用されているものは、ほかにもあらゆる種類の水準がある。今後、確信度や本論とは異なる水準、そして人工言語習得研究に見られるような多値反応による主観的変数の測定方法を使用した追行実験が行われるべきである。

第二に、個人差についてである。本論が行った拡散過程モデルによる解析は、実験参加者毎のデータを無作為にプール化した後に、それを大被験者としてみなし、それら大被験者毎の母数の推定値から中央値を得ることによって、集団全体の母数を推定した。これは、あくまでも統計的な適切性の確保を目的とした処遇であり、本来であれば望ましいものではない。たとえば、各実験参加者毎の母数を個別に推定することが、より望ましいことは明白である³。ただし、48 項目からなる本実験の実験計画から見て、各実験参加者毎に固有の母数を推定することは解析上困難である。これは、拡散過程モデルの

² さらにいえば、ここでの判断行動は、選択肢所与の強制択一のみを示しており、あらゆる判断行動をも含むものでもない。また、自然な言語行動において、文法性に関する判断行動が一般的であるというわけではない。

³ 本論では、大被験者をもちいた上で複数人に固有な母数を推定し、その平均的値をもとめることで標本全体の推定値を得た。ただし、個人差について考慮する場合、一元観がまさに主張するように、個人に固有な母数を推定すべきである。

解析に莫大な試行数を要するからであり、大被験者という処遇は、純粹にこのためである。しかしながら、このことによって推定結果にバイアスが含まれることは否定できない。たとえば、大被験者毎に推定された項目間散布度に関する母数は、実際の項目間散布度のみではなく、大被験者内に含まれる個人間の散布度と明らかに交絡している。しかし、そうはいつでも、これは大局的に見て、拡散過程モデルが当該のデータに対して優れた数理的近似を示すという主張自体の妨げにはならないと考えている。また、研究仮説 2 について、特に反応時間が規則反応率に及ぼす影響における個人間分散を検討すると、少数ではあるが、一定の個人において、係数の符号が逆転する、つまり、反応時間が短ければ、主観的変数が直感反応である確率の上昇を示す個人が存在する可能性も示唆された。本論は、このことを説明する理論的基盤をもたない。今後、上記のような個人差についての研究を進める必要がある。

第三に、文法性による結果の差である。GJT における正答には、本来二種類のパターンがある。つまり、文法文を容認する反応 (hit)、非文法文を却下する反応 (correct rejection) である。これらは、認知主義的にいって、少なくとも同種の認知過程によるものだとは考えられない。今後、このような潜在的要因についての検討も必要である。

最後に、競合しうる数理的モデルについてである。まず、拡散過程モデルは、一元観のコンセプトに沿う既存の数理モデルのひとつであるものの、直接的に一元観の主張を表すものではないことに注意が必要である。たとえば、一元観は、証跡を個人と項目の組み合わせにおいて固有であると見ているが、一般的な拡散過程モデルが、必ずしもそのような含意を含むわけではない (e.g., Ratcliff, 1978; Ratcliff & McKoon, 2004)。

一元観の主張を直接的に表す数理モデルは、3つの観測変数、すなわち正答率、反応時間、主観的変数に対して影響を及ぼすマルチレベルの分布混在型因子分析モデル⁴であろう。ただし、これらの変数は、第一に試行にネストしており、次に項目および個人にネストしている。さらにそれぞれの変数は、連続量と二値といったように、明らかに分布が異なっており、解析が極めて困難である。今後、このような数理モデルの開発も望まれるところであろう。

一方、明示的知識および暗示的知識の二分法に関する従来の因子分析モデル (R. Ellis, 2005) は、第三章で述べたように、検証するまでもないが、正答率、反応時間、主観的変数といった観測が、2つの離散的な潜在クラスによって説明されるという数理モデルは関心の対象になる⁵。これは一種の混合分布モデルか潜在クラスモデルに帰着するであろう。

⁴ たとえば、ここでいうように、連続型変数である潜在変数が、二項分布に従う正答率および規則反応率、そして指数正規合成分布といった分布に従う反応時間に影響をおよぼすモデルも想定できる。つまり、ひとつの連続型変数が、正答率、規則反応率、反応時間分布に影響を及ぼすとするモデルである。

⁵ これは、カテゴリカル変数である潜在変数が、二項分布に従う正答率および規則反応率、そして指数正規合成分布などといった分布に従う反応時間に影響をおよぼすモデルのことである。つまり、ひとつのカテゴリカル変数 (明示的知識および暗示

うが、先の例と同様に、データの複雑な階層性や、解析に用いる分布が混在している点から見て、明らかに解析困難である⁶。

また、本論では詳しく取り上げなかったが、項目反応理論には、正誤反応のみならず、反応時間を取り入れるモデルがいくつか存在する。これらのモデルにおけるモデリング・コンセプトは、一元観がもつ主張と親和性が高い。たとえば、永岡・植野（1991）は、個人に対して固有である速力（ τ ）という母数を想定した項目反応理論を開発している。また、Tuerlinckx and De Boeck（2005）や van der Maas et al.（2011）は、拡散過程モデルを項目反応理論に拡張した数理モデルを開発している。その他にも、正誤反応と反応時間の両方を取り入れた項目反応理論ベースのモデルは数多い。それらの数理モデルの詳細については、van der Linden（2009）に詳しい。さらに、Shigemasu and Ueno（1993）は、確信度を取り入れた項目反応理論のモデルを提案しており、これと似たような試みには、永岡・植野（1992）などがある。

これらの数理モデルは、現在のところ外国語教育研究において適用例がないが、一元観との関係性を検討することと合わせて、今後、応用の可能性を探っていくべきであろう。国内における外国語教育研究は、その歴史を鑑みるならば、いうまでもなく学際志向であるのだから、今後、第二言語習得研究に限らず、心理統計や数理心理学、そして情報工学的なアプローチを応用していくこともひとつの有効な方略であろう。

