

証拠に基づく医療 (EBM) との比較を通じて 理論言語学の方法論を更に見直す*

What does **evidence-based medicine** (EBM) mean to theoretical linguistics?

黒田 航 Kow Kuroda

杏林大学 医学部 Medical School, Kyorin University

Created on 2015/07/08; Modified on 2015/07/11, 12, 15, 08/06, 07

1. はじめに

私が認知言語学に係わるようになって、二十年以上経過した。だが、私は今でも認知言語学者の言動に、事ある度に次の残念な態度を認める:

- (1) 認知言語学の基礎は (生成言語学の場合と違って) 認知心理学的洞察にあるのだから、基礎の妥当性を実験で確かめる必要がない¹⁾。

この問題を私は機会がある度に指摘して来た (e.g., (黒田 2005; 黒田 2007)) が、それに特に効果があったようにも見えない。その理由を自分なりに考えた結果、私なりに一つの結論に至った。それは次である:

- (2) 認知言語学の多数派とは、そもそも方法論の縛りを受けないで=好き勝手に研究した人たちなので、件の「墮落」は避けようがない。

状況を変える努力が空しいと言うつもりはないが、正直に言うと私も最近はどうでも良くなって来ていた。

同じやり方で成果が上がらないのは目に見えている。という訳で、本発表ではいつもと違う事をしてみようと思った。他の研究分野で起こった構造的に同じ現象から何か学べないか? と思った。以下ではそれを実践したい。

私が展開しようと思っている事は次の通り:

*本考察は、著者の Japan Cognitive Linguistics Association (JCLA), 2015 のワークショップ「方法論」での発表のために準備され、増補された。

¹⁾ここで私が念頭に置いている「実験」の意味は、板倉 (1992) で確認できる。

- (3) 医療的治療 (diagnosis) と科学的立証 (account) のアナロジーに基づいて言語学の将来を見通す

- a. 科学での説明 (理論) の選択と医療での治療 (方法) の選択は相同である。いずれも不完全な情報 (e.g., 交錯した証拠) が与えられている状態で、最適な説明法や対処法を選択する事に本質がある。
- b. 医療の基本方針が、直観/権威に基づいた医療 (authority/intuition-based medicine: A/IBM) が (十分な) 証拠に基づいた医療 (evidence-based medicine: EBM) に発展的したように、言語学の基本方針も、直観/権威に基づいた立証 (authority/intuition-based account: A/IBA) から、(十分な) 証拠に基づいた立証 evidence-based account: EBA) に発展する必要がある。
- c. そのためには言語学も、医療と同じように証拠の扱い方に関する方法論を確立すべきである²⁾。

その前に私が批判する予定の (認知) 言語学の研究のフォーマットを簡単に確認して置こう。

2. (認知) 言語学者はどんな風に研究しているのか?

本論を展開する前に、前提の確認のために、(認知) 言語学者の研究の進め方の典型をここで素描し

²⁾疫学の基本的な考え方を知るには、ロスマン (2002) を参考にした。

よう。

2.1 論証の基本形

(認知) 言語学者の様々なモノゴトの論じ方、いわゆる論証法を見た結果、私はそこに三つの典型を認める(ただし、心理実験に係わる研究は対象に含めていない)。それらは(4)と(5)と(6)に示す通りである。なお、(5)は検証という上位の研究サイクルに(4)を埋め込んだもので、(6)は(5)の変種である(これは弁証法(dialectics, dialectic process)と呼んでも良いのかも知れない³⁾)。

