

## 外国語教育研究と信号検出理論

草薙 邦広

名古屋大学大学院生／日本学術振興会

後藤 亜希

名古屋大学大学院生

---

### 概要

本稿の目的は、信号検出理論 (signal detection theory, SDT) の数理的基盤について紹介し、外国語教育研究における信号検出理論の応用可能性を議論することである。外国語教育研究やその関連諸分野では、語彙性判断課題や文法性判断課題といった各種の判断課題の成績をデータとして取りあつかう場合が多いが、古典的テスト理論の要領で、正答率のみをその主たる分析対象としている。しかし、正答率のみを対象とする分析は、刺激の観察者が一定の水準に対して反応バイアスをもつ場合に、結果の解釈が不適切になる可能性がある。一方、信号検出理論は、正答率のみならず、 $d'$  や  $A'$  といった弁別力指標 (sensitivity index),  $c$ ,  $c'$  や  $\beta$  といった反応バイアス, または ROC などといった多様な分析観点をもたらす。よって、信号検出理論を分析に応用することは、外国語教育研究にとって利点となりうる。

**Keywords:** 信号検出理論, 弁別力, 反応バイアス, 判断課題, 研究方法論

---

### 1. 背景

外国語教育研究およびその関連諸分野では、語彙性判断課題や文法性判断課題といった各種の判断課題 (judgment task) ないし弁別課題 (discrimination task) を実施することがある。ここでは、語彙性判断課題を例とする。典型的な語彙性判断課題は、実在語 (e.g., *cat*, *dog*) および擬似語 (または非単語) といった複数の水準をもつ。刺激の観察者 (実験参加者) は、実在語と擬似語をひとつずつ無作為に提示され、その刺激が実在語であるかどうかを判断する。観察者が当該言語の語彙知識, 正書法知識, 音韻的知識, または形態的知識を十分にもつ場合、実在語を実在語であると判断する確率, 実在語以外の刺激を実在語ではないと判断する確率, そしてそれらの和や平均の値は高い値を示すと予想できる。一方、観察者が上記のような知識を十分にもたない場合、実在語を実在語であると判断する確率, そして実在語以外の刺激を実在語ではないと判断する確率, そしてそれらの

和や平均の値は、逆に低い値を示すと予想できる。よって、これらの値における分散を、知識や技能の程度を表すものであると捉えることは、因果推論として自然である。外国語教育研究では、この推論に沿って、実在語と擬似語を弁別するパフォーマンスの程度（弁別力）を観測値として得て、個人の知識や技能について議論する。

ここで、観察者が実在語を実在語であると判断する確率  $H$  (*hit ratio*) は、(1) 式のように条件つき確率で定義できる。

$$H = P(\text{"real word"}|S_1) \quad (1)$$

一方、実在語以外の刺激を実在語ではないと判断する確率  $CR$  (*correct rejection*) は (2) 式のように定義できる。

$$CR = P(\text{"psudo word"}|S_2) \quad (2)$$

ただし、 $S_1$  は実在語を、 $S_2$  は擬似語といった実在語以外の刺激を示す。

一般にもちいられる正答率 (*accuracy, correct response ratio, propotion correct, P(c)*) とは、これら2つの加重平均であり、水準間の試行数が同数のとき、(3) 式のように表すことができる。いうまでもなく、水準間の試行数は同数が望ましい。

$$P(c) = \frac{1}{2}(H + CR) \quad (3)$$

ここでの正答率は、古典的テスト理論における正答率と等しい。しかしながら、正答率は、判断課題における弁別力の数理的表現として常に適切であるとはいえない。たとえば、ある観察者の弁別結果が、 $H = .50$ ,  $CR = .50$  であったとき、 $P(c) = .50$  である。しかし、別の観察者の弁別結果が、 $H = 1.00$ ,  $CR = 0$  であったときも、同じ値の  $P(c) = .50$  になる。このとき後者の例では、すべての施行について実在語と判断していることになり、前者が両方の水準に対して同様に反応している傾向であるのに対して、著しい差異がある。正答率のみによる課題成績の検討は、反応バイアスを評価することができない。

