

言語表現の容認度とは何か？また何であるべきか？

言語学者であるはずなのに、

容認度判断が何であるかに自信をもって答えられない(大半の)人々への手引き

黒田 航

Created: 2011/3/21; Modified: 2011/4/3, 6, 7, 8, 9, 10, 11, 15; 2016/07/18

1 始めに¹⁾

この解説書では容認度評定 (acceptability rating/evaluation) の十分に一般的で曖昧性のない定式化を提示する。その定式化は文法性判断 (grammaticality judgment) の概念を含むくらい一般的なものである。

なお、本論文は黒田 (2009) の議論を発展させたものであり、黒田 (2011) の補遺として書かれた。必要に応じて黒田 (2009, 2011) を参照することをお勧めする。

1.1 逸脱性の評定値としての容認度

(理論) 言語学はしばしば架空の表現を相手にする。例えば (1) の対と (2) の対を比較し、(2b) が (1b) の表現に較べて逸脱しているという事実を記述したり、論証に利用したりする:²⁾

- (1) a. 彼は強盗に襲われた。
b. 彼は強盗に襲いかかられた。
- (2) a. 彼は不安に襲われた。
b. ?*彼は不安に襲いかかられた。

(2b) は (日本語の表現として見る限り) 逸脱している。

黒田 (2011) は (2b) のような例を反実例 (anti-examples) と呼んで、(1a), (1b), (2a) のような実例 (examples) から区別した。簡単に言うと、実例は用例が確

¹⁾本稿の執筆時に中村文紀 (慶応大学大学院) と吉川正人 (慶応大学大学院) からの意見が参考になった。この場を借りて感謝の意を表したい。それでも本稿に残る過誤の責任は筆者が一人で負うものである。

²⁾本稿で挙げる例は出典を明示しない限り、筆者による自作例である。

認できる実際の事例 (actual instances), 反実例は用例が確認できない架空の事例 (fictive instances) である。

反実例の利用は現代 (理論) 言語学ではあたり前の営為だが、その理論的基盤は必ずしも明確ではない。本文書の目的は、容認度評定の定式化を通じてそれを明らかにすることも含む。

1.2 逸脱の表わし方 (第一次近似)

言語学者は、その事実を記述するために特殊な記号を慣用的に用いる。具体的には、おおよそ次のような規約があるが、その評定の方法は体系化されておらず、評定結果も安定しているわけではない:

- (3) a. どんな人も、どんな文脈でも容認しないほど逸脱した表現に “*” を割当てる。
b. 多くの人が特殊な文脈を除いて容認しないくらい逸脱した表現に “*?” を割当てる。
c. 多くの人が文脈によって容認するくらい逸脱した表現には “?” を割当てる。

問題をはっきりさせるために、逸脱表現を分類する指針を示しておこう。

1.3 表現の分類

表現を分類する際に重要な特徴は、次の四つである:

- (4) a. 逸脱があるか? [±deviant]
b. 結果が意図的なものか? [±intended]
c. 意図は詩的なものか? [±poetic]

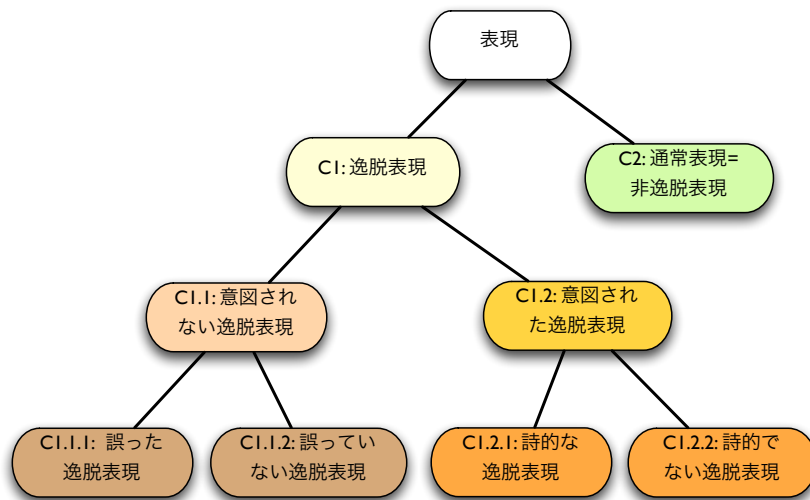


Figure 1: (4) の 4 特徴による表現の分類

d. 誤りを含むか? [±errorneous]

意図的に産出される場合と非意図的に産出される場合がある (図 1 の C1.2 と C1.1 の区別) . 更に, 意図的に産出される (図 1 の C1.2) 場合, 詩的な効果を狙った場合 (図 1 の C1.2.1) とそうでない場合 (図 1 の C1.2.2) がある . 非意図的に産出される (図 1 の C1.1) 場合, 誤りである場合 (図 1 の C1.1.1) と誤り未満である場合 (図 1 の C1.1.2) に分けられる .

具体例を示すと, (1) と (2) の四つの表現で, (1a), (1b), (2a) はいずれも C2 のクラスに属する表現である . これに対し, (2b) は C1.1.2 のクラスに属する表現である .

参考までに C1.1.1, C1.2.1, C1.2.2 の例も幾つか示しておこう:

(5) C1.1.1 の例

- a. *は彼に強盗れた襲わ .
- b. *襲われた強盗に彼は [(1a) と同義表現として]

(6) C1.2.1 の例

- a. ?幾時代かがありまして//?*茶色い戦争ありました (中原中也「サーカス」)
- b. 月が出た、//丘の上に人が立つてゐる。//??帽子の下に顔がある。(萩原朔太郎「蛙の死」)

c. ?その絵は医者にかかっていた .

(7) C1.2.2 の例³⁾

- a. ??疲れた時にでもクリック連打して時間をもてあそびます (Web から採取)
- b. ?*常に傍観者であればいいとの門外感さえもあつた (Web から採取)

これらはどれも逸脱しているが, 全部が同じように逸脱しているわけではない . 非文法性という概念との係わりがあるので, この点を明らかにしておこう .

1.3.1 (非) 文法性とは何か? また何であるべきか?

C1.1.1 は厳密に非文法的 (ungrammatical) と言ってよい表現であるが, その他の C1.1.2, C1.2.1, C1.2.2 の表現をすべて非文法的と言うのは非文法性の本来の定義から逸脱している . 少なくとも辞書を文法の一部として認めない限り, C1.1.2, C1.2.1, C1.2.2 が非文法性であると主張することはできない .

更に仮に C1.1.2 の例が非文法性だとしても, それは C1.2.1 と C1.2.2 の例が非文法的であることは保証しない . この意味で, 文法性判断で C1.1.2, C1.2.1, C1.2.2 の三つを区別しないのは方法論的に破綻している .

³⁾このような事例の扱いについては, 黒田・寺崎 (2010) を参照されたい .