- (4)
- 説明すべきデータの提示: 幾つかの事例 $A = \{a_1, a_2, \dots, a_n\}$ を示す (n はせいぜい10の桁の数。下手すれば1の桁の数)。
 - 説明理論の提示: A を自分の提案する仮説 H が説明できると主張する(が、多くの場合、その主張の妥当性は直観に委ねられていて、妥当性のちゃんとした検証(e.g., 証明)がある訳ではない)。
 - 説明対象の追加: H で説明できる他の事例 $B = \{b_1, b_2, \dots, b_m\}$ を追加する。
 - 弱い結論: H は、導入を動機づけた事例集合 A だけでなく、他の事例集合 B も説明できるのだから、(動機づけになった基本データを説明するより)一般性が高いと主張する。
 - 強い結論: (4c)の事例の追加+(4d)の弱い結論のサイクルを k 回して ($k = 3 \sim 5$)、 H は十分に一般性の高い仮説だと結論する。

(5) 反証か確証が目的の論証

- 先行研究 R の示した仮説 H を示す(この際、 R が元の論文や研究書で挙げた例が念入りに引用されるが、 R の示した H の定義の引用はどちらかと言うとなおざり)。
- (追加) 確証(confirmation)が目的の場合

³⁾(5)を弁証法と呼んで良いかは不明である。

- H にうまく当てはまる例 $N = \{n_1, n_2, \dots, n_k\}$ を追加で示す。
 - H は正しいと主張する。
 - N が意外な事実である場合には、 H は正しいだけでなく、新しい応用=説明力があると主張する。
- c. 反証(disconfirmation)が目的の場合
- 説明すべきデータの提示: H にうまく当てはまらない例 $C = \{c_1, c_2, \dots, c_k\}$ を示す (k は1か10の桁)。 C のような例を H の反例(counterexamples)と呼ぶ。
 - H は正しくないと結論し、 C を(4b)の A と同じく扱う。
 - 説明理論の提示: C を説明できる仮説 H' (= H の対案)を示す(が、多くの場合、その主張の妥当性は直観に委ねられていて、妥当性のちゃんとした検証(e.g., 証明)がある訳ではない)。
 - 説明対象の追加: H' で説明できる他の事例 $D = \{d_1, d_2, \dots, d_l\}$ を追加する。
 - 弱い結論: H' は、導入を動機づけた事例集合 C だけでなく、他の事例集合 D も説明できるのだから、(動機づけになったデータ C の説明より)一般性が高いと主張する。
 - 強い結論: この事例追加のサイクルを k 回して ($k = 3 \sim 5$)、 H' は一般性の高い仮説だと結論する。

[(4)と(5)の手順のダイアグラム化したものを追加]

(5(c)v)と(5(c)vi)に関して注意すべき事が一つある。 H' が C, D, \dots を説明する事が(不十分ながらも)示されたとしても、 H が元々説明していた A, B, \dots を H' が(少なくとも H と同程度に)説明しているのかは、独立した検証が必要である。それなのに、この検証はしっかり実行されない傾向

があり、これが偽の正例や偽の負例の発生を招いている事は稀でない—と言うより、常態である。

想像するに、 H' が H の反証になっている事で、それが H より一般的になっている事が自明視される傾向があるのだろう。この自明視の傾向は言語学全体に汚染しているが、根拠がない。この事は明記に値する。

2.2 記述上の一般化

更に (5) の変種として、記述的一般化=記述上の一般化 (descriptive generalization) と呼ばれる手順を追加する。その手順は次の通りである。

- (6) a. 説明すべきデータの提示: 幾つかの事例 $A = \{a_1, a_2, \dots, a_n\}$ を示す。
- b. 説明理論の提示 1: A を仮説 H で説明できると主張する (が、多くの場合、その主張の妥当性は直観に委ねられていて、妥当性のちゃんとした検証 (e.g., 証明) がある訳ではない)。
- c. 説明すべきデータの追加: A と別に事例 $B = \{b_1, b_2, \dots, b_m\}$ を示す。
- d. 説明理論の提示 2: A の場合と異なり、 B は H で説明できず、それを説明するには自分の提案する別の仮説 H' が必要であり、かつ十分である事を示す (なお、 H と H' は相互に矛盾する事も場合も、矛盾しない場合もある)。
- e. 弱い結論 = 説明の統合:
 - i. H か H' の一方だけでは、 A と B の両方を説明する事はできない (従って、 H か H' かの二者択一には意味がない) 事を確認し、
 - ii. その上で、二つを統合する仮説 I があれば、 A と B のいずれも説明できる事を示す (が、厳密な証明がある事は稀)。
- f. 強い結論: I が、 H か H' のいずれか一方よりも一般的な説明を提供していると主張する。