一方、信号検出理論 (*signal detection theory*) は、弁別力のみならず、反応バイアスについても取り扱うものである。信号検出理論を応用する最も単純な利点として、分析観点の多様化をもたらすことがあげられる。次節からは、信号検出理論の基本的な性質について見ていく。

## 2. 等分散正規分布モデルにもとづく信号検出理論の導出

信号検出理論は、もともとは通信工学に起源をもつものであるが、1960年代にGreen and Swets (1966) によって精神物理学または数理心理学に導入されたとされる (Macmillan & Creelman, 2005)。現在においても、数理心理学や認知心理学、特に記憶や意識についての研究などで頻繁に使用される方法論である。信号検出理論の第一の特徴として、上記のような正答率とは、全く異なるアプローチをもって弁別力を表現することがあげられる。また、「判断する」という認知過程自体を数理的にモデル化する点もその顕著な特徴である。ここでは、多々ある信号検出理論のモデルの中から、もっとも基本的な等分散正規分布モデル (Equal-variance Gaussian model) を導出することで、その機能の一端を示したい。

等分散正規分布モデルとは以下のようなものである。まず、提示された刺激は観察者に対して、連続的な一次元的心理量を喚起させると考える。さらに、擬似語は、実在語に対して相対的に弱い心理量を喚起させるものであり、実在語は、擬似語よりも強い心理量を喚起させると仮定する。これらの心理量はそれぞれの水準ごとに正規分布を成し、水準間で分散が等しいと仮定する。

さらに、弱い心理量を喚起させる分布をノイズ分布、強い心理量を喚起させる分布をシグナル・ノイズ分布と呼ぶ。このように定義されるモデルの視覚的表現を図1に示す。このような図は判断空間 (decision space) とも呼ばれている。

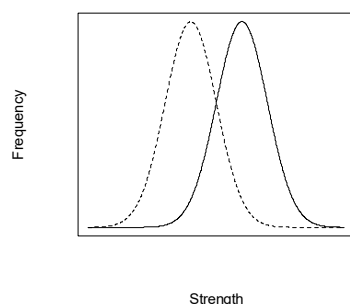


図 1. ノイズの分布 (破線) とシグナル・ノイズの分布 (実線)

ここで、ある刺激に対する心理量上の 1 点、判断基準 (criterion,  $c$ ) について考える。その 1 点を超えるものを、観察者は「実在語である」と判断し、超えないものを「実在語ではない」と判断すると仮定する (図 2)。

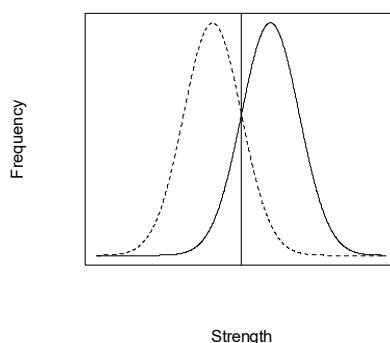


図 2. ノイズの分布とシグナル・ノイズの分布に判断基準を加えたもの

このとき、ノイズ分布の内側であり、かつ判断基準よりも小さい部分の面積は、「擬似語に対して擬似語であると判断する確率」であり、これは  $CR$  である。また、シグナル・ノイズ分布の内側であり、かつ判断基準よりも大きい部分の面積は、「実在語に対して実在語であると判断する確率」であり、これは  $H$  である。正規分布の内側の面積は 1 であるから、 $1-CR$  の値は、「擬似語に対して実在語であると判断する確率」であり、これを  $FA$  (false alarm ratio,  $F$ ) と呼ぶ。同様に、 $1-H$  の値は、「実在語に対して擬似語であると判断する確率」であり、これを  $M$  (miss ratio) と呼ぶ。

(1) 式および (2) 式と同様に、条件つき確率で  $FA$  と  $M$  を表すと、それぞれ (3) 式および (4) 式ようになる。

$$FA = P(\text{"real word"}|S_2) \quad (3)$$

$$M = P(\text{"pseudo word"}|S_1) \quad (4)$$

また、 $H$  に該当する観察数を *hit*,  $M$  に該当する観察数を *miss*,  $FA$  に該当する観察数を *false alarm*,  $CR$  に該当する観察数を *correct rejection* と呼ぶ。このとき、観察者の弁別結果は、以下のようなクロス表で表すことができる (表 1)。もちろん、*hit* を実在語の刺激数で割ったものが  $H$ , *miss* を実在語の刺激数で割ったものが  $M$ , *false alarm* を擬似語の刺激数で割ったものが  $FA$ , *correct rejection* を擬似語の刺激数で割ったものが  $CR$  となる。

表 1.