1.4 課題の定式化

以上の予備的考察の下で、容認度評定の十分な定式化では次の (8) に挙げた事柄を問題にしよう:

- (8) a. 容認度を表現する値は {good, bad} (= {1, 0}) のみか? そうではなく程度をもつものであれば, それを表わすにはどうしたらよいか?
- b. 誰にとっての容認度か? 異なる二人の容認度はいつも同じだと期待してよいのか? 複数の評定者間の評定の変異は考慮しなくてよいのか?

以下, §2 で (8a) の程度の問題を, §3 で (8b) の変異の問題を扱う.

2 容認度評定とは何か?

2.1 準備

容認度が程度の表わすことができるように定式化するのに必要な幾つかの概念と用語を定義する.

2.1.1 用語の定義

表現 e が評定者 (rater) r にとって容認 (accept) できる程度 $a(r, e)$ を明示的に与えることを, “ r にとっての e の容認度 (acceptability) の評定 (rating/evaluation)” と呼ぶことにする.

$a(r, e)$ は尺度 m の上で表現される. 一般的な理解では m は間隔尺度 (interval scale) だが, それは順位尺度 (ordinal scale) であっても構わないし, 可能であれば比率尺度 (ratio scale) であってもよい. 後に §2.3.1 で順位尺度としての解釈を提示するが, 当面は間隔尺度だと理解しておいてよい.

2.1.2 例 1

評価者 r が, 刺激として与えられた表現 e に, r がそれを許容できる程度に応じて主観的に “容認可能” (good) か “容認不可能” (bad) のいずれかの評価を割当てるのは, 容認度評定の一種である.

2.1.3 例 2

評価者 r が, 刺激として与えられた表現 e に, r がそれを許容できる程度に応じて主観的に “完全に容認不可能” か “ほぼ容認不可能” か “ほぼ容認可能” か “完全に容認可能” のいずれかの評価を割当てるのは, 容認度評定の一種である.

2.1.4 例 3

評価者 r が, 刺激として与えられた表現 e に, r がそれを許容できる程度に応じて主観的に “5 点”, “4 点”, “3 点”, “2 点”, “1 点” のいずれかの点を割当てるのは, 容認度評定の一種である. この場合, 点が増えると容認度が単調に増すようにしてもよいし, 点が増えると容認度が単調に減るようにしてもよい.

2.1.5 例 4: 標準化された容認度評定

更に, 適当な操作によって $a(e)$ の最大値が 1.0 で, 最小値が 0 になるようにしたもの, つまり値域が 0 から 1.0 の閉区間 $[0, 1.0]$ に収まるようにしたものを, 標準化された容認度評定 (standardized acceptability rating/evaluation) と呼ぶ.

以下では標準化された容認度評定と標準化されていない容認度評定を区別する. 単に容認度を言った場合には標準化を前提としないとする.

2.1.6 文法性判断と例 1 の同一視

文法性判断と容認度評定は概念的には同一ではない. もっとも厳密な意味での文法性判断は語の意味を一切参照しないでなされなければならないが, それが現行の言語学の研究で実践されているようには見えない. 文法性判断と容認性評定は実質的に区別されていないということである. 従って, それらを同一視しても特に問題は生じないだろう.

以上の準備の下で, 容認度の一般的な表現を考える.

Table 1: 容認度 $a_{i,j}$ と $E \times C$ との対応づけ ($E = (e_1, \dots, e_n), C = (c_1, \dots, c_m)$)

	c_1	...	c_{j-1}	c_j	c_{j+1}	...	c_m
e_1	$a_{1,1}$...	$a_{1,j-1}$	$a_{1,j}$	$a_{1,j+1}$...	$a_{1,m}$
\vdots	\vdots	\ddots	\vdots	\vdots	\vdots	\ddots	\vdots
e_{i-1}	$a_{i-1,j-1}$	$a_{i-1,j}$	$a_{i-1,j+1}$...	$a_{i-1,m}$
e_i	$a_{i,j-1}$	$a_{i,j}$	$a_{i,j+1}$...	$a_{i,m}$
e_{i+1}	$a_{i+1,j-1}$	$a_{i+1,j}$	$a_{i+1,j+1}$...	$a_{i+1,m}$
\vdots	\vdots	\ddots	\vdots	\vdots	\vdots	\ddots	\vdots
e_n	$a_{n,j-1}$	$a_{n,j}$	$a_{n,j+1}$...	$a_{n,m}$

2.2 特定の r による表現 e の評定と文脈との関係

まず評定者の変異を無視した表現 e の容認度評定 $a(e)$ を考える。この際、 $a(e)$ への文脈効果を明示することが重要である。これをしないことが容認度を巡る言語学内の議論が混乱していることの原因である。

2.2.1 容認度の全体空間 (r を固定した場合)

特定の表現 e への文脈 c_1, c_2, \dots の影響をモデル化するために、次のように考える：

- (9) E と C の列挙可能性の条件: n 個の表現の集合 $E = (e_1, e_2, \dots, e_n)$ と m 個の文脈の集合 $C = (c_1, c_2, \dots, c_m)$ を考える。 E, C はいずれも有限個であり、かつ列挙可能であると想定する。
- (10) 文脈依存性の条件: 容認度評定 $a(e)$ は常に (e_i, c_j) という対に行われると考える。この時、対 (e_i, c_j) への容認度評定の結果を $a_{i,j}$ で表わす。

これらの仮定から、 $E = (e_1, e_2, \dots, e_n)$ のそれぞれの要素は、 m 個の文脈列 $C = (c_1, c_2, \dots, c_m)$ 上の容認度分布と対応づけられる。これは図式的には、表 1 で表わされる。具体例を §2.5 で示す。

(10) の文脈依存条件は少なくとも次を意味する：

- (11) a. $a(e)$ の適切な表現は $a(e_i, c_j)$ である ($e_i \in E, c_j \in C$)。これは本稿では主に $a_{i,j}$ と書かれる。

- b. 一般には同一の e_i が異なる文脈 $c_j \neq c_k$ で異なる容認度を持つ (i.e., $a_{i,j} \neq a_{i,k}$)。

(11a) は定義から明らかなので強調の必要はないが、(11b) については事情が異なる。実際のところ、筆者が知る限り、言語学では (11b) が認識されていない。これが容認度という用語に関する概念的混乱の原因の一つであることは確実だと思われる。

2.2.2 確率分布としての容認度

表 1 の $a_{i,j}$ を e_i のノルム (i.e., $\sum_{j=1}^m a_{i,j}$) で割って標準化すると、 e_i の文脈列上の確率分布が得られる。この分布は意味論の観点で特に重要である。

表 1 の $a_{i,j}$ を c_j のノルム (i.e., $\sum_{i=1}^n a_{i,j}$) で割って標準化すると、 c_j の表現列上の確率分布が得られる。この分布は言語の語用論=実用論 (pragmatics) の観点で特に重要である。