[(6) の手順のダイアグラム化したものを追加]

(6e) の H と H' を基にして I を売るプロセスは弁証法的であり、そうして得られた I は、 H と H' の記述的一般化=記述上の一般化と呼ばれる。

記述的一般化に関する一般的な注意が一つある。(6e) の段階で、 I が偽の正例や偽の負例を発生させていない事を独立に確認する必要がある。それなのに、この検証はしっかり実行されない傾向があり、これが偽の正例や偽の負例の発生を招いている事は稀でない—と言うより、常態である。

想像するに、 I が H と H' の一般化になっている事で、自動的に説明力の向上がもたらされると錯覚される傾向がある。この自明視の傾向は言語学全体に汚染しているが、根拠がない。この事は詳しい検討に値するので、そうする。

2.3 一般化の (悪性の) 副作用 = 偽の正例と偽の負例の発生

一般化には多くの場合、悪性の副作用が付随し、しばしば見過ごされる。特に、研究者が不注意な場合、一般化の悪性の副作用は気づかれない。

2.3.1 偽の正例と偽の負例とは何か?

悪性の副作用は、偽の正例 (false positives) や偽の負例 (false negatives) の呼び込みである。それぞれの定義は次に示す通りである:

- (7) a. ある分類基準を設け、すべての事例の集合 U を幾つかのカテゴリ C_1, C_2, \dots, C_n に分ける。
- b. いずれかのカテゴリ C_i に属する例を C の正例 (positives) と言い、他のすべての例を C_i の負例 (negatives) と言う。
- (8) a. 本来の正例 x が正しく正例だと分類される場合、 x を真の正例と言う。
- b. 本来の正例 x が誤って負例だと分類される場合、 x を偽の負例と言う。

- c. 本来の負例 x が誤って正例だと分類される場合, x を偽の正例と言う.
- d. 本来の負例 x が正しく負例だと分類される場合, x を真の負例と言う.

2.3.2 過剰般化とは何か?

偽の正例か偽の負例のいずれか (あるいは両方) を発生させる一般化を, 過剰般化 (overgeneralization) と呼ぶ.

2.3.3 (4) 研究の実例

[補筆予定]

2.3.4 (5) 研究の実例

[補筆予定]

2.3.5 (6) 研究の実例

[補筆予定]

2.4 過剰般化の横行を見咎める

説明するまでもなく明らかな事だと思うが, 過剰般化は妥当な説明を提供しない. 実用性も低い. だが, 言語学では過剰般化が横行している.

横行の原因は何か? 私の見る限り, 言語学では過剰般化が咎められていないからである. いや, 咎められていないだけでなく, 一部では奨励されてすらいる. そんな状態で言語学のマトモな経験科学化が可能であるとは思われない. それが私が方法論的基盤を確立しなければダメだと思うに至った最大の理由である.

幸い, EBM をモデルにする事で, その目的は理論的な困難なく追求できる.