弁別結果のクロス表

観察者の判断	刺激の特性	
	実在語	擬似語
実在語	<i>Hit</i>	<i>False Alarm</i>
擬似語	<i>Miss</i>	<i>Correct Rejection</i>

さて、 $H$  は、正規分布関数と判断基準で囲まれた面積であるため、 $F(x|S_1)$ を積分することによって、(5) 式のように表現することもできる。

$$H = \int_c^{\infty} f(x|S_1) dx \quad (5)$$

同様に、 $M$  は (6) 式のように表現できる。

$$M = 1 - \int_c^{\infty} f(x|S_1) dx \quad (6)$$

$FA$  は、 $F(x|S_2)$ を積分することによって、(7) 式のように、 $CR$  は (8) 式のように表現できる。

$$FA = \int_{-\infty}^c f(x|S_2) dx \quad (7)$$

$$CR = 1 - \int_{-\infty}^c f(x|S_2) dx \quad (8)$$

この仮定のもとで、平均の差  $d'$  をもとめたい。それぞれの正規分布関数における標準偏差 ( $\sigma$ ) を便宜的に 1 と置き、さらに 2 つの分布の交点を 0 と仮定すると、以下の関係が成立するはずである。研究者によっては、ノイズ分布の平均を 0 と置く場合があるが、以降の計算上大差ない。ノイズ分布  $z(p)$  とは  $p$  における標準正規分布の逆数である。ちなみに、標準正規分布の逆数は、Microsoft Excel では、NORMSINV 関数でもとめられる。

$$-\frac{d'}{2} - c = z(FA) \quad (9)$$

$$\frac{d'}{2} - c = z(H) \quad (10)$$

これを以下のように整理していく。

$$\begin{aligned}c &= -z(FA) - \frac{d'}{2} \\ &= -z(H) + \frac{d'}{2}\end{aligned}\tag{11}$$

であるから,

$$\begin{aligned}-z(FA) - \frac{d'}{2} &= -z(H) + \frac{d'}{2} \\ -\frac{d'}{2} - \frac{d'}{2} &= -z(H) + z(FA) \\ -d' &= -z(H) + z(FA) \\ d' &= z(H) - z(FA)\end{aligned}\tag{12}$$

と計算できる。

よって (12) 式からもとまる  $d'$  は、ノイズ分布とシグナル・ノイズ分布の中心傾向の差の大きさを示す。この指標は、弁別力指標 (sensitivity index) とも呼ばれ、信号検出理論において、判断ないし弁別の成績を示すもっとも基本的な指標である。 $d'$  が 0 のとき、その観察者は全く弁別力がないことを意味し、 $d'$  の値が大きいときには高い弁別力を、逆に値が小さいときには低い弁別力を示す。つまり、判断空間におけるノイズ分布とシグナル・ノイズ分布が離れていればいるほど、また、共有する面積が少なければ少ないほど、弁別力が高いということである。ここで、信号検出理論では、ノイズ分布とシグナル・ノイズ分布、判断基準といったパラミタは、観察者や条件に由来するものと考えることに注意されたい。

また、これまでの導出過程から、判断基準 ( $c$ ) の値をもとめることもできる。まずは、(9) 式を  $c$  について整理すると (11) 式のようになる。(11) 式から、

$$c = -z(FA) - \frac{d'}{2}$$

であり、(12) 式より、

$$d' = z(H) - z(FA)$$

なので、これより、

$$\begin{aligned}c &= -z(FA) - \frac{z(H) - z(FA)}{2} \\ &= -\frac{-2z(FA) - z(H) + z(FA)}{2} \\ &= -\frac{-z(FA) - z(H)}{2} \\ &= -\frac{(z(H) + z(FA))}{2}\end{aligned}\tag{13}$$

となる。

このようにしてもとめられる判断基準  $c$  が 0 のとき、反応バイアスがないことを示す。 $c$  が正の値を取るとき、実在語であると回答する方に、負の値を取るときに擬似語であると回答する方に反応バイアスをもつことを示す。繰り返しになるが、このような反応バイアスは、正答率のみによる分析から検討することはできない。