2.2.3 注意 1

(10) は容認度評定のモデル化の核心となる仮定である。これは真偽の不確かな想定ではあるが、それでも妥当性の正当化が必要なほど突拍子もない仮定だとは思われない。

2.2.4 注意 2: C の列挙可能性

方法論的により厄介なのは (9) の仮定である。これは強い仮定であるが、本稿では証明は与えられない。

C の列挙可能性に関しては特に課題が多い。 C の表示方法として何が適切を含め、これがうまく行くために必要になる多くのことがわかっていない。それでも、 C を中本ら (2005) や Nakakoto and Kuroda (2008) や Kuroda et al. (2007) が試みたように特徴空間として表わすのが有効だろうとは予測できる。

だが、それ以上に一般的な形で (9) が妥当であることを示すのは実際には相当に大変である。それは別に機会に試すことにする。従って、現時点では、(9) は完全に天下りの想定である。

概論として言うと、以下で見るように (9) からは有益な予測が生まれるが、その一方で、これは多くの言語学者の想定外であり、反発する人びとも多いと思われる。流派に拠らず、言語学者は表現の数は無限であり、文脈の数は無限であると考えたがる。 E や C がおのおの無限個存在するという仮定は、(9) よりも弱い仮定である。従って、他の条件が同じであれば (9) の方が望ましい。

2.3 容認度分布の構造

表 1 で記述される構造を評価するために尺度を導入する。この目的のために必要最低限な仮定は次の二つである:⁴⁾

(12) 表現 e_i について、文脈 c_1, c_2, \dots, c_m ごとの容認度 $a_{i,1}, a_{i,2}, \dots, a_{i,m}$ は半順序集合をなす。つまり、次のいずれかが成立する (ただし $j \neq k$):

- a. $a_{i,j} < a_{i,k}$
- b. $a_{i,j} = a_{i,k}$
- c. $a_{i,j} > a_{i,k}$

(13) 文脈 c_j について、表現 e_1, e_2, \dots, e_n ごとの容認度 $a_{1,j}, a_{2,j}, \dots, a_{n,j}$ は半順序集合をなす。つまり、次のいずれかが成立する (ただし $i \neq k$):

- a. $a_{i,j} < a_{k,j}$
- b. $a_{i,j} = a_{k,j}$

⁴⁾ 推移率 $a_{i,j} < a_{i,k}$ かつ $a_{i,k} < a_{i,l}$ ならば $a_{i,j} < a_{i,l}$ の成立は前提としている。

$$c. a_{i,k} > a_{k,j}$$

i を固定するにせよ j を固定するにせよ、容認度は表 1 で表わされる $E \times C$ の空間の中の順序尺度として理解できる。

2.3.1 容認度を順序尺度として理解する

文脈 c_j に対する表現 e_i の容認度 $a_{i,j}$ を測る尺度は通常は間隔尺度⁵⁾だと考えられているが、そうである理論的必然性はない。実際、ここでの記述は順位尺度であれば十分だと想定していることに等しい。

これから次の成立が期待できる:

- (14) a. 文脈 c_j と対応づけられた半順序集合 $(a_{i,j}, <)$ と別の文脈 c_k と対応づけられた半順序集合 $(a_{i,k}, <)$ は一般には異なる ($j \neq k$)。つまり $(a_{i,j}, <) \neq (a_{i,k}, <)$
- b. 表現 e_i と対応づけられた半順序集合 $(a_{i,j}, <)$ と別の表現 e_k と対応づけられた半順序集合 $(a_{k,j}, <)$ は一般には異なる ($j \neq k$)。つまり $(a_{i,j}, <) \neq (a_{k,j}, <)$

2.3.2 曖昧性の解消とは何か?

e を固定し c を特定するプロセスは、 e の曖昧性解消 (disambiguation) に対応する。これが必要なのは、 C 上に一定の閾値を越える容認度をもった E の要素が複数個存在するからである。

2.3.3 選択/生成とは何か?

曖昧性解消は聞き手の理解の一部をなす課題である。これと反対に c を固定し e を特定するプロセスは話し手の産出の一部をなす課題である。これは e の選択 (selection) や生成 (generation) と呼ばれる。これが必要なのは、 E 上に一定の閾値を越える容認度をもった C の要素が複数個存在するからである。

⁵⁾ なお、比率尺度を構成する方法が考えられないわけではないが、それは容認度の使用条件ではない。

Expression	Index	rain/rain (7a)	rain/fine (7b)	fine/rain (8a)	fine/fine (8b)	Average	Stdev
今日はいい天気ですね	(9a)	1	5	3	4	3.25	1.71
今日はあいにくの天気ですね	(9b)	4	3	5	1	3.25	1.71
昨日はいい天気でしたね	(10a)	1	3	5	4	3.25	1.71
昨日はあいにくの天気でしたね	(10b)	4	5	3	1	3.25	1.71
今日もいい天気ですね	(11a)	3	2	1	5	2.75	1.71
今日もあいにくの天気ですね	(11b)	5	1	2	3	2.75	1.71
昨日はいい天気ですね	(12a)	1	1	3	2	1.75	0.96
昨日はあいにくの天気ですね	(12b)	2	3	1	1	1.75	0.96
Average		2.63	2.88	2.88	2.63		
Stdev		1.67	1.60	1.60	1.67		

Figure 2: $a_{i,j}$ と $E \times C$ の対応 [$E = \{(19), (20), (21), (22)\}$, $C = \{(17), (18)\}$]: 皮肉 (irony) として解釈される場合には評定値として 3 を与えている .

2.3.4 原則「形が違えば意味が違う」の成立条件

(14a) が意味するのは「完全に同義な表現は存在しない」という原理 (anti-synonymity 条件) である . これは「形が違えば意味が違う」という原則と等価である . 別の言い方をすれば , 「形が違えば意味が違う」という原則は (14a) の帰結 (あるいは (14a) から得られる定理) である .

更に論を進めると次のことがわかる :

- (15) a. (14a) を満足しない場合 , E には完全に同義な表現の対が少なくとも一つ存在する . その場合 , E は冗長である .
- b. 言語がもっとも簡潔であるならば , それは (14a) は満足している必要がある .

2.3.5 (14b) の (不) 成立の効果

(14b) が満足されない体系で何が起るのかのは , よくわからない .

2.4 容認度評定で評定者が実際にやること

以上の考察から , e の容認度評定で評定者 r が実際にやっていることは , 次であると予想ができる :

(16) e の容認度評定で r は暗黙のうちに

- a. C の部分集合 C' を選択し ,

- b. C' の範囲でだけ $a(e, c_j)$ ($c_j \in C'$) を評定し ,
 $A' = \{a(e, c_1), a(e, c_2), \dots, a(e, c_{m'})\}$ を得る .
c. A' の最大値 $\max(A')$ を e の容認度として返す .

この時 , $\max(A')$ が閾値を越えるかは C' の大きさに依存する .

厄介なのは , §3.1.1 で述べるように , C' として利用可能な範囲が r によって違う可能性が高いという点である .

2.5 具体例

抽象的な定義だけではわかりにくいので , 具体例を挙げて説明しよう .