表 1: EBM が想定する証拠のレベルの分類

Level	内容と例
1a	無作為化のある比較治験 (RCT) のメタ分析 (meta-analysis)
1b	少なくとも 1 つの無作為化のある比較治験 (randomized comparative trial: RCT)
2a	ランダム割付を伴わない同時コントロールを伴うコホート研究 (前向き研究, prospective study, concurrent cohort study など)
2b	ランダム割付を伴わない過去のコントロールを伴うコホート研究 (historical cohort study, retrospective cohort study など)
3	症例対照研究 (case-control study, 後ろ向き研究)
4	処置前後の比較などの前後比較, 対照群 (control group) を伴わない研究
5	症例報告やケースシリーズ
6	専門家個人の意見 (専門家委員会報告を含む)

3. 証拠に基づいた医療 (EBM) に学ぶ

認知言語学であれ生成言語学であれ, 私は言語学の研究を見ていると, 方法論的基盤が極度に脆弱だと感じる. 例を挙げて具体的に言うと, 言語学では,

- (9) a. 調査で無作為化が減多に実施されない.
- b. 複数の説明の間のトレードオフが減多に比較されない.
- c. 自作例がコーパスで使用が確認された事例と同程度に信頼性のある証拠として扱われる⁴⁾.

なぜか? 私見では, それは言語学の証拠 (evidence) の扱い方に根本的に難点があるからである. それが方法論を主題にしたワークショップで私を取り上げたい論点である.

言語学では, 次の意味で証拠の扱い方が素朴過ぎる. 具体的に言うと, 次の通りである:

- (10) a. すべての証拠が同程度に信頼できる訳ではない事を知らず, どの証拠がどれくらい信用できるかに関する経験的見積りや蓄積がない.

⁴⁾例えば Usage-based model の提唱者が, 立証で自作例 (self-invented data) しか用いない.

b. その結果，信用できる証拠と信用できない証拠の区別がついていない．

(11) その事と相関して，立証が確認バイアス (confirmation bias) (Baron 1981; Mitroff 1981; Bierhoff and Klein 1989; Klayman 1995; Evans 1990; ギロビッチ 1993) にまみれている．

確認バイアスとは，人は証拠集めで無意識かつ体系的に，自説に好都合な事例に注目し，不都合な事例を無視する傾向の事である．

(11)の難点は別の機会 (e.g., (黒田 2011)) で指摘したので，この場では論点を (10) に絞る．以下では，論点を明確にするために，証拠に基づいた医療 (evidence-based medicine: EBM) を例に取って，アナロジーを展開しようと思う．

3.1 EBMの基本的な考え方

EBMは，臨床的意思決定 (clinical decisions) の支援を目的にしている．診断 (diagnosis) は臨床的意思決定の代表例だが，臨床的意思決定は診断に尽きる訳ではない．

Table 1⁵⁾ は EBM が認定している証拠の種類を，信頼度の順に並べたものである．信頼性の Level は大別で 1, 2, 3, 3, 5, 6 の六つ，細別で 1a, 1a, 2a, 2b, 3, 4, 5, 6 の八つある．

医療とのアナロジーに基づいて言語学の立証法について私見を述べる前に，基本的な考え方を説明しよう．

3.1.1 (臨床) 治験

治療法 T の (臨床) 治験 ((clinical) trial) とは， T の真の効果を測るための対照実験の事で，次のように実行される：

(12) a. 無作為に選ばれた被験者 P を対象群 P_1 と統制群 P_2 に分ける．

b. P_1 に治療 T を実施し， P_2 に別の治療 T' を実施する．i) T に効果がある事を確認するには， T' に偽薬 (placebo) を使う⁶⁾．ii) 効果が確認されている二つの治療 T_1 と T_2 を比較する時には， P_1 と P_2 に T_1 と T_2 を実施する．

3.1.2 コホート研究と症例対照研究

要因 f に関するコホート研究 (cohort study) とは，実世界で f の影響を受けた集団 P_1 と受けていない集団 P_2 を自然状態で一定の期間追跡調査し， f の影響を比較する研究．コホート研究は要因対照研究 (factor-control study) とも呼ばれる．

コホート研究には，調査が時間を逆上って行われる逆行=逆向きのコホート研究 (retrospective, historical) と，時間に沿って行われる順行=前向きのコホート研究 (prospetive) の区別がある⁷⁾．