さらに弁別力指標および反応バイアスは、ROC (receiver operating characteristics または、relative operating characteristics) という視覚的表現からも理解することができる。ROC は、横軸に  $FA$  を、縦軸に  $H$  を取った 2 次元空間である (図 3)。この 2 次元空間において、直線  $y = x$  は、 $d' = 0$ 、つまり弁別力がないことを示す。この直線よりも左上にある曲線は、この直線から見て、それぞれ  $d' = 1$ 、 $d' = 2$ 、 $d' = 3$  の弁別力をあらわす。

また、直線よりも右下の観察者の反応が付置される場合は、負の弁別力を表しうるが、研究者が実質的な興味をもつことは少ない。一方、 $y = -x + 1$  の直線は、 $c = 0$ 、つまり反応バイアスがないことを示す。この直線よりも右上の空間は、実在語を選好する反応バイアスを示し、左下の空間は擬似語を選好する反応バイアスを示す。

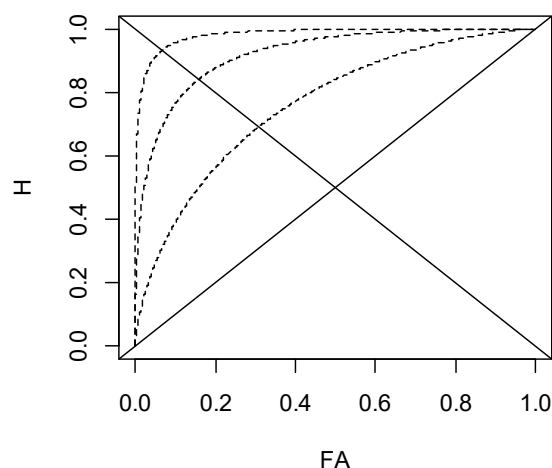


図 3. ROC を表す 2 次元空間

また、 $H$  および  $FA$  を  $z$  変換（標準正規分布の逆数を得る操作）したものを軸として ROC を描けば、 $d'$  は直線で表現できる。

ここまでが等分散正規分布モデルの基礎である。

### 3. 信号検出理論の指標とその周辺

#### 3.1 関連するモデル

信号検出理論に属するモデルは等分散正規分布モデルに限らず、その理論的な枠組みだけでも非常に多岐にわたる。単一高閾値モデル (single high threshold model)、低閾値モデル (low threshold model)、Luce の選択理論 (choice theory) などは弁別力や判断基準に対して異なる指標を使用する。これらは、Macmillan and Creelman (2005) に詳しいため、興味をお持ちの読者は参照してほしい。また、非等分散モデル、多肢選択モデル、ノンパラメトリックなモデル、ROC を駆使したさまざまな計算方法などは、認知心理学においても、古くから広く使われているモデルである。

さらに、信号検出理論は、一般化線形モデル (GLM) のひとつとして扱うこともでき (e.g., DeCarlo, 1998)、変量効果を含む一般化線形混合効果モデル (GLMM) を応用することができる。近年では、ベイズ統計学やマルコフ連鎖モンテカルロ法 (MCMC) による推定を援用した信号検出理論モデルも提案されている (e.g., Lee, 2008)。

信号検出理論に関連する数理心理学的モデルに目を向けると、Ratcliff による拡散モデル (diffusion model) などに代表されるような、逐次抽出過程モデル (sequential sampling model) は、信号検出理論を拡張し、弁別力および判断基準だけではなく反応時間分布の生成も同時に説明するモデルであるといえる (e.g., Ratcliff, 1998, 2002; Ratcliff, Gomez & McKoon, 2004)。

認知心理学において、記憶研究では R-K 手続き (remember-know procedure) を援用



する実験 (e.g., Yonelinas, 2002) が信号検出理論とその拡張的モデルを頻繁に使用する。人工言語学習や明示的および暗示的知識に関する研究では、判断の確信度や、判断に伴う主観的状态 (subjective state) と判断の関係性をモデル化するために、第二種の信号検出理論 (type II signal detection theory) やその発展的モデルを使用している (e.g., Barret, Dienes & Seth, 2013)。もちろん、これらの手法も、近年では一般にベイズ統計にもとづく分析が従来の分析法に取って代わるようになってきており、今後しばらくは、この方向性で進むと考えられる (Dienes, 2016)。