（d）一元観の抽象性

繰り返しになるが、一元観は、一種の数理的なモデリング・コンセプトであり、それ自体は非常に抽象的なものである。先述の通り、具体的な数理モデルを表すわけではない。そのために、このモデルは、特に以下の面において非常に抽象的なものに留まる。

第一に、証跡が、個人と項目の組み合わせにおいて固有である、という主張における、項目の定義は実に曖昧である。一元観は表象主義を採用しているため、少なくとも理論的には、項目を知識表象と読み替えることができる。よって、紐付けられる知識表象が同一と考えられる複数の言語項目、ないし実際の刺激文について、証跡の値は同一であると考えべきである。ただし、本論では、このような理論的な含意に反して、刺激文における証跡の差を積極的に捨象している。また、表象主義の伝統に則って、知識表象を有限離散、かつ、個々人間で共有しているものと仮定しているが、この態度について、

的知識みが、正答率、規則反応率、反応時間分布に影響を及ぼすとするモデルである。これは上記のモデルと直接的に適合度を比較することができる。

⁶ 潜在クラスモデルは、観測変数をカテゴリカル変数とし、潜在変数も同様にカテゴリカル変数とする潜在変数モデルのひとつである。一方、ここでの混合分布モデルは連続型変数をカテゴリカル変数とし、潜在変数をカテゴリカル変数とするモデルとして捉えることができる。

実証性が高いわけではない。今後、連合主義⁷をも視野に入れ、一元観の理論的含意自体についても更に検討する必要があるだろう。

第二に、前項とも関連するが、主観的変数の水準について、具体的な予測をもたないという点が曖昧なままである。早急に具体的な予測を可能にする理論的な背景の整備が必要であろう。

最後に、証跡の時系列変化、すなわち証跡の遷移と遷移制約についての実証を、本論はまったく実施していない。この点に関して、厳密な経験主義的態度を取ることを推奨しているが、もちろん、今後の実証研究における最優先事項である。

6.3 一元観がもたらす効用と今後の研究

本論には、上記のような限界点があり、反論の余地が数多いものの、一元観は、第三章にて述べたように、さまざまな場面の教育的応用において、明示的知識および暗黙的知識の二分法よりも、より高い効用を示すことが期待される。その効用化の方法に関する大枠については、まさに第三章の通りであるが、ここでは、本モデルの応用研究についての展望を示したい。

もっとも効率的な効用化が期待できる分野は、いうまでもなく、e テスティングである。従来の外国語におけるテストの大半が、なんらかの判断行動の結果にもとづくものであり、判断行動に付随する観測データは、二値反応にしる、反応時間にしる、現在では非常に得やすい種類のデータであるといえる。反応時間に関しては、従来、主流であった紙面でのテストを使用する際には記録し難いものであったが、e テスティングを使用することによって、容易に記録することが可能となっている。それだけでなく、サーバ上において、データベースといった形式を使用することで、それらのデータを効率的に保存・加工することも可能である。同様に、e テスティングの技術のもとでは、判断行動に関する主観的変数、たとえば確信度や主観的困難度評定も、非常に得やすい、ありふれたデータの種類となる。このようなデータは、WBT や CALL 教材の爆発的普及に伴い、外国語教育の実践現場において、ビッグデータとは当然いわないまでも、近年では、相当な量が累積されていると考えられている（草薙, 2017d）。一元観は、そのような豊富な情報を数理的に集約するための、ひとつの認識論的基盤になる。そしてそのような大規模のデータに関する数理的な集約は、さまざまな教育的意思決定に影響を及ぼすであろう。

そのほかにも、第三章で述べたように、集団間における測定不変性と構造不変性の問題の解消、成果の一元化による評価の効率化、そしてパネルデータ⁸といった時系列情報

⁷ コネクションイズムなどとして知られる立場のひとつであり、表象主義と対比される場合もある。表象主義は、少なくとも記号論的で明示的な表象を仮定しない。今後理論面において、このような立場と一元観の関係を議論する必要がある。また、特に ACT* などといった計算論的な枠組みとの関連も検証する必要がある。

⁸ 時系列データとクロスセクションデータの両方をもつものがパネルデータと呼ばれる。つまり、一般的にいわれるところの縦

取得のコストダウン，といった効用が期待できる。これらの応用的側面についても，特にe テスティングの応用を中心にしながら取り組んでいくとよいだろう。

6.4 結語

最後に，結語として，本論全体の流れをまとめる。

本論は，第一に，「外国語における文法知識が二つに分けられる」という言明を中心にして，外国語の文法知識におけるこれまでの研究を概観した。次に，第二言語習得研究における明示的知識および暗示的知識の二分法に焦点を当て，この二分法の歴史的成立経緯とその問題点について，数理的アプローチの観点から批判を加えた。その批判についての概要は，以下の通りである。

- (a) 第二言語習得研究は，成立当初より，学習者と母語話者の運用の差に注目してきた
- (b) 1970年代，中間言語可変性研究より，個人内における運用の差に研究者の関心が移行した
- (c) 中間言語可変性研究の中で，複数の観測に対して，「構造的仮定」が導出された
- (d) これは，集団レベルである学習者と母語話者の運用の差を，個人内に投影したものであった
- (e) しかしながら，構造的仮定は，個人の能力と課題の困難度を混同した
- (f) また，SATとして知られるありふれた現象について考慮されたものではなかった
- (g) 2000年代，R. Ellisらによって，構造的仮定を実証したとする心理測定的研究が発表され，明示的知識および暗示的知識の二分法が広く知られるようになった
- (h) その際，「構造の優先」とでもいうべき，実証性の非常に低い言説が件の二分法に含まれた
- (i) それは主に，認知心理学における明示的知識および暗示的知識の特性を，未検証のまま統合したものであった
- (j) 明示的知識および暗示的知識が二種類に分けられることの証拠として示された心理測定的研究自体にも，さまざまな問題点があった