(17)–(18) に概略的に定義した 4 文脈 (あるいは前提条件) が与えられていると想定し , (19)–(22) に挙げた 8 表現の容認度がどうなるかを表わすとどうなるか?

(17) 話し手を聞き手が話している日を d , その前日を d^* とし , $W(d)$ は d の天気の色を表わすとする .

- a. d^* は雨天 , かつ d は雨天 ($W(d^*) = W(d)$)
b. d^* に雨天 , かつ d は晴天 ($W(d^*) \neq W(d)$)

(18) a. d^* は晴天 , かつ d は雨天 ($W(d^*) \neq W(d)$)

- b. d^* は晴天 , かつ d は晴天 ($W(d^*) = W(d)$)

(19) a. 今日はいい天気ですね .

- b. 今日はあいにくの天気ですね .
- (20) a. 昨日はいい天気でしたね .
- b. 昨日はあいにくの天気でしたね .
- (21) a. 今日もいい天気ですね .
- b. 今日もあいにくの天気ですね .
- (22) a. ??昨日はいい天気ですね .
- b. ??昨日はあいにくの天気ですね .

5段階評定法を使った容認度評定の結果の一例は表2のようになるだろう .

2.5.1 含意

表2を構成する容認度評定は筆者の判断による . これ为代表性をもつかどうかは独立検証の必要がある . とはいえ , 表2が妥当な表示であるなら , それは少なくとも次のことを意味する :

- (23) 行について容認度の値を比較するのは , 文脈群に対する特定の表現の適応度 (fitness) を測ることに等しい . その観点から次のことが確認できる :
 - a. 最大値 (e.g., 5) と最小値 (e.g., 1) は常に存在する .
 - b. 最大値 (e.g., 3) が閾値 (e.g., 3) を超えるかどうかは場合による . 例えば (22a) と (22b) の最大値は3で十分に高いとは言えない .
 - c. 文脈の異なりについて表現の容認度の平均値を比較すると , 容認度の高い表現 (e.g., (19a), (21a)) と低い表現 (e.g., (22a), (22b)) が区別できる . 容認度の分布の分散が大きいということは , 表現の使用文脈の選択性が高いということである .
 - d. 任意の二つの表現が異なる文脈で同じぐらい容認できる (あるいは容認できない) という可能性は常に存在する . 例えば (22a) と (22b) と

では1という値を与える文脈が二つあることに注意 .

- (24) 列について容認度の値を比較するのは , 表現群に対する特定の文脈の利用可能度 (availability) (=文脈の表現に対する選択性 (selectivity) の逆) を測ることに等しい . その観点から次のことが確認できる :

- a. 最大値 (e.g., 5) と最小値 (e.g., 1) は常に存在する .
- b. どの文脈にも最大値が十分に容認できる (e.g., 4以上) 表現が少なくとも一つ存在する . より一般的に言うと , どんな表現も十分に容認できるようにならない文脈は存在しないだろう . それが存在することは , おそらく言語表現の定義に反する .
- c. 表現の異なりについて文脈の容認度の平均値を比較すると , 利用可能性の高い=選択性の低い文脈 (e.g., (18a)) と低い文脈=選択性の高い (e.g., (18b)) が区別できる . 容認度の分布の分散が大きいということは , 文脈の選択性が高いということである .
- d. 任意の二つの表現が同一の文脈で同じぐらい容認できる (あるいは容認できない) という可能性は常に存在する . 例えば (17b) と (18a) では5という値を与える表現が二つあることに注意 .

2.5.2 (19a) の解釈の詳細

皮肉の解釈がどうなるかを考えよう . 例えば (19a) が皮肉として発話される場合を考えよう .

これが皮肉になるのは , 文脈が (18a) の場合である . この認識を特定の聞き手 h による e の曖昧性の解消だと考えた場合 , h の課題は , (17a), (17b), (18b), (18b) から最適な文脈を一つ選ぶことである .

Expression	Index	rain/fine (7b)	fine/rain (8a)	fine/fine (8b)	rain/rain (7a)	Average	Stdev
今日はいい天気ですね	(9a)	5	3	4	1	3.25	1.71
昨日はあいにくの天気でしたね	(10b)	5	3	1	4	3.25	1.71
昨日はいい天気でしたね	(10a)	3	5	4	1	3.25	1.71
今日はあいにくの天気ですね	(9b)	3	5	1	4	3.25	1.71
今日もいい天気ですね	(11a)	2	1	5	3	2.75	1.71
今日もあいにくの天気ですね	(11b)	1	2	3	5	2.75	1.71
昨日はあいにくの天気ですね	(12b)	3	1	1	2	1.75	0.96
昨日はいい天気ですね	(12a)	1	3	2	1	1.75	0.96
Average		2.88	2.88	2.63	2.63		
Stdev		1.60	1.60	1.67	1.67		

Figure 3: $a_{i,j}$ と $E \times C$ の対応 [$E = \{(19), (20), (21), (22)\}$, $C = \{(17), (18)\}$]: 皮肉 (irony) として解釈される場合には評定値として 3 を与えている。

(17b) と (18b) は字義通りの解釈である。だが、発話者 s が最適な選択をしたと想定する限り、(18b) よりも (17b) が選好される必要がある。それは (18b) を選んだとき、その条件下で (19a) は (21a) よりも容認度が低いからである。§2.5.3 で後述するが、 C と E の選択の相関から最適な解釈と選ぶという方法は、双方向的最適性理論 (Bidirectional Optimality Theory) の考え方と同じである。

次に (19a) が皮肉として理解される場合を考える。この表現が皮肉になるのは、文脈として (18a) が選ばれる場合である。文脈に (17b) か (18b) が選ばれた場合に皮肉にならない理由は皮肉の定義から明らかである⁶⁾。

ここで考えるべきことは、文脈に (17a) を選ぶことが皮肉にならない、少なくとも文脈に (18a) を選ぶ場合に比べて皮肉になりにくいという点である⁷⁾。それはなぜか?

提案したモデルでの説明は単純明快である。 e を (19a) に固定した条件で、(17a) との対の容認度=1 は (18a) との対の容認度=3 より低いからである。これは助詞「は」が対比を含意をもち、昨日の天気と今日の天気が違うという含意が成立するからである。

もちろん「表 3 の容認度分布がそうなるように仕組み

⁶⁾ 正確には文脈に (18b) が選ばれた場合に皮肉にならない理由はそれなりに複雑である。少なくとも助詞「は」が含意する対比に矛盾を読みこんだ場合、皮肉の要素がないわけではない。その説明はここでは試みないで、意欲ある読者の課題とする。

⁷⁾ 先行研究を網羅的に調べたわけではないが、筆者が知る限り、この点を解明した先行研究はない。

れたからそうではないか?」という論点先取の難点はあるかも知れない。だが、表 3 を筆者が構成している時、特に皮肉の解釈の説明のことは考えていなかった。やったことは、単に (19a) との (17)–(18) にある 4 文脈との対への容認度評定の結果が (19a) について半順序集合になるようにしただけである。