これらのすべてを詳しく紹介することは、本稿の限界を遥かに超えるものであるが、以下には、研究における実用性を重視し、一般的な信号検出理論における指標とその周辺をいくつか紹介する。

### 3.2 反応バイアス指標 $c'$ および $\beta$

反応バイアスには、 $c$  のほかにも相対判断基準位置 (relative criterion location) と呼ばれる  $c'$  や尤度比にもとづく判断基準である  $\beta$  とよばれる指標がある。 $c$  は、 $d'$  が極端な値を取ると過大な値を返す傾向があるため、 $d'$  の値で調整した判断基準が必要な場合がある。 $c'$  は、まさに  $d'$  で調整した  $c$  であり、(14) 式で定義される。

$$c' = \frac{c}{d'} \quad (14)$$

また、 $\beta$  は、ノイズ分布とシグナル・分布の尤度比にもとづくものであり、(15) 式で計算できる。

$$\beta = e^{cd'} \quad (15)$$

ここでの  $e$  は、ネイピア数 (2.71...) である。よって、自然対数の形式をとった  $\ln(\beta)$  は、(16) 式でもとまる。

$$\ln(\beta) = cd' \quad (16)$$

一般的にこれらの指標の方が、 $c$  よりも反応バイアスを表すものとして頻繁にもちいられるようである。

### 3.3 ノンパラメトリックな指標

これまで紹介した指標は、等分散正規分布モデルにもとづき、パラメトリックな指標であるが、信号検出理論の指標の中には、ノンパラメトリックな指標もある。弁別力指標のひとつである  $A'$  は、ROC 空間における一点よりも下の面積を推定することによってもとめられる (Grier, 1971)。Grier による定義は、(17) 式である。

$$A' = \frac{1}{2} + \frac{(H - FA)(1 + H - FA)}{4H(1 - FA)} \quad (17)$$

これは、 $H$  が  $FA$  よりも大きい場合に限り、逆の場合は (18) 式を使う。

$$A' = \frac{1}{2} + \frac{(FA - H)(1 + FA - H)}{4FA(1 - H)} \quad (18)$$

場合によっては、 $d'$  よりも  $A'$  を使うことが推奨される。

### 3.4 選択理論とオッズ比

数理心理学者である Luce の選択公理 (choice axiom) にもとづく選択理論は、等分散正規分布モデルによる信号検出理論と非常に似通った指標を取りあつかう。選択理論における弁別力指標  $\alpha$  は、(19) 式で定義される。

$$\alpha = \left[ \frac{H(1 - FA)}{(1 - H)FA} \right]^{\frac{1}{2}} \quad (19)$$

この指標を対数形式であらわすと、(20) 式になる。

$$\ln(\alpha) = \frac{1}{2} \ln \left( \frac{H}{1 - H} \right) - \frac{1}{2} \ln \left( \frac{FA}{1 - FA} \right) \quad (20)$$

$d'$  に定数をかけると  $\ln(\alpha)$  に漸近する性質が知られており (Macmillan and Creelman, 2005, p. 95-96)、数理的に  $d'$  と  $\ln(\alpha)$  の関係は深い (cf. DeCarlo, 1998)。

### 3.5 $d'$ および $c$ の標本誤差と信頼区間の構成

信号検出理論でもとめられる指標は、もちろん標本値であり標本誤差がある。 $d'$  の標

準誤差は、漸近的な方法でもとめることができる (Macmillan & Creelman, 2005, p. 325)。まず、ある比率  $p$  の、標準正規分布逆数における観測値  $z(p)$  の標本分散は (21) 式で表せられる。

$$\text{var}[z(p)] = \frac{p(1-p)}{N[\phi(p)]^2} \quad (21)$$

ここから、(22) 式のように  $d'$  の標本分散をもとめられる。

$$\text{var}(d') = \frac{H(1-H)}{N_2[\phi(H)]^2} + \frac{FA(1-FA)}{N_1[\phi(FA)]^2} \quad (22)$$