これらの状況にもとづいて，本論は，明示的知識および暗示的知識の二分法に対抗する認識論的基盤として，一元観を提案した。一元観は，あくまでも，判断行動に関する抽象的なモデリング・コンセプトであり，具体的には以下のような含意をもつ。

- (a) 一元観は，表象主義にもとづく
- (b) 一元観は，個人と知識表象の組み合わせにおいて固有な証跡があると仮定する
- (c) 証跡は，連続的な潜在変数である
- (d) 証跡は，判断行動における正答率，反応時間，主観的変数の原因である

断データと横断データの組み合わせである。近年は，統計モデリングの高度化によって，パネルデータが重宝されるようになってきている。

- (e) 証跡は、課題の困難度や課題条件とは独立である
- (f) 一元観は、数理心理学や認知心理学における基本的な既知事実と親和性がある

この一元観に関するもっとも基本的な証拠を得るための実験を行ったところ、以下のようなことが明らかになった。

- (a) 外国語における GJT から得られる観測は、拡散過程モデルという数理モデルによって優れた数理的近似が与えられる
- (b) GJT において、正答と短い反応時間は、主観的変数が規則反応率である確率を上昇させる

これらのことから、本論は、明示的知識および暗示的知識の二分法よりも、一元観が、外国語の文法知識のあり方に関して、優れた認識論的基盤になる得るもののひとつだと主張した。ただし、一元観は現状において不完全なものであって、その是非については、今後、多数の研究が必要となるだろう。

まさに結語となるが、本論は全編を通して、当該分野の主流ではない、数理的アプローチの思想的背景と、その研究方法に全面的に則っている。そして、本論が行った批判や主張のほとんどは、まさに数理的アプローチや、そこに含まれる思想的特徴を反映したものである。前述の通り、数理的アプローチは、帰結主義を積極的に受け入れており、一元観などといった本論での主張は、その帰結によって評価されるべきであると考えている。今後、一元観が、将来的に我々の教育実践において高い効用をもたらす、という一種の帰結が、本論と本論を支えた研究アプローチの評価を最終的に規定するものになると著者は考えている。

参考文献

欧文文献

- Abutalebi, J., Cappa, S. F., & Perani, D. (2001). The bilingual brain as revealed by functional neuroimaging. *Bilingualism: Language and Cognition*, 4, 179–190.
- Anderson, J. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103, 411–423.
- Bailey, N., Madden, C., & Krashen, S. D. (1974). Is there a “natural sequence” in adult second language learning?. *Language Learning*, 24, 235–243.
- Banks, W. P. (1970). Signal detection theory and human memory. *Psychological Bulletin*, 74, 81–99.
- Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language*, 68, 706–755.
- Bowels, M. (2011). Measuring implicit and explicit linguistic knowledge: What can heritage language learners contribute? *Studies in Second Language Acquisition*, 33, 247–271.
- Carr, T. H., & Levy, B. A. E. (1990). *Reading and its development: Component skills approaches*. Academic Press.
- Clahsen, H., & Felser, C. (2006a). Grammatical processing in language learners. *Applied Psycholinguistics*, 27, 3–42.
- Clahsen, H., & Felser, C. (2006b). How native-like is non-native language processing?. *Trends in Cognitive Sciences*, 10, 564–570.
- Cleeremans, A., & Jiménez, L. (2002). Implicit learning and consciousness: A graded, dynamic perspective. In R. M. French & A. Cleeremans (Eds.), *Implicit learning and consciousness: An empirical, computational and philosophical consensus in the making*. Hove, UK: Psychology Press.
- Dewhurst, S. A., & Conway, M. A. (1994). Pictures, images, and recollective experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1088–1098.
- Dickerson, L. (1975). The learner's interlanguage as a system of variable rules. *TESOL Quarterly*, 9, 401–407.
- Dienes, Z. (2004). Assumptions of subjective measures of unconscious mental states: Higher order thoughts and bias. *Journal of Consciousness Studies*, 11, 25–45.

- Dienes, Z. (2008) Subjective measures of unconscious knowledge. *Progress in Brain Research*, 168, 49–64.
- Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: Distinguishing structural knowledge and judgment knowledge. *Psychological Research*, 69, 338–351.
- Dulay, H. C., & Burt, M. K. (1973). Should we teach children syntax?. *Language Learning*, 23, 245–258.
- Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, 5, 155–174.
- Ellis, N. C. (1998). Emergentism, connectionism and language learning. *Language Learning*, 48, 631–664.
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit knowledge. *Studies in Second Language Acquisition*, 27, 305–352.
- Ellis, R. (1985). Sources of variability in interlanguage. *Applied Linguistics*, 6, 118–131
- Ellis, R. (2004). The definition and measurement of 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. (2007). Confirming the operational definitions of explicit and implicit knowledge in Ellis (2005). *Studies in Second Language Acquisition*, 29, 119–126.
- Ellis, R., Loewen, S., Elder, C., Erlam, R., & Philip, J. (2009). *Implicit and explicit knowledge in second language learning, testing, and teaching*. Multilingual Matters Ltd.
- Erlam, R. (2009). The elicited oral imitation test as a measure of implicit knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam & J. Philip. (2009). *Implicit and Explicit Knowledge in Second Language Learning, Testing, and Teaching* (pp. 32–65). Multilingual Matters Ltd.
- Farah, M. J. (1994). Neuropsychological inference with an interactive brain: A critique of the “locality” assumption. *Behavioral and Brain Sciences*, 17, 43–61.
- Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28, 3–71.
- Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. *Advances in Psychology*, 134, 217–236.

- Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*, 21–38.
- Gutiérrez, X. (2012). Implicit knowledge, explicit knowledge, and achievement in second language (L2) Spanish. *The Canadian Journal of Applied Linguistics*, *15*, 20–41.
- Gutiérrez, X. (2013). The construct validity of grammaticality judgment tests as measures of implicit and explicit knowledge. *Studies in Second Language Acquisition*, *35*, 423–449.
- Han, Y. and Ellis, R. (1998). Implicit knowledge, explicit knowledge, and general language proficiency. *Language Teaching Research*, *2*, 1–23.
- Hawkins, R. (2001). The theoretical significance of Universal Grammar in second language acquisition. *Second Language Research*, *17*, 345–367.
- Heitz, R. P. (2014). The Speed-Accuracy Tradeoff: History, Physiology, Methodology, and Behavior. *Frontiers in Neuroscience*, *8*, 150.
- Hulstijn, J. (2005). Theoretical and empirical issues in the study of explicit and implicit second language learning: Introduction. *Studies in Second Language Acquisition*, *27*, 129–140.
- Isemonger, M., I. (2007). Operational definitions of explicit and implicit knowledge. *Studies in Second Language Acquisition*, *29*, 101–118.
- Jeon, E. H., & Yamashita, J. (2014). L2 Reading comprehension and its correlates: A meta-analysis. *Language Learning*, *64*, 160–212.
- Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, *25*, 603–634.
- Jiang, N. (2007). Selective integration L2 knowledge in adult second language learning. *Language Learning*, *57*, 1–31.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–99.
- Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology*, *3*, 1–24.
- Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, *28*, 107–128.
- Kelley, H. H., & Michela, J. L. (1980). Attribution theory and research. *Annual Review of Psychology*, *31*, 457–501.
- Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.
- Krashen, S. (1994). The input hypothesis and its rivals. In N. Ellis (Ed.), *Implicit and explicit*

- learning of languages* (pp. 45–77). San Diego, CA: Academic Press.
- Kruschke, J. K. (2010). *Doing Bayesian data analysis: A tutorial with R and BUGS*. New York, NY: Academic Press.
- Kusanagi, K. (2014). Speeded effect on accuracy, sensitivity, response bias and reaction time of L2 learners' grammaticality judgments: Using signal detection theory. *JABAET Journal*, 18, 37–54.
- Kusanagi, K., & Fukuta, J. (2015). Japanese EFL learners' implicit knowledge of prenominal adjective orders: A priming study. *International Journal of Curriculum Development and Practice*, 17, 1–13.
- 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. *Annual Review of English Language Education in Japan*, 24, 205–220.
- Lardiere, D. (2000). Mapping features to forms in second language acquisition. In Archibald, J., (Ed), *Second language acquisition and linguistic theory* (pp. 102–29). Oxford: Blackwell.
- Larsen-Freeman, D. E. (1975). The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly*, 9, 409–419.
- Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.
- Lilienfeld, S. O., Sauvigne, K. C., Lynn, S. J., Cautin, R. L., Lutzman, R. D., & Waldman, I. D. (2015). Fifty psychological and psychiatric terms to avoid: A list of inaccurate, misleading, misused, ambiguous, and logically confused words and phrases. *Frontiers in Psychology*, 6, 1100.
- 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* (pp. 66-94). Multilingual Matters Ltd.
- Luce, D. (1986). *Response times: Their role in inferring elementary mental organization*. New York, NY: Oxford University Press.
- Martin, A. D., Quinn, K. M., & Park, J. (2011). MCMCpack: Markov chain monte carlo in R. *Journal of Statistical Software*, 42, 1–21.
- McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information-processing perspective. *Language Learning*, 33, 135–158.

- Mealor, A. D., & Dienes, Z. (2012). Conscious and unconscious thought in artificial grammar learning. *Consciousness & Cognition*, *21*, 865–874.
- Mealor, A. D., & Dienes, Z. (2013). The speed of metacognition: Taking time to get to know one's structural knowledge. *Consciousness & Cognition*, *22*, 123–136.
- Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first- and second-language learners. *Psychological Bulletin*, *140*, 409–433.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*, 741–749.
- Murdock, B. B., & Dufty, P. O. (1972). Strength theory and recognition memory. *Journal of Experimental Psychology*, *94*, 284–290.
- Norman, D. A., & Wickelgren, W. A. (1969). Strength theory of decision rules and latency in retrieval from short-term memory. *Journal of Mathematical Psychology*, *6*, 192–208.
- Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, *15*, 202–206.
- Perner, J., and Dienes, Z. (1999) Higher order thinking. *Behavioural and Brain Sciences*, *22*, 164–165.
- Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychological Review*, *117*, 864–901.
- Prévost, P. and White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, *16*, 103–33.
- Pustejovsky, J. (1991). The generative lexicon. *Computational Linguistics*, *17*, 409–441.
- R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922.
- Ratcliff, R., & Rouder, J.N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*, 347–356.
- Ratcliff, R., & Rouder, J.N. (2000). A diffusion model account of masking in two-choice letter identification. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 127–140.