この段階で無意識の作為を否定することは不可能だが、そんな辻褃合わせめいたことをしなくても自然な説明が得られる体系であると期待する方が実りが多いと筆者は考える。それはここで取り上げた (19a) だけでなく、先に挙げた例の他の場合や取り上げていない別の表現を同様に検討すれば明らかになるだろう。

2.5.3 関連研究と関連分野

以上のように表現のインデックス i と文脈のインデックス j の直積で特徴づけられた容認度分布は、R. Blutner が提唱している双方向最適性理論 (Bidirectional Optimality Theory: BOT) (Blutner and Zeevat 2004; Blutner, Hoop, and Hendriks 2006) と共通するところが多い。双方向の最適化は $a_{i,j}$ の i と j の両方の次元で最適化を求めることに等しい。このことからわかるように、 $a_{i,j}$ の活用は意味論研究に留まるものでなく、逆に語用論研究への示唆が大きい。この点で、多くの言語学者が要請する意味論と語用論の分離は、実現不可能な単なる理想論であ

る可能性が示唆されているとも言える。

2.5.4 $E \times C$ の局所的な調節ゲームへの分解

表 2 の容認度分布は、 E 成分と C 成分の適当な交換により表 2 のように、局所的な調節ゲーム (coordination games) (Lewis 1969; Sugden 2004) の和集合として表わすことが可能である (ただし、表 2 と表 3 の違いは成分の交換だけなので、両者の数学的構造は同一である)。成分の交換がによって局所的な調節ゲームを見出せるということは、全体が多価値的な空間にあるということである。

2.5.5 記述ゲーム

表 2 の構造をゲーム理論 (Bacharach 1976; Cubitt and Sugden 2003; Blackwell and Girshick 1954; von Neumann and Morgenstern 1972; Maynard-Smith 1982; Sugden 2004) の対象としての“ゲーム”として解釈するには参加者 (“プレイヤー”) を認定しなければならない。十分に精緻な定式化ではないが、ここで案だけを述べる。

自然科学者 (e.g., 物理学者, 生物学者) が自然を相手に記述ゲーム (description game) をしていると考える。この時、記述は E が記述のための表現 (expressions) で、 C が記述される真実 (truths) であるようなゲームとなる。記述ゲームは調節ゲーム (Lewis 1969; Sugden 2004) の一種である。

2.6 容認度を比較する二つの次元

表 1 の $a_{i,j}$ で表現される容認度分布を分析する方法は二つある: (A) i を定項で j を変項だと考える方法と、(B) i を変項で j を定項だと考える方法の二つである。A は特定の表現の容認度が文脈によって変化することを意味し、B は特定の文脈で最低な表現が変化することを意味する。十分な容認性判断のためには、これら二つの変化を同時に追いかけることが必要である。このことを念頭に置きながら、絶対低な容認不可能性という概念の意味を検討する。

2.6.1 容認不可能な表現は存在するか?

(22a) や (22b) のような表現の扱いが問題になる。これに関して、次の問いに答えを出そう:

(25) 絶対に容認不可能な表現は存在しないか?

これに答えるために考慮すべきことは、次である:

(26) 文脈の数を十分に増やせば、(22a) や (22b) のような表現の容認度が 5 や 4 になるような特定の文脈が存在する。

事実として言えば、(26) の主張はおそらく正しい。具体的に言うと、(27) の文脈では (22a) や (22b) の容認度は 5 になる:

(27) A が B に報告するために遠隔地 P の天気を数ヶ月分モニターしている。記録を見ていると昨日の前まではずっと雨天 (か晴天) が続いていた。昨日になって天気が雨天から晴天に (あるいは晴天から雨天に) 変わったのを、A が B に (22a) (か (22b)) と行って報告する。

この事実は (25) を確証する。

だが、それで問題は解決するのだろうか? 実はそうではない。それを以下で確認する。

2.6.2 C について容認度を比較する

言語 L の表現の全体集合を $E = \{e_1, e_2, \dots, e_n\}$ とし、文脈の全体集合を $C = \{c_1, c_2, \dots, c_N\}$ とする。 $e_i \in E$ と $c_j \in C$ の対の容認度を $a_{i,j}$ とする。この時、(25) が述べているのは (28) である:

(28) i を固定し j を可変とした時に、

- a. $a_{i,j}$ の値は j について半順序集合をなし、
- b. i ごとに十分に $a_{i,j}$ 大きくなるような j が少なくとも一つ存在する。

これは表 1 で $a_{i,j}$ を行ごとに比較することに相当する。

ここで、 j を固定し i を可変としたらどうなるかを考えるのは重要である。

2.6.3 E について容認度を比較する

考察すればすぐにわかるように, $a_{i,j}$ について, i が可変で j を固定の時には (29) が成立する:

(29) j を固定し i を可変とした時に,

- a. $a_{i,j}$ の値は i について半順序集合をなし,
- b. j ごとに十分に $a_{i,j}$ が大きくなるような i が少なくとも一つ存在する.

これは表 1 で $a_{i,j}$ を列ごとに比較することに相当する.

(29) は完全な容認不可能性の否定を述べた (25) の真偽にかかわらず真である. 従って, 仮に (25) が真であっても (29) を記述することは必要であり, かつ有意義なことである.

2.7 まとめ

これまでの考察が意味することは次である:

(30) (25) が真で, 絶対的に容認不可能な表現が存在しないというのが事実だとしても, それとは独立に (29) が述べる容認度の順序づけを説明しなければならない.

より具体的に言えば, 言語科学者がすべきことは, 表 1 で表わされる $E \times C$ の空間中の容認度 $a_{i,j}$ の分布を考えることである. そのために, 特定の表現 e の容認度を数多くの異なる文脈 c_1, c_2, \dots と対応づけて評価するだけでなく, 特定の文脈 c_j で e がどれぐらいの適応度 (degree of fitness) をもつか, すなわち他の候補表現 e_1, e_2, \dots に較べて e_i がどれぐらいの競争力をもつかを評価する必要がある.

最期に評定者の変異の扱いについて考察する.

3 評定者の変異の扱うための拡張

$a_{i,j,k}$

以上の考察では評定者の変異は捨象されていた. だが, 一般に異なる評定者 r, r' が同一の表現 e に与えた容認

度評定の結果 $a(e, r), a(e, r')$ が同一になる保証はどこにもない. これは多人数の行動データを利用したデータ主導の研究では致命的な欠陥となる. 本稿は最期にこの問題を扱うが, 議論は十分に掘り下げない.

3.1 変異の原因の関する短い考察

理論言語学の多くの流派は, 異なる評定者 r, r' が同一の表現 e に与えた容認度評定の結果 $a(e, r), a(e, r')$ が同一でないという経験的事実に対して, 次の二つの極端な反応をする:

- (31) a. $a(e, r) \neq a(e, r')$ であるという事実をしばしば (言語学の経験科学化に不可欠な理想化 (idealization) と称して) 単純化のために無視する.
- b. $a(e, r) \neq a(e, r')$ は自明だと言って, まじめにモデル化しない.