よって、(23) 式は標準誤差を表す。

$$SE(d') = \sqrt{\text{var}(d')} \quad (23)$$

ここから正規分布を使って、95%信頼区間 (CI) を構成すると、(24) 式の通りになる。

$$95\% \text{ CI } (d') = d' \pm 1.96 SE(d') \quad (24)$$

$c$  の標準誤差も、(25) 式でもとまる。

$$SE(c) = \sqrt{0.25 \text{ var}(d')} \quad (25)$$

ただし、一般的にこの方法には問題があるともいわれているので、注意が必要であり (Macmillan & Creelman, 2005)、より正確な計算法もいくつか提案されている。

#### 4. 分析の例

ここからは、具体的な数値例を元に、信号検出理論指標の計算をおこなう。ある観察者に対して語彙性判断課題をおこなったとし、以下のような反応が得られたとする。

表 2.

ある観察者における反応の例

観察者の判断	刺激の特性	
	実在語 ( $k = 50$ )	擬似語 ( $k = 50$ )
実在語	29	18
擬似語	21	32

注. このデータは実際のものではない。

次に、それぞれの反応の比率をもとめる。すると表 3 のようになる。

表 3.

反応の比率

$H$	$M$	$FA$	$CR$
.58	.42	.36	.64

ここから、弁別指標の  $d'$  をもとめる。手順は (26) のとおりである。

$$\begin{aligned}
 d' &= z(H) - z(FA) \\
 &= z(.58) - z(.36) \\
 &= 0.20 - (-0.36) \\
 &= 0.56
 \end{aligned}
 \tag{26}$$

この値は多少の弁別力を示すものと解釈される。ただし、この値は真値と等しいとは限らないことに注意が必要である。

次に、判断基準  $c$  を計算する。手順は (27) のとおりである。

$$\begin{aligned}
 c &= -\frac{(z(H) + z(FA))}{2} \\
 &= -\frac{(0.20 + 0.36)}{2} \\
 &= -0.08
 \end{aligned}
 \tag{27}$$

この値はやや負の方向にバイアスがあると解釈される。ただし、弁別力と同様に、この値は真値と等しいとは限らないことに注意が必要である。

また、この観察者の反応を ROC の空間に布置すると、図 4 のようになる。

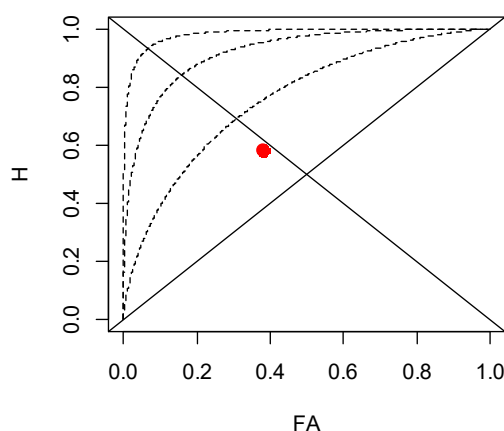


図 4. ROC の 2 次元空間における付置の例

信号検出理論指標の計算は、このように比較的容易であるが、計算を行うためのツールやソフトウェアも複数ある。まず、Macmillan and Creelman (2005) は、Microsoft Office Excel でこのような計算をおこなう方法を紹介している。統計環境 R (R Core Team, 2015) では、標準的な関数の組み合わせで計算ができるほか、信号検出理論による分析専用の *psyphy* パッケージ (Knoblauch, 2014) も公開されている。

さらに、第一著者のホームページ<sup>1</sup>では、Microsoft Excel による計算用シートと、R プラットフォーム上で動作する計算用関数を無償公開している。特に後者は、複数の観察者における反応を、古典的テスト理論のように正誤の 2 値で入力した行列と、刺激の特性を入力したベクトルを指定することで、本稿で紹介したほぼすべての指標を一覧として出力するだけでなく、ROC の 2 次元空間上に各観察者を付置した図も出力する。付録には、この関数のコードを掲載している。

このように、信号検出理論による分析を外国語教育研究者が導入することは容易であり、研究活動上の導入コストも高くない。

## 5. 信号検出理論の外国語教育研究における応用可能性

外国語教育研究では、信号検出理論をもちいた分析がそれほどの市民権を得ているとはいえない。昨今の国内研究では、音声、語彙や文法を扱う研究に一部研究例が見られるものの (e.g., Goto, 2016; Kusanagi, 2014)、未だ研究例は少ないといわざるをえない。