- Ratcliff, R. & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*, 333–367.
- Ratcliff, R., & Starns, J. J. (2009). Modeling confidence and response time in recognition memory. *Psychological Review*, *116*, 59–83.
- Ratcliff, R., & Starns, J. J. (2013). Modeling confidence judgments, response times, and multiple choices in decision making: Recognition memory and motion discrimination. *Psychological Review*, *120*, 697–719.
- Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging*, *19*, 278–289.
- Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning*, *63*, 595–626.
- Rebuschat, P., Hamrick, P., Riestenberg, K., Sachs, R., & Ziegler, N. (2015). Triangulating measures of awareness. *Studies in Second Language Acquisition*, *37*, 299–334.
- Revelle, W. (2016) *psych: Procedures for Personality and Psychological Research*, Northwestern University, Evanston, Illinois, USA,
- Rosenthal, D.M. (2000), Metacognition and higher-order thoughts (extended response to commentaries), *Consciousness & Cognition*, *9*, 231–42.
- Schneider, W. (1987). Connectionism: Is it a paradigm shift for psychology?. *Behavior Research Methods, Instruments, & Computers*, *19*, 73–83.
- Scott, R. B., & Dienes, Z. (2008). The conscious, the unconscious, and familiarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *34*, 1264–1288.
- Segalowitz, N. S., & Segalowitz, S. J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics*, *14*, 369–385.
- Shigemasu, K., & Ueno, M. (19993). A new item response model with parameters reflecting state of knowledge, *Behaviormetrika*, *20*, 161–169.
- Shiotsu, T., & Weir, C. J. (2007). The relative significance of syntactic knowledge and vocabulary breadth in the prediction of reading comprehension test performance. *Language Testing*, *24*, 99–128.
- Singmann, H., Brown, S., Gretton, M., & Heathcote, A. (2017). *rtdists: Response Time Distributions*. R package version 0.7-3.
<https://CRAN.R-project.org/package=rtdists>
- Sorace, A. (1985). Metalinguistic knowledge and language use in acquisition-poor

- environments. *Applied Linguistics*, 6, 239–254.
- Spada, N., Shiu, J. L. J., & Tomita, Y. (2015). Validating an elicited imitation task as a measure of implicit knowledge: Comparisons with other validation studies. *Language Learning*, 65, 723–751.
- Spielberger, C. D. (1966). The effects of anxiety on complex learning and academic achievement. *Anxiety and Behavior*, 361–398.
- Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of Learning and Memory*, 82, 171–177.
- Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences*, 93, 13515–13522.
- Suzuki, Y. (2017). Validity of new measures of implicit knowledge: Distinguishing implicit knowledge from automatized explicit knowledge. *Applied Psycholinguistics*, 38, 1229–1261.
- Suzuki, Y., & DeKeyser, R. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning*, 65, 860–895.
- Tamura, Y., & Harada, Y., Kato, D., Hara, K., & Kusanagi, K. (2016). Unconscious but slowly activated grammatical knowledge of Japanese EFL learners: A case of tough movement. *Annual Review of English Language Education in Japan*, 27, 169–184.
- Tamura, Y. & Kusanagi, K. (2015a). Asymmetrical representation in Japanese EFL learners' implicit and explicit knowledge about the countability of normal/material nouns. *Annual Review of English Language Education in Japan*, 26, 253–268.
- Tamura, Y., & Kusanagi, K. (2015b). Measuring Japanese learners' explicit and implicit knowledge of constraints on verb semantics: A case of assertive predicates in English as a foreign language. *International Journal of Curriculum Development and Practice*, 17, 25–37.
- Tarone, E. (1979). Interlanguage as chameleon. *Language Learning*, 29, 181–191.
- Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics*, 4, 142–163.
- Tarone, E. (1988). *Variation in interlanguage*. London: Edward Arnold.
- Tuerlinckx, F., & De Boeck, P. (2005). Two interpretations of the discrimination parameter. *Psychometrika*, 70, 629–650.
- Tulving, E. (1985). How many memory systems are there?. *American Psychologist*, 40, 385–398.

- Vafae, P., Suzuki, Y., & Kachisnke, I. (2017). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition, 39*, 59–95.
- van Der Linden, W. J. (2009). Conceptual issues in response-time modeling. *Journal of Educational Measurement, 46*, 247–272.
- van der Maas, H. L., Molenaar, D., Maris, G., Kievit, R. A., & Borsboom, D. (2011). Cognitive psychology meets psychometric theory: on the relation between process models for decision making and latent variable models for individual differences. *Psychological Review, 118*, 339–356.
- van Gelderen, A., Schoonen, R., De Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2004). Linguistic Knowledge, Processing Speed, and Metacognitive Knowledge in First-and Second-Language Reading Comprehension: A Componential Analysis. *Journal of Educational Psychology, 96*, 19–30.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4-70.
- Volkman, J. (1934). The relation of the time of judgment to the certainty of judgment. *Psychological Bulletin, 31*, 672–673.
- Voskuilen, C., & Ratcliff, R. (2016). Modeling confidence and response time in associative recognition. *Journal of Memory and Language, 86*, 60–96.
- Wagenmakers, E. J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology, 21*, 641–671.
- White, L. (2003). *Second language acquisition and universal grammar*. Cambridge University Press.
- Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41*, 67–85.
- Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review, 114*, 152–176.
- Yamashita, J., & Shiotsu, T. (2017). Comprehension and knowledge components that predict L2 reading: A latent-trait approach. *Applied Linguistics, 38*, 43–67.
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*, 441–517.

Ziori, E., & Dienes, Z. (2012). The Time Course of Implicit and Explicit Concept Learning. *Consciousness & Cognition, 21*, 204–216.