理由こそ違おうが, どちらも容認度評定の個体変異のモデル化ができていない点では変わりがない. 従って, 優劣はつけようがない.

以下ではこの問題を解決するための定式化をするが, その前に理由を考察しておくのは有意義である.

3.1.1 $a(e, r) \neq a(e, r')$ となる理由

それにしても, 異なる評定者が同一の表現に異なる容認度を与えるのはなぜなのか?

この原因として考えられるのは, 次の二つである:

- (32) 表現 e を固定しても, 評定者 r と評定者 r' とは,
 - a. c_1, c_2, \dots, c_m と e との対応づけを容認度に対応させる評価関数 $a(e_i, c_j)$ が異なる.
 - b. 評価関数 $a(e_i, c_j)$ が同一であっても, r は r' は異なる C 評価に使っている範囲が異なる. 具体的に言えば, r と r' はいずれも C の全体を考慮していない.

方法論的には (32a) は (32b) よりも強い条件であり、他の条件が同じであれば、(32b) による説明を優先すべきである。従って、次のように仮定することは合理的だろう:

(33) 評定者は一般に、 e の容認度評定で C の部分空間 $C' \subset C$ しか考慮していない。

3.1.2 確証バイアスの影響

確証バイアス (confirmation bias) (ギロピッチ 1993) は (32a) と (32b) の二つの次元で影響する。一般に (32a) は (32b) による説明が十分でない場合にのみ考慮されるべきであるが、確証バイアスが強い場合にはその原則が成立しなくなる可能性がある。確証バイアスは言語学の一部の学派では常習性のある傾向なので、この点には特に注意が必要である。

3.1.3 (33) への対処法

ここで気をつけるべきなのは次の点である:

(34) e の容認度評定で、評定者 r と評定者 r' が利用している文脈をそれぞれ D と D' とする ($D, D' \subset C$ で $D \neq D'$)。

- a. r の e の容認度分布は $a_{i,j}$ ($j \in D$) であり、 r' の e の容認度分布は $a_{i,j'}$ ($j' \in D'$) である。
- b. $D \neq D'$ であるならば、 $a_{i,j} \neq a_{i,j'}$ は自然な結果である。

$a_{i,j} \neq a_{i,j'}$ なのであれば、 r と r' の容認度評定を一つの代表値 (e.g., 平均値か最大値) で近似した場合に、それらが異なるのは異常な事態ではない。

(33) が原因となって生じる容認度分布の個体変異を減らすためになすべきことは、刺激提示の際に C' を明示することである。

とはいえ、本稿ではそのための方略を示すことはしない。それを実現するためには十分に強力な C の指定法/表示法の確立が必要である。それは現状では望みがたい。

その代わりに、以下では一定の誤差の範囲で (32a) が妥当である場合のモデル化を示す。

3.2 $a_{i,j,k}$ への拡張による $a_{i,j}$ への r の変異の取り込み

r と r' が C 全体を使って容認度評定をしていると想定した上で考慮すべきことは、評価関数それ自体の変異である。この場合には、評価関数の変異をモデル化すれば十分である。

表 1 で表わされるのは、 $a_{i,j}$ を (特定の評定者 r にとっての) 表現 e_i の文脈 c_j での容認度である。これを r を変項化して一般化する。

3.2.1 C 上成分のベクトル化

C に対応づけられた e_i の容認度評定の列を $A_i = (a_{i,1}, \dots, a_{i,j}, \dots, a_{i,m})$ (m は C の要素の数) とし、評定者 $r_k \in R$ の異なりの次元 R を追加すると、 $a_{i,j,k}$ を構成できる。これによって評定者の違いによる表現 E の容認度分布の変異を考えることができる。

具体的には、容認度は n 個の表現列 E_n 、 m 個の文脈 C 、 N 個の評定者列 R を次元とする 3 次元空間中の (確率) 分布 $a_{i,j,k}$ となる。これは表 2 で表わされる。この時、 C の効果は A によってベクトル化されているので、文脈の異なりは A に隠れている。

3.2.2 $a_{i,j,k}$ の幾何学的イメージ

幾何学的に言うと、表 2 が表わす $a_{i,j,k}$ は 3 次元のうち、 A が奥行きに相当する次元 k を表わす。別の言い方をすると、表 2 は表 1 の表わす平面が N 枚積み重なった状態であると理解するとよい。

3.3 R を追加して何がわかるようになるか?

表 2 の表わす構造で追跡可能になるのは、 $i \in E$ と $k \in R$ の次元の (共) 変動と $j \in C$ と $k \in R$ の次元の (共) 変動である。これらを追跡することは何を教えてくれるのか?

$a_{i,j,k}$ の三つの次元 i, j, k は独立に動かすことができるが、言語学では k で表わされる評定者間の (共) 変動は捨象して i と j の (共) 変動に注目することが多い。これは $a_{i,j}$ が特定の評定者の内部モデルを表現すると理解

Table 2: $A_{i,j}$ と $R \times E$ の対応 ($R = (r_1, \dots, r_N)$, $E = (e_1, \dots, e_m)$)

	e_1	...	e_{j-1}	e_j	e_{j+1}	...	e_m
r_1	$A_{1,1}$...	$A_{1,j-1}$	$A_{1,j}$	$A_{1,j+1}$...	$A_{1,m}$
\vdots	\vdots	\ddots	\vdots	\vdots	\vdots	\ddots	\vdots
r_{i-1}	$A_{i-1,j-1}$	$A_{i-1,j}$	$A_{i-1,j+1}$...	$A_{i-1,m}$
r_i	$A_{i,j-1}$	$A_{i,j}$	$A_{i,j+1}$...	$A_{i,m}$
r_{i+1}	$A_{i+1,j-1}$	$A_{i+1,j}$	$A_{i+1,j+1}$...	$A_{i+1,m}$
\vdots	\vdots	\ddots	\vdots	\vdots	\vdots	\ddots	\vdots
r_N	$A_{N,j-1}$	$A_{N,j}$	$A_{N,j+1}$...	$A_{N,m}$

され、言語学者の関心が言語の心内モデルの記述や説明であることが多いからである。

とはいえ、 k の変動に言語学的な意味がないということではなく、 k と i の (共) 変動と k と j の (共) 変動とは社会言語学 (sociolinguistics) の研究では重要な要素になる。ただ、この可能性は筆者の技量の限界もあって、本稿では具体的に検討することができない。

(37) = F2

- a. 太郎が 校庭を 走る
- b. 稲妻が 北の空を 走る
- c. 戦慄が 永田町を 走る
- d. 汗が 太郎の額を 走る

3.4 $E \times R$ 上の $a_{i,j,k}$ の具体例⁸⁾

抽象的な説明だけでは理解が難しいと思うので、 $a_{i,j,k}$ の具体例の一つ挙げよう。ただし詳細な統計的な解析はしていない。

16人の評定者に(35)の4段階の規準で(36)–(39)の12文を評定してもらった結果を表4に示す。これは $E \times C \times R$ の $E \times R$ 上への写像である。