このことの理由として考えられるのは、第二言語習得研究といった関連諸分野においても、一部の研究領域を除いて、信号検出理論の数理的基盤が十分に浸透していないことがあげられる。第二言語習得研究は、従来から心理言語学や神経科学との交流が盛んであ

るものの、認知心理学、特に記憶や意識に関する研究、そして数理心理学との交流が比較的遅れている。心理言語学は、1990 年代後半より反応時間や読解時間といった行動データを、神経科学は、いうまでもなく従来、生理データを主として扱うため、判断成績などの分析手法が、これらの分野から第二言語習得研究を通して外国語教育研究に輸入されることはあまりなかった。

次に、信号検出理論や数理心理学を理解するためには、数理的な素養が必要のようにみえることも、理由のひとつだと考えられる。外国語教育研究者は、得てして数理的な職業訓練を受ける機会に恵まれないため、数理的アプローチを避ける傾向にあると考えられる。

しかし、これらのような状況は、今後変化していくと考えられる。現在の第二言語習得は、数理心理学はともかくとして、記憶や意識に関する研究との交流が徐々に活発になってきている。これは特に、人工言語学習や明示的および暗示的知識に関する研究領域に顕著な傾向であり、認知心理学における意識の研究に触発された第二言語習得研究や外国語教育研究も国内外において近年見られるようになった (e.g., Rebuschat, 2014; Tamura et al., 2016)。さらに、R といった極めて優れた研究ツールの普及が、信号検出理論といった数理的アプローチを取る研究者を増やしていくものと思われる。

ただし、信号検出理論による分析に適したデータが、必ずしも外国語教育研究で得やすいものとは限らない点に注意が必要である。信号検出理論を使用する数理心理学などは、反応パターンを主な関心とするため、典型的には、「小被験者数、大試行数」のデータを得る。一方、外国語教育研究やその関連諸分野では、被験者間分散やその共変量も関心となるため、被験者間分散やそれに対する回帰係数の推定の都合もあって、「中被験者数、中試行数」のデータを得がちである。また、学習者である被験者に、大きな実験的負担をかけることは憚られるのであり、さらに十分な被験者数の確保も他分野に比べ比較的困難であるので、結果として「小被験者数、小試行数」といったデータしか手に入らない場合も多い。このようなデータは、いうまでもなく信号検出理論の分析には向いていない。外国語教育研究に信号検出理論を応用するためには、まずは試行数を十分に増やすなどして、実験を適正に計画する必要がある。一般に信号検出理論による分析では、観察者毎の反応をプールすることは避けられているが、推定精度を上げるためにはやむを得ない場合もあると考えられる。いずれにせよ、このような処置も含めて、外国語教育研究独自の研究目的に合わせた適切な処置を選択するよう心がけるべきである。

## 6. 総括

本稿では、信号検出理論の数理的基盤、特に等分散正規分布モデルの導出を行い、外国語教育研究における応用可能性を論じた。信号検出理論は、従来までなされてきた分析よりも分析観点の多様化をもたらすという点において、手法上の利点が見込まれる。更に、

現在はツールの発展などで分析自体の導入コストが低くなってきている。実験計画などに多少の留意点は残るものの、概して、これからの外国語教育研究における有望な分析手法であるといえる。