和文文献

- 草薙邦広 (2013). 「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—」『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』 46–67.
- 草薙邦広 (2015). 「明示的および暗示的知識とはなにか, どうやって測るか, なにに役立つか」外国語教育メディア学会中部支部主催新人奨励賞記念講演会. 名古屋大学.
- 草薙邦広 (2017a). 「外国語の読解時における相の強制現象—ベイズ統計によるアプローチ—」『中部地区英語教育学会紀要』 46, 33–38.
- 草薙邦広 (2017b). 「確率分布から見る外国語教育研究データ」『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報告論集』10, 1–40
- 草薙邦広 (2017c). 「オンライン学習履歴データの統計的取り扱い」『広島外国語教育研究』 20, 231–244.
- 草薙邦広・川口勇作 (2015). 「文法性判断の確信度と明示的および暗示的知識」『中部地区英語教育学会紀要』 44, 65–72.
- 草薙邦広・後藤亜希 (2016). 「外国語教育研究と信号検出理論」『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報告論集』 8, 20–36.
- 久保拓弥 (2012). 『データ解析のための統計モデリング入門—一般化線形モデル・階層ベイズモデル・MCMC(確率と情報の科学)』 東京:岩波書店.
- 豊田秀樹 (2015). 『基礎からのベイズ統計学:ハミルトニアン・モンテカルロ法による実践的入門』 東京:朝倉書店.
- 豊田秀樹 (2017). 『実践ベイズモデリング—解析技法と認知モデル—』 東京:朝倉書店.
- 永岡慶三・植野真臣 (1991). 「回答所要時間における項目応答理論」『行動計量学』 18, 1–8.
- 永岡慶三・植野真臣 (1992). 「自信-正誤反応における項目応答理論」『電子情報通信学会論文誌(A)』 J75-A, 2, 407–413.
- 松浦健太郎 (2016). 『Stan と R でベイズ統計モデリング』 東京:共立出版.

謝辞

数多くの先生方、同僚、友人、そして家族の励ましがなければ、本論を完成させることはできなかった。その中でもとりわけ、2010年より入学した名古屋大学において、主指導教官として長年に亘って指導を頂いた山下淳子先生には、言葉に出来ぬほどの感謝とともに、お会いしたまさにその瞬間から、今も変わらない尊敬の念を常日頃感じ続けている。一研究者としてのみならず、ひとり人間として、さまざまな側面において、私のロールモデルであり続けていただいた。また、博士課程在学中や満期退学後にも、度々体調を崩したり、研究上の悩みを抱えていたときもあった。そのようなときも、山下先生は継続的に、私に対して声をかけて下さった。言葉に出来ぬほどの感謝の念、というものを初めて感じさせてくださったのは山下先生だった。

同様に、副指導教官として修士論文、そして博士論文の指導を頂いた杉浦正利先生にも、同様の感情をもっている。杉浦先生がご指導くださったことは、私の研究の随所に色濃く残っており、今の自分の研究スタイルは、杉浦先生の影響なしでは語れないものである。杉浦先生への感謝の念は、私が研究に取り掛かるすべての瞬間、そこにある。

博士後期課程在学中にご指導を頂いた稲垣俊史先生も、私の研究に関して多大な時間を割いて頂いた。また、三輪晃司先生には本論の審査に加わって頂き、貴重なコメントを頂戴した。お二人の先生にも心から感謝を述べたい。

そのほかにも、勤務先の上司である岩崎克己先生と達川奎三先生には、本論文の執筆に関して、絶え間ない励ましの言葉を頂いた。勤務先を同じくし、そして名古屋大学の先輩でもある森田光宏先生と阪上辰也先生にも、先輩らしい、実に温かい励ましを頂いた。鬼田崇作先生と長い間交わした議論は、本論の質を高めてくれたことと思う。諸先生方の励ましにお礼を申し上げたい。

名古屋大学の木下徹先生と尾関修治先生は、この研究に関して直接指導を頂いたわけではなくとも、いつも私のことを気遣って下さっていた。お二人の励ましも、本論の完成にはなくてはならない存在だった。また、名前を挙げきれないほど数多くの先輩、同輩、後輩が私の研究に影響を及ぼしている。彼らと同時代に勉強できたことは、自分の生涯の宝物である。

最後になるが、妻の優佳は同年代の研究者として、そして家族として、全局的に私をサポートしてくれた。彼女の支えにも感謝したい。

付録

付録 A 本実験における刺激

項目番号	提示文	文法性	種類
1	The man played soccer very well when he was young.	文法文	副詞
2	The woman caught a cold because of the bad weather.	文法文	冠詞
3	Bill wanted to know where Jack had been.	文法文	埋め込み
4	If he came to the city, he would enjoy the local food.	文法文	仮定法
5	If she could speak English, she could make many friends.	文法文	仮定法
6	The boys must not leave this building now.	文法文	助動詞
7	The old woman can not cook the traditional foods.	文法文	助動詞
8	Emily is going to meet her son's teacher next week.	文法文	所有格
9	The local leader says that he was an actor.	文法文	三単現
10	The passenger always takes the same seat on the train.	文法文	三単現
11	The man says he wants to get married next year.	文法文	補語
12	Her parents decided to buy a new house.	文法文	補語
13	The old man always likes watching television.	文法文	副詞
14	She always speaks English quickly in class.	文法文	副詞
15	The young people take pride in their local football team.	文法文	冠詞
16	The boss was in the bath at the moment of the decision.	文法文	冠詞
17	Sam wanted to know what they had told John.	文法文	埋め込み
18	The girl didn't know what she should say at the time.	文法文	埋め込み
19	If he had a ticket, he could enter the hall.	文法文	仮定法
20	The girls must not show their homework to each other.	文法文	助動詞
21	The boy is living in his rich uncle's house this month.	文法文	所有格
22	Nancy spoke to the professor's secretary yesterday.	文法文	所有格
23	The old lady cooks rice for dinner every day.	文法文	三単現
24	His brother wants to buy a computer this weekend.	文法文	補語