因果関係 $f \rightarrow g$ を判定する症例対照研究 (case-control study) とは，(主に実世界で) 効果 g が観察されている時，その原因が f であるかを判別する手段である． f の影響を受けた集団 P_1 と受けていない集団 P_2 の g の発生率 r_1, r_2 を比較し， r_1 が r_2 より有意に大きいならば f に対して g の原因の一つであると結論する．

3.1.3 メタ分析 (meta-analysis)

メタ分析 (meta-analysis) は，複数の無作為化ありの治験同士，あるいはコホート研究同士を比較して，錯綜した結果を統合する作業の事である．体系的レビュー (systematic review) と同じ考え方で実行されるが，メタ分析の方が意味する範囲が狭い．

メタ分析の作業はなぜ必要なのか？それは，どんな測定 (観察だけでなく実験を含む) にも無作為な誤り — 疑陽性 false positives と偽陰性 false negatives — が不可避免的に含まれるので，その影響

⁵⁾Wikipedia の <<https://ja.wikipedia.org/wiki/根拠に基づいた医療>> の記事 (2015/07/11) を参考に作成．Table 2 も同様．

⁶⁾治療には偽薬効果 (placebo effects) という (不思議な) 現象があるので， T の効果を測る実験でも，無為でなくダミーの (薬効のないはずの) 治療を実行しないとけない．

⁷⁾言語研究でコホート研究に対応する研究がどんな研究なのか，私も見当がついていない．

を抑制する必要があるからである。一般論として、誤差の影響はサンプルの規模が小さい程大きくなる。

3.1.4 証拠の信頼度=根拠の強さの階層化

EBMの基準では、証拠は(13)にあるようにA, B, Cの三段階に評価される。ただし、Grade Aには少なくとも1つのLevel 1 (=1aか1b)の研究が、Grade Bには少なくとも1つのLevel 2 (=2aか2b)の研究がある事が条件となる。

(13) 「根拠の強さ」の分類と具体内容

Grade A: 言い切れる強い根拠あり; Grade B: 言い切れる根拠あり; Grade C: 言い切れる根拠なし

EBMでは、以上に示した体系的な方法論に従って、治療 T について Table 2 にあるように勧告がなされる。

以上で簡単に紹介したEBMのありようが、(認知)言語学に対してどんな意味を持つか考えよう。

3.2 EBMをモデルに言語学の実態を知る

EBMは臨床的意思決定(代表例が診断)の支援を目的にしている。ここで科学的立証(scientific account)と医学/臨床的診断(medical/clinical diagnosis)の間にアナロジーが成立すると考えてみよう。具体的に言うと、説明を妥当とし、受入れるという科学的意思決定が、診断を下すという医学的意思決定とアナロジー上の対応物だと考えて、それから何が言えるか考えよう。

適当な解釈の後なら少なくとも次が言えるはずである:

- (14) a. 特に第三者の検証を得ていない独自判断 = 自作例の容認度判断は(高名な研究者に拠るものであろうと)、実質的にLevel 6であり、言語学に権威主義が蔓延しているのは一目瞭然。
- b. 言語学の立証の大半は、Table 1のEBMの信頼性の水準で言うとLevel 5, 6であり、意欲的な立証でもLevel 4止まり。

(12)の無作為化された比較試験に相当する事を言語学で実践すると(15)になる。

- (15) a. 比較すべき二つの説明 T_1, T_2 がある(例えば T_1 が認知言語学による説明, T_2 が生成言語学による説明)。
- b. T_1 と T_2 が共通に説明できる現象を P とする。 P の事例を T_1, T_2 の説明と独立に = 理論に中立に選定し、これを p とする。
- c. T_1 による p の説明 $A_1(p)$ と T_2 による p の説明 $A_2(p)$ を独立の基準(e.g., 何らかの統計指標)で評価し、差の有無を検討する。
- d. メタ分析は、 P から p を選定する仕方を変え p_1, \dots, p_n を得て、それらに基づいて行われた結果を比較する事で実現できる。