(38) = G1

- a. 校庭を 太郎が 走る
- b. 北の空を 稲妻が 走る
- c. 永田町を 戦慄が 走る
- d. 太郎の額を 汗が 走る

(35) 3点: まったく違和感を感じない

2点: 軽く違和感を感じるが、言おうとしていることは簡単にわかる

1点: 強く違和感を感じるが、言いたいことがまったくわからないわけではない

(39) = G2

- a. 太郎が 校庭に 走る
- b. 稲妻が 北の空に 走る
- c. 戦慄が 永田町に 走る
- d. 汗が 太郎の額に 走る

0点: 何を言っているかわからないか、明らかに異常なことを言っていると思う

(36) = F1

- a. 校庭に 太郎が 走る
- b. 北の空に 稲妻が 走る
- c. 永田町に 戦慄が 走る
- d. 太郎の額に 汗が 走る

⁸⁾本節で使用するデータは筆者が2006年にMorphology and Lexicon Forumで口頭発表した研究で使ったデータに新規のデータを追加したものである。

Expression	Index	r1	r2	r3	r4	r5	r6	r7	r8	r9	r10	r11	r12	r13	r14	r15	r16	Av.	Stdev	
太郎が 校庭を 走る	F1a	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	0.0	
永田町に 戦慄が 走る	G2c	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	1.0	2.9	0.5	
北の空に 稲妻が 走る	G2b	2.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	1.0	2.8	0.5
校庭を 太郎が 走る	F2a	3.0	3.0	3.0	3.0	3.0	3.0	3.0	2.0	2.0	3.0	2.0	3.0	2.0	2.0	3.0	2.0	2.6	0.5	
北の空を 稲妻が 走る	F2b	3.0	3.0	3.0	3.0	3.0	3.0	2.0	3.0	2.0	2.0	1.0	3.0	2.0	2.0	2.0	3.0	2.5	0.6	
稲妻が 北の空を 走る	F1b	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	1.0	2.0	2.0	2.0	2.0	1.0	1.0	3.0	2.4	0.8	
戦慄が 永田町に 走る	G1c	3.0	2.0	3.0	3.0	3.0	1.0	3.0	1.0	2.0	2.0	1.0	3.0	2.0	1.0	3.0	1.0	2.1	0.9	
永田町を 戦慄が 走る	F2c	3.0	2.0	3.0	1.0	1.0	1.0	2.0	3.0	2.0	3.0	3.0	2.0	1.0	2.0	2.0	2.0	2.1	0.8	
戦慄が 永田町を 走る	F1c	3.0	3.0	3.0	1.0	2.0	2.0	3.0	3.0	1.0	2.0	1.0	1.0	1.0	1.0	2.0	3.0	2.0	0.9	
太郎の額を 汗が 走る	F2d	3.0	3.0	3.0	2.0	3.0	2.0	2.0	2.0	2.0	1.0	1.0	3.0	1.0	1.0	1.0	2.0	2.0	0.8	
稲妻が 北の空に 走る	G1b	3.0	3.0	3.0	3.0	3.0	2.0	3.0	1.0	2.0	0.0	2.0	1.0	3.0	1.0	1.0	1.0	2.0	1.0	
太郎の額に 汗が 走る	G2d	1.0	3.0	3.0	2.0	3.0	2.0	3.0	3.0	3.0	1.0	2.0	2.0	1.0	2.0	0.0	1.0	2.0	1.0	
汗が 太郎の額を 走る	F1d	3.0	3.0	2.0	1.0	3.0	3.0	2.0	1.0	2.0	1.0	2.0	2.0	0.0	1.0	0.0	2.0	1.8	1.0	
汗が 太郎の額に 走る	G1d	3.0	2.0	2.0	2.0	3.0	1.0	2.0	1.0	2.0	1.0	1.0	2.0	1.0	1.0	3.0	3.0	1.8	0.8	
太郎が 校庭に 走る	G1a	3.0	1.0	0.0	1.0	1.0	3.0	2.0	1.0	1.0	0.0	0.0	0.0	1.0	3.0	3.0	3.0	1.4	1.2	
校庭に 太郎が 走る	G2a	2.0	1.0	0.0	1.0	1.0	1.0	1.0	1.0	1.0	0.0	2.0	1.0	0.0	1.0	3.0	1.0	1.1	0.8	
	Av.	2.8	2.6	2.5	2.2	2.6	2.3	2.5	2.1	2.0	1.7	1.8	2.1	1.6	1.8	1.9	2.0			
	Stdev	0.6	0.7	1.0	0.9	0.8	0.9	0.6	1.0	0.7	1.1	0.9	1.0	1.0	0.9	1.1	0.9			

Figure 4: $E \times R$ の例: 評定の平均値の降順に並べている

3.4.1 表 4 の容認度分布の個体変異が示唆すること

表 4 の容認度評定の分布を見ると、少なくとも次のことがわかる:

- (40) a. F1a のような理想的な反応 (標準偏差=0) もあるが、これは例外である。
- b. 行の標準偏差を見ればわかるように、評定者間の値の分散は決して小さくない。Stdev の最小値は F1a の 0 だが、最大値は G1a の 1.2 である。
- c. もっとも容認度の平均が低い G2a でも 3 点をつける評定者がいる。
- d. r1 のように相当に寛容な評定者もいれば、r10、r13 のように非寛容な評定者もいる。これが意味することは、評定者が一人しかいないような容認性判断はまったく信頼できないということである。

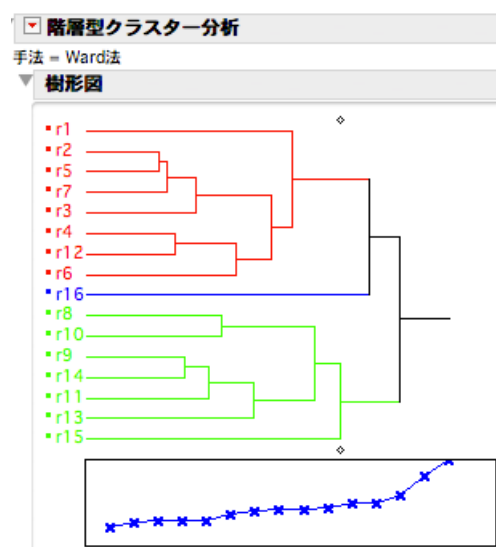


Figure 5: 表 4 のデータに基づく R のクラスター化 (JMP 5.0.1J で作成)

3.4.2 評定者のクラスター化

表 4 に示した容認度評定の結果を使って評定者を分類することもできる。例えば表 4 に示したデータを階層的クラスター分析 (Ward 法) にかけた結果が図 5 の分岐図である。分析では $E = \{F1a, F1b, \dots, G2d\}$ を特徴として $R = \{r1, r2, \dots, r16\}$ をクラスター化している。

三つのクラスターが得られているが、それが何を意味するのは必ずしも明らかではない。ただ、§3.3 で述べたことを思い出すと、世代や階級や方言の違いがクラスター化に影響している可能性はある。それを調べるにはもっと個人の情報が必要になる。

3.4.3 容認可能/容認不可能の二者択一は可能か?