25	She writes very well English but she can not speak.	非文法文	副詞
26	They will leave soon this room to take lunch.	非文法文	副詞
27	The citizens can make a use of the bus system.	非文法文	冠詞
28	The policeman took look at the map before going out.	非文法文	冠詞
29	Tom wanted to know whether was she going.	非文法文	埋め込み
30	She has no idea about what should she do next.	非文法文	埋め込み
31	If she had another bag, she can pack all the books.	非文法文	仮定法
32	The students can to read the difficult books well.	非文法文	助動詞
33	Joseph is going to meet the local leader son in two days.	非文法文	所有格
34	She visited her father office at ten yesterday.	非文法文	所有格
35	The big boy eat sushi as much as possible.	非文法文	三単現
36	The student wants finding a new job next month.	非文法文	補語
37	They enjoyed very much wine last night.	非文法文	副詞
38	The man recognized her at the glance when she entered the room.	非文法文	冠詞
39	She wanted to know why had he studied English.	非文法文	埋め込み
40	If he were richer, she will marry the man.	非文法文	仮定法
41	If the man had a chance, he will be the president.	非文法文	仮定法
42	The worker can to speak English very well.	非文法文	助動詞
43	The leaders must to attend the meeting tomorrow.	非文法文	助動詞
44	He is going to his sister place tomorrow morning.	非文法文	所有格
45	Mary lives in the countryside but work in the city center.	非文法文	三単現
46	The woman sometimes play the piano in the room.	非文法文	三単現
47	The young man decided buying a car next week.	非文法文	補語
48	The group decided going to the zoo next month.	非文法文	補語

付録 B ベイズ統計の概説

ここでは、本論が利用したベイズ統計に関する概説を行なう（Kruschke, 2010; Lee & Wagenmakers, 2014; 久保, 2012; 豊田, 2015, 2017; 松浦, 2016 など参照のこと）。

ベイズ統計は、**ベイズの定理**を統計的推定に応用した統計的方法論のひとつであり、そのベイズの定理とは、

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} \quad (1)$$

である。これは、事象 A と事象 B の合同確率 $P(A \wedge B)$ が、事象 B という条件下における事象 A の条件付き確率 $P(A|B)$ と、事象 B の周辺確率 $P(B)$ の積であること、つまり、

$$P(A \wedge B) = P(A|B)P(B) \quad (2)$$

から導出される。ベイズの定理自体は、このように条件付き確率、周辺確率、そして合同確率の関係を表しているにすぎないが、これを統計的推定に応用することができる。

まず、(1) 式において、事象 A をデータ (D) と考える。次に事象 B を仮説 (H) または仮説を表す母数と考える。すると、

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)} \quad (3)$$

となる。ここで、 $P(H|D)$ は、データの下での仮説が正しい確率を表す。これをベイズ統計では、事後確率といい、特に分布の形として表すときは、**事後分布**と呼ぶ。事後分布をもとめることが分析の目標である。次に、 $P(D|H)$ は、ある仮説の下において、データが得られる確率を表す。これを**尤度**と呼ぶ。 $P(H)$ は、仮説が正しい確率を表し、これを事前確率および**事前分布**と呼ぶ。事前分布はしばしば**主観的確率**とも呼ばれる。 $P(D)$ は、データが得られる確率であるが、しばしば応用に関する限りは無視してよい。

事後分布は、得られたデータという条件下における仮説の正しさを示すのだから、まさに統計的推定における研究者の最終的関心である。事前分布は、データを得る前の仮説の正しさを示すという点において、主観的である。尤度は、ある母数の下で、観測がどの程度あらわれるかを計算するものである。

(3) から、**事後分布** \propto **尤度** \times **事前分布** と考えることができる。すなわち、推定する事後分布は、データから計算される確率と研究者がもつ主観に比例 (\propto) する。さらに言い

換えるならば、事後分布とは、研究者がある仮説に対してもつ信念と、得られたデータとの整合性を掛け合わせたものである。具体的にいて、事後分布は、ある現象に近似する数理モデルの母数について検討される。たとえば、任意のデータが正規分布に従うとして、その母平均 (μ) を推定しようとする。ここで、研究者は、まず事前分布を設定する。この事前分布は、母平均がどれくらいの値を取るかを示したものである。図 1 に、母平均に関する事前分布の例を示す。ここでの事前分布ならば、母平均は 50 が一番もってもらしく、0 や 100 といった値を取るとはあまりないだろう、といった主観を示している。

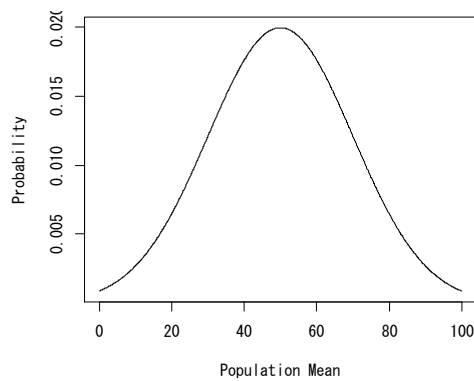


図 1. 事前分布の例

実際のデータからは、尤度がもとめられる。尤度は、たとえば母平均が特定の値である場合に、実測のデータが得られる計算上の確率である。母平均の値を項に取った関数として尤度を表すものを尤度関数という。図 2 に、あるデータにおける（対数）尤度関数の例を示す。

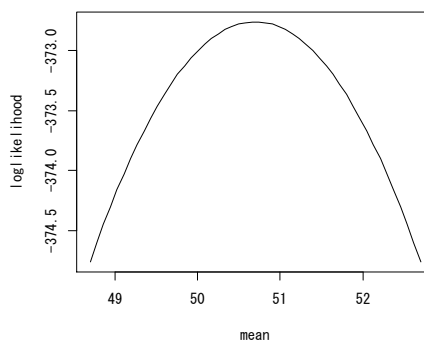


図 2. 尤度関数の例

ここで、先述の通り、事後分布は、事後分布と尤度に比例するため、直感的には、ベイズ推定における母平均は、図1と図2の組み合わせになる。しかしながら、かならずしもその事後分布を表す式が解析的にもとまるとは限らないため、しばしば、**マルコフ連鎖モンテカルロ法 (MCMC)** を使用し、事後分布に比例する分布からシミュレーション計算によってサンプルを得る。これを **MCMC サンプル** という。MCMC サンプルを莫大な数得て、そのサンプルの分布を検討することは、事後分布を検討することにほぼ等しい。たとえば、図3は、指数正規合成分布におけるそれぞれの母数についての事後分布から得られたサンプルと MCMC の様子を示す図であるが、この図より、母数 μ において、 μ が 3.83 周辺の値を取ることは、3.7 や 4.0 といった値を取るよりもかなりもっもらしい、といった柔軟な意思決定をすることができる。また、これに対応するトレース図から見ても、この推定は安定していることが視覚的に確認できる。