なお、言語データのアノテーション(annotation)に無作為化実験と同様の価値がある事は指摘に値する。

3.2.1 アナロジーの根拠: 自然科学化の限界

医療と言語学の目的の違いを重視すると、私が描いているアナロジーが成立するか?と疑問をもつ人がいても不思議ではない。医療は応用科学の代表例であり、言語学は自然科学の一例であるか、そうでなくても人文社会科学の一例である。基礎科学の説明が失敗しても、人が死ぬ事は滅多にない⁸⁾。医療の目的は実用(厚生を増大は実用の最たるもので、その中でも人命救済は究極の厚生増大)であるが、科学の目的は実用ではないという差を考えると、それはもっともな懸念である。

その懸念が妥当だとしても、それは私が素描したアナロジーが成立する余地がないという事を意味しない。アナロジーが成立する条件は一つではない。医療と科学の間に目的の相同性と別の次元で相同性が成立すれば、二つの事象間にアナロジーは成立する。

⁸⁾絶対にはないとは言えないのだが、それが起こっても基礎科学に責任を問うのは難しい。

表 2: 推奨の強さの分類と具体的内容

Grade A	強い科学的根拠があり，行うよう強く勧められる
Grade B	行うよう勧められる
Grade C1	行うことを考慮しても良いが，十分な科学的根拠がない
Grade C2	科学的根拠がないので，勧められない
Grade D	無効性あるいは有害性の科学的根拠があり，行わないよう勧められる

具体的に言うと，医療は(16)に示す一点で，自然科学と一線を画している営為である．これを理由に私は社会科学や人文科学と医療との間に興味深いアナロジーが設定可能だと考える．

- (16) 医学では様々な理由(その一つが人道的な理由)により，比較実験や追加試験の実行可能性が強く制限されている．

医療科学は多くの自然科学と異なり，患者を比較実験の対象にできない．これは医療や医学の完全な自然科学化が不可能であり，かつ望ましくない事を意味する．

論を進めて，私はこの事が，医学や医療が社会科学(e.g., 経済学)や人文科学(e.g., 歴史学)と意外な形(権威主義の影響が大きい，デマの影響を受けやすいなどの点)で親和性が高い事の根本的理由でもあると推測する．

この事を私は重視し，言語学の方法論の充実を希望するなら，(17)の事を積極的に認め，その上で方法論を整備すべきだと主張したい．

- (17) 研究対象 X が歴史をもつ複雑系であるなら， X の研究の自然科学化=厳密科学化には限界がある(例えば X =言語)．

3.2.2 EBM の受容の二つの相

EBM の目標には(18)に示す二つの基調がある(二つは表裏一体な現象である):

- (18) 医学的判断をする際の
- 直観(intuition)への過度の依存からの脱却
 - 権威(authority)への依存からの脱却

(18a)は洗練された統計を用いる事の効用として語られ(e.g., (Ayres 2008))，(18b)はTable 3⁹⁾の威光に基づく医療(eminence based medicine)で揶揄されている．Table 1のLevel 4, 5, 6が直観基盤で，権威基盤の医療の温床だという事を指摘すれば事足りる．この点は理論言語学への示唆に富む事なので，詳しく説明する．

表 3: Basis of clinical practice

Basis for decisions	Marker	Measuring device	Unit of measurement
Evidence	Randomised controlled trial	Meta-analysis	Odds ratio
Eminence	Radiance of white hair	Luminometer	Optical density
Vehemence	Level of stridency	Audiometer	Decibels
Eloquence (or elegance)	Smoothness of tongue or nap of suit	Teflometer	Adhesin score
Providence	Level of religious fervour	Sextant to measure angle of genuflection	International units of piety
Diffidence	Level of gloom	Nihilometer	Sighs
Nervousness	Litigation phobia level	Every conceivable test	Bank balance
Confidence	Bravado	Sweat test	No sweat