注

1. <https://sites.google.com/site/kusanagikuni/home/tips/sdt>

## 参考文献

- Barrett, A. B., Dienes, Z., & Seth, A. K. (2013). Measures of metacognition on signal-detection theoretic models. *Psychological Methods, 18*, 535–552.
- DeCarlo, L. T. (1998). Signal detection theory and generalized linear models. *Psychological Methods, 3*, 186–205.
- Dienes, Z. (2016). How Bayes factors change scientific practice. *Journal of Mathematical Psychology*. Advance online publication. doi:10.1016/j.jmp.2015.10.003
- Goto, A. (2016). The effects of prosodic cues on auditory sentence processing: An analysis focusing on early and late closure. *LET Journal of Central Japan, 27*.
- Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.
- Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin, 75*, 424–429.
- Knoblauch, K. (2014). *psyphy: Functions for analyzing psychophysical data in R*. R package version 0.1–9. <http://CRAN.R-project.org/package=psyphy>
- 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.
- Lee, M. D. (2008). BayesSDT: Software for Bayesian inference with signal detection theory. *Behavior Research Methods, 40*, 450–456. doi:10.3758/BRM.40.2.450
- Macmillan, N. A., & Creelman C. D. (2005). *Detection theory: A user's guide* (2nd ed). Hillsdale, NJ: Erlbaum.
- R Core Team (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <http://www.R-project.org/>.
- Ratcliff, R. (1998). The role of mathematical psychology in experimental psychology. *Australian Journal of Psychology, 50*, 129–130. doi:10.1080/00049539808258788
- Ratcliff, R. (2002). A diffusion model account of reaction time and accuracy in a two choice brightness discrimination task: Fitting real data and failing to fit fake but plausible data.

*Psychonomic Bulletin and Review*, 9, 278-291.

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, 111, 159-182.

Rebuschat, P. (2014). *Investigating Implicit and Explicit Language Learning*. Routledge.

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*. *ARELE*, 27, 169-184.

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.

## 付録

### 信号検出理論用分析コードの例

```
sdt<-function(res,key,correc=T,plot=T){
  mp<-sum(key==1)
  mn<-sum(key==0)
  acc<-rowMeans(res)
  rt<-function(x,key){
    h<-sum(x[key==1]==1)
    m<-sum(x[key==1]==0)
    cr<-sum(x[key==0]==1)
    fa<-sum(x[key==0]==0)
    list(h,m,cr,fa)
  }
  t4<-
data.frame(t(matrix(unlist(apply(res,1,FUN=rt,key)),4,length(res[,1]
))))
  colnames(t4)<-c("Hit","Miss","CR","FA")
  r4<-data.frame(t4[,1:2]/mp,t4[,3:4]/mn)
  if(correc==T){
    r4[,1:2][r4[,1:2]==1]<-(mp-.5)/mp
    r4[,1:2][r4[,1:2]==0]<-.5/mp
    r4[,3:4][r4[,3:4]==1]<-(mn-.5)/mn
    r4[,3:4][r4[,3:4]==0]<-.5/mn
  }else{
```



```
    }
    p<-(t4[,1]+t4[,4])
    n<-(t4[,2]+t4[,3])
    rp<-(t4[,1]+t4[,4])/length(key)
    rn<-(t4[,2]+t4[,3])/length(key)
    posneg<-
data.frame("Positive"=p,"Negative"=n,"Positive.Ratio"=rp,"Negative.
Ratio"=rn)
    d.prime<-qnorm(r4[,1])-qnorm(r4[,4])
    criterion<-0.5*(qnorm(r4[,1])+qnorm(r4[,4]))
    beta<-exp(d.prime*criterion)
    A.prime<-1/2+( (r4[,1]-r4[,4])*(1+r4[,1]-
r4[,4])/(4*r4[,1]*(1-r4[,4])))
    B<-(r4[,1]*(1-r4[,1])-r4[,4]*(1-r4[,4]))/(r4[,1]*(1-
r4[,1])+r4[,4]*(1-r4[,4]))
    logOR<-log(r4[,1]*r4[,4]/r4[,2]/r4[,3])
    sdti<-data.frame(d.prime,A.prime,criterion,beta,B,logOR)
    if(plot==T){
    plot(r4[,4],r4[,1],xlim=c(0,1),ylim=c(0,1),xlab="FARatio",yla
b="Hit Ratio",pch=20, cex=1.5,main="ROC Space")
    for(i in 1:3){
        lines(pnorm(seq(-4,4,.01),i,1),pnorm(seq(-
4,4,.01),0,1),lty=2,col="blue")
    }
    for(j in 1:3){
        lines(pnorm(seq(-4,4,.01),0,1),pnorm(seq(-
4,4,.01),j,1),lty=2,col="blue")
    }
    abline(0,1,lty=2,col="blue")
    }else{
    }
    list("Accuracy"=acc,"Reaction.Table"=t4,"Reaction.Ratio"=r4,"
Positive.Negative"=posneg,"SDT.indices"=sdti)
    }
```