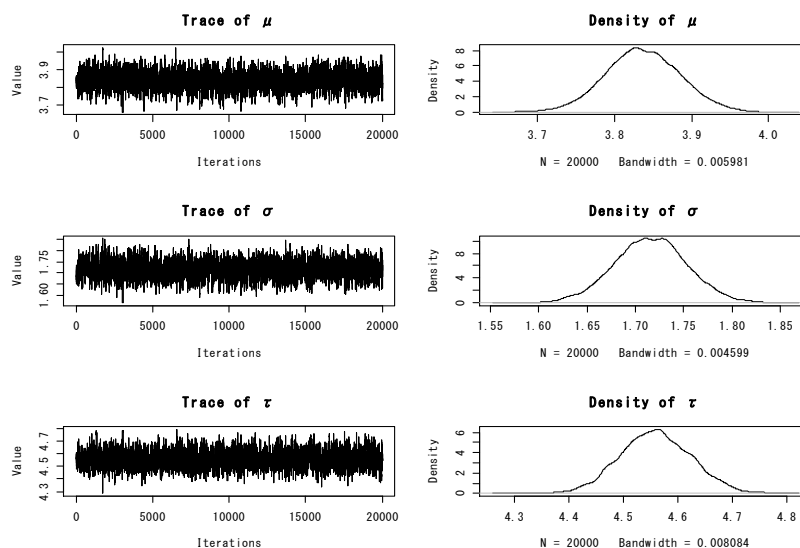


図3. 事後分布とトレース図の例

このように、ベイズ統計は、ある数理モデルにおける母数の事後分布を検討することを主たる目的としているが、しばしば、母数の事前分布の選択やその母数の設定に主観が残るため、それが批判の対象となることもある。そのような場合、事後分布にどのような結果も及ぼさないような事前分布を選択することがある。これを**無情報事前分布**などと呼ぶ。

本論においても、ほとんどの母数に対して無情報事前分布を選択し、できるだけ事前分布の影響を避けるように分析を行った。

付録 C 拡散過程モデルによる本実験のデータのシミュレーション

拡散過程モデルによって、本論文の実験データにおける正答率、正答の反応時間分布、誤答の反応時間分布について数値シミュレーションすることができる。ここでは、統計解析環境 R と rtdists パッケージを使用したコードを示す。なお、シミュレーション試行数を 10,000 としている。

```
#使用するパッケージ
library(rtdists)
library(MCMCpack)
library(retimes)

#シミュレーション
set.seed(0)
sim<-rdiffusion(n=10000,#シミュレーション試行数
  a=5.31,#境界母数
  v=0.34,#漂流母数
  t0=0.75,#非決定時間
  sv=0.33,#漂流母数のばらつき
  st0=4.68)#非決定時間のばらつき

#正答の反応時間におけるヒストグラム
hist(sim[sim[,2]=="upper",1],
  breaks=50,
  main="",
  xlab="Reation Time",
  col="lightgray")

#誤答の反応時間におけるヒストグラム
hist(sim[sim[,2]=="lower",1],
  breaks=50,
  main="",
  xlab="Reation Time",
  col="lightgray")
```

```

#反応時間分布に対して指数正規合成分布をフィット
#指数正規合成分布の尤度関数を作成

llf<-function(beta,x){
  sum(log(dexgauss(x,beta[1],beta[2],beta[3])))
}

#MCMC
model<-MCMCmetrop1R(llf,
  theta.init=c(3,2,4), #初期値
  x=sim[,1],
  burnin=500,
  mcmc=20000)

summary(model)

#シミュレーションデータから拡散過程モデルの母数をリカバリー
#拡散過程モデルの尤度関数を作成
ll_diffusion <- function(pars, rt, response) {
  densities <- tryCatch(
    ddiffusion(rt, response=response,
      a=pars[1], v=pars[2], t0=pars[3],
      sz=pars[4],
      st0=pars[5], sv=pars[6]),
    error = function(e) 0)
  if (any(densities == 0)) return(1e6)
  return(-sum(log(densities)))
}

#初期値を設定（ここでは真値）
start <- c(5.31,0.34,0.75,0,4.68,0.33)
names(start) <- c("a", "v", "t0", "sz", "st0", "sv")

```

```
#最尤推定（計算時間が莫大になるため、ここではデータの一部のみを使用）
fit <- nlminb(start,
  ll_diffusion,
  lower = 0,
  rt=sim[1:500,1],
  response=sim[1:500,2])

#母数がおおよそリカバリーできることを確認できる
round(fit$par,3)
```

付録 D 研究仮説 2 に関する階層ベイズモデルのコード

研究仮説 2 の解析に関して、MCMCpack パッケージにおける MCMChlogit 関数を使用した。具体的なコードは以下のとおりである。ここでは、データフレーム型の変数 `dat` に、ID (被験者番号)、res (正答/誤答)、subj (規則反応/直感反応)、rt (反応時間) が、ロング型の要領で格納されているとする。

```
#パッケージ
library(MCMCpack)

#コード
fit<-MCMChlogit(subj~res+log(rt),
  group="ID",
  random=~res+log(rt),
  data=dat,
  r=3,
  R=diag(0.01,3),
  chain=1,
  mcmc=2000,
  burnin=5000,
  thin=10
)

#結果の要約
summary(fit$mcmc)
```