3.2.3 EBM の特殊性

証拠に基づいた医療(EBM)は医療の世界に無条件に受け入れられた訳ではない．それが広く受け入れられるためには，疫学(epidemiology)の進展が必要だった．

疫学は相当に特殊な形で証拠を扱う．一言で言うと疫学は限られた証拠から最大限に活用する事を目的とした方法論である．ファング(2011, loc. 1003)の一節で上手に要約されているので，引用しよう．

⁹⁾Isaacs & Fitzgerald (1999) から引用．

疫学に携わる統計学者の重要な実績は、彼らが次のような難題に日々直面していることを考えると、いっそう感心させられる。

- 最小限のデータ (10 件を下回るケースから判断を下さなければならない)
- 緊急性 (患者が死にかけている)
- 不完全な情報 (聞き取り調査には「思い出せない」という回答も)
- 信頼できない情報 (人は想像で語るときがある)
- 原因を必ず見つけなければならない (この至上命題がさまざまな間違いにつながる)
- 間違いがもたらす結果の重大さ (説明するまでもない)

これは、統計学者にとって当たり前の世界ではない。他の統計学者はもっと寛大な状況に恵まれている。

- 豊富なデータ (文字どおり数百万件を分析する)
- 時間的余裕 (結論は繰り返し検証され、精緻なものになっていく)
- パターンだけに注目する (原因はほとんど気にしない)
- リスクが小さい (誰かが死ぬわけではない)

この緊張関係は、理論言語学者とコーパス言語学者との緊張関係を彷彿とさせて、示唆的である。

理論言語学者の仕事は、精神性の面では疫学者の仕事に近い。この対応づけが妥当なら、理論言語学者は疫学者の限られた証拠から最大限に活用する事を目的とした方法論から多くの事が学べると私は思う。

3.2.4 EBM の浸透から学べる事

実を言うと、医学界の EBM の受容は無条件ではない。それに反対する医師は今でも少なくない (Groopman 2008; Hickey and Roberts 2011)。批判には正当な面と不当な面がある。正当な面として、批判は診断の際の直観の重要性を説き、(18a) の行き過ぎに警鐘を鳴らしている。診断は現代科学で説明し切れない芸術 (art) の側面を強くもっている (Gawande 2003; Sanders 2010) ので、単なる技術 (technology) を見なすのがムリだからだ。医

療では直観がなしで済ませられないものなのである。稀な症例の診断の場合には、特にそうだ。

不当な面は、EBM の目標を (例えば AI による) 完全自動診断と誤って認定し、その限界を衝くものである。だが、その批判は EBM の本来の目的からすると藁人形論法である。EBM は臨床的な意思決定の支援のために存在するものであって、それに完全に取って替わるものだとは考えられていない。具体的には (19) のようなシステムであって、Step 1-Step 3 で完結する訳ではない。

(19) EBM を構成する 5 つの段階¹⁰⁾

- Step 1 目の前の患者についての問題の定式化
- Step 2 定式化した問題を解決する情報の検索
- Step 3 検索して得られた情報の批判的吟味
- Step 4 批判的吟味した情報の患者への適用
- Step 5 上記 1-4 の step の評価

だが、EBM に対する不当な批判と同質の批判が言語学でもコーパス言語学に対して生じているのは興味深い。コーパスは正例しか含んでいない (と言うより、素朴実在論の見地ではコーパスにある事例はすべてが正例扱いされなければならない)。正例だけを集めても負例との境界はわからない。負例は人手で作成しなければならない。これに直観の助けが必要なのは論を待たない。だが、それはコーパスの利用を控える理由にはならない。コーパスにある事例のみに依存した研究は、作例のみ依存するよりははまだマシだと言うべきである。権威主義 (と行き過ぎた直観基盤主義) は非科学的研究の温床だからである。

3.2.5 医療と人文科学の意外な共通性

医療科学と人文科学の意外な共通点を最後に一つ指摘して終わりにする。

- (20) 医療科学と人文科学は、稀な現象を無視しない事も美德にする (稀な疾患は無視して構わない疾患ではない。稀な現象は無視して構わな

¹⁰⁾Wikipedia の <<https://ja.wikipedia.org/wiki/根拠に基づいた医療>> の記事 (2015/07/11) を参考にした。

い現象ではない) . これは他の多くの自然科学
や社会科学の研究と異なっている点である .

参考文献

- Ayres, I. (2008). *Super Crunchers: Why Thinking-By-Numbers is the New Way To Be Smart* (Reprint edition ed.). Bantam. [邦訳: イアン・エアーズ『その数学が戦略を決める』(訳: 山形浩生), 文春文庫, 2010.]
- Baron, J. (1981). An analysis of confirmation bias.
- Bierhoff, H. W. and R. Klein (1989). Expectations, confirmation bias, and suggestibility. In V. A. Gheorghiu, P. Netter, H. J. Eysenck, and R. Rosenthal (Eds.), *Suggestion and Suggestibility*, pp. 337–346. New York: Springer.
- Evans, B. (1990). *Bias in Human Reasoning: Causes and Consequences*. Psychology Press. [refers to the confirmation bias?].
- Gawande, A. (2003). *Complications: Surgeon's Notes on an Imperfect Science*. Picador.
- Groopman, J. (2008). *How Doctors Think*. Mariner Books.
- Hickey, S. and H. Roberts (2011). *Tarnished Gold: The Sickness of Evidence-based Medicine*. CreatSpace Independent Publishing Platform.
- Isaacs, D. and D. Fitzgerald (1999). Seven alternatives to evidence based medicine. *The British Medical Journal* 319(7225), 1618. EBM.
- Klayman, J. (1995). Varieties of confirmation bias. *Psychology of Learning and Motivation* 32, 384–418.
- Mitroff, I. I. (1981). Scientists and confirmation bias. In R. D. Tweeney, M. E. Doherty, and C. R. Mynatt (Eds.), *On Scientific Thinking*, pp. 170–175. Columbia University Press.
- Sanders, L. (2010). *Every Patient Tells a Story: Medical Mysteries and the Art of Diagnosis*. Harmony.
- ギロピッチ, T. (1993). 人間この信じやすきもの: 迷信・誤信はどうして生まれるか. 新曜社. [原典: Thomas Gilovich (1993). *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life*, Free Press. Discusses confirmation bias].
- ファンク, カ. (2011). ヤバい統計学. CCCメディアハウス. [Funk, Kaiser: *Numbers Rule Your World*, 1st edition. McGraw-Hill Education, 2010. の矢野野薫による日本語訳.]
- ロスマン, K. J. (2002). ロスマンの疫学: 科学的思考への誘い. 篠原出版新社. [Epidemiology: An Introduction, Oxford University Press. の日本語訳 (2004 年出版).].
- 板倉, 聖. (1992). 新哲学入門: 楽しく生きるための考え方. 東京: 仮説社.
- 黒田, 航. (2005). 「言語の生得性」に関する (得てして不毛な議論に関する) 覚書: 拙論「認知言語学の言語習得へのアプローチ」の補遺. [URL: <http://clsl.hi.h.kyoto-u.ac.jp/~kkuroda/papers/why-innateness-matters.pdf>].
- 黒田, 航. (2007). 純粹内観批判: 生成言語学の対抗馬であるだけでは認知言語学は言語の経験科学になれない. URL: <http://clsl.hi.h.kyoto-u.ac.jp/~kkuroda/papers/critique-of-pure-introspection.pdf>.
- 黒田, 航. (2011). 自作例を使った研究の基礎. In 敬. 中本 and 在. 李 (Eds.), *認知言語学研究の方法: 内省・コーパス・実験*, pp. 29–63. ひつじ書房.