表 4 の容認度評定の分布から、平均評定値が 2.0 を越える F1a (3.0 [0.0]), G2c (2.9 [0.5]) G2b (2.8 [0.5]), F2a (2.6 [0.5]), F2b (2.5 [0.6]), F1b (2.4 [0.8]), G1c (2.1 [0.9]), F2c (2.1 [0.8]) が容認可能な表現とし、平均評定値が 1.5 に満たない G1a (1.4 [1.2]), G2a (1.1 [0.8]) を容認不可能判断するのが妥当な判断ではないかと思う。なお、容認度の平均値が 1.5 と 2.0 の間にある F1c (2.0 [0.9]), F2d (2.0 [0.8]), G1b (2.0 [1.0]), G2d (2.0 [1.0]), F1d (1.8 [1.0]), G1d (1.8 [1.2]) を容認可能か不可能かの二者択一で決めるには難しい。少なくともそれをするための実証的な証拠が理由が十分だとは言えない。

3.4.4 直観による解析と統計処理による解析のズレ⁹⁾

以上の評価は(それなりに熟練した)言語学者としての筆者の判断だが、それはどれくらいデータの統計解析に支持されるものだろうか? これを確認するために R を F1a, F1b, ..., G2d を独立変数としてクラスター化したのと同じに、 E を r_1, r_2, \dots を独立変数としてクラスター化できる。その結果を図 6 に示す。

図 6 のクラスター分析 (Ward 法) の結果は、容認度分布から E が (41) の三つのクラスターに分れることを意味している:¹⁰⁾

- (41) a. C1: {F1a, {G2c, G2b}}
- b. C2: {{{{F2a, F2b}, F1b}, {F2c, F1c}},
{G1c, G1b}, {{F2d, G1d}, F1d}}, G2d}}

⁹⁾2011/4/11 に追加。

¹⁰⁾クラスター C2 の下で全体が 4 分割になっているのが、評定課題が 4 段階だったことの副作用なのかどうかはわからない。それを確かめるには、同一の被験者に同一の課題を 3 段階か 5 段階評定を行なわせた結果と比較する必要がある。

Figure 6: 表 4 のデータに基づく E のクラスター化 (JMP 5.0.1J で作成)

c. C3: {G1a, G2a}

それぞれについては、C1 は容認可能と判断して問題の生じない表現、C3 は容認不可能と判断して問題の生じない表現、C2 は容認可能や容認不可能と判断すると誤認リスクが生じる表現だと解釈できるだろう。

C2 を更に $C2a = \{\{F2a, F2b\}, F1b\}, \{F2c, F1c\}\}$, $C2b = \{\{G1c, G1b\}, \{F2d, G1d\}, F1d\}, G2d\}$ の二つに分けるとすると、C2a は一定の誤認リスクで容認可能と判断してよいもの、C2b は一定の誤認リスクで容認不可能と判断してよいものとなる。

この結果は先に筆者が述べた結果と部分的にしか合っていない。合っている部分は、C3 の分離である。合っていない部分は C1 と C2 の区別である。後知恵的に言うと、C1 が別格であるという直観はあったが正当化が難しかったので、筆者の分類結果は C1 と C2a を混合するような中途半端な分類になっている。C2b がうまく分離できていないのはその副作用である。

参考までに外れ値に該当する可能性がある r16 を外して行なったクラスター分析の結果を図 7 に示す。C1, C2, C3 の区別に特に違いは生じていない。

以上のことが意味していることは次である：

(42) 熟練した言語学者の直観に基づく分類はデータの統計処理で支持される保証がない。

では、このようなズレが生じた時にどうするべきか？ 率直でない人の答えには「高度に政治的」な配慮が混入するだろうから、答えは一通りではない。だが筆者の答えは簡単である：十分な量の十分に信頼できるデータがあり、十分に妥当な解析法が使われているならば、人の直観は統計解析に勝てない。自戒を込めて言うと、これは Ayres (2008) が主張している論点と一致する。

4 終わりに

この試論で筆者は容認度評定の定式化を与えた。それは表現の全体集合 $E = (e_1, e_2, \dots, e_n)$ 、文脈の全体集合 $C = (c_1, c_2, \dots, c_m)$ 、評定者の全体集合 $R = (r_1, r_2, \dots, r_N)$ の三つの次元から構成される空間 $E \times C \times R$ と容認度分布 $a_{i,j,k} = a(e_i, c_j, r_k)$ を対応づける方法である。得られた定式化は特殊な場合として文法性判断を含むくらい一般的なものである。

提案された定式化の応用例として、8 つの表現: (19a) = “今日はいいい天気ですね”, (19b) = “今日はいいいにくの天気ですね”, (20a) = “昨日はいいい天気でしたね”, (20b) = “昨日はいいいにくの天気でしたね”, (21a) = “今日もいいい天気ですね”, (21b) = “今日もいいいにくの天気ですね”, (22a) = ?? “昨日はいいい天気ですね”, (22b) = ?? “昨日はいいいにくの天気ですね” と 4 つの文脈: (17a) = 発話前日は雨天で発話当日は雨天, (17b) = 発話前日は雨天で発話当日は晴天, (18a) = 発話前日は晴天で発話当日は雨天, (18b) = 発話前日は晴天で発話当日は晴天との相互作用を検討し、調節ゲームの一種として理解できることを見た。

最期に、 $a_{i,j}$ に評定者の変異を表わす次元を追加する拡張を論じた。その応用例として、F1a = “太郎が校庭を走る”, F1b = “稲妻が北の空を走る”, F1c = “戦慄が永田町を走る”, F1d = “汗が太郎の額を走る”, F2a = “校庭を

Figure 7: 表 4 の r16 を除いたデータに基づく E のクラスター化

太郎が走る”, F2b = “北の空を稲妻が走る”, F2c = “永田町を戦慄が走る”, F2d = “太郎の額を汗が走る”, G1a = “太郎が校庭に走る”, G1b = “稲妻が北の空に走る”, G1c = “戦慄が永田町に走る”, G1d = “汗が太郎の額に走る”, G2a = “校庭に太郎が走る”, G2b = “北の空に稲妻が走る”, G2c = “永田町に戦慄が走る”, G2d = “太郎の額に汗が走る”の12文の容認度を16人の票停車が(3点:まったく違和感を感じない; 2点:軽く違和感を感じるが, 言おうとしていることは簡単にわかる; 1点:強く違和感を感じるが, 言いたいことがまったくわからないわけではない; 0点:何を言っているかわからないか, 明らかに異常なことを言っていると思う)の4規準で評定した結果の値の分布を分析し, 得られた結果から表現を容認可能と容認不能に二者択一的に分類することは現実な要求でないことが通例である可能性を指摘した. 更に言語学者の直観による分類がデータの統計処理で支持される保証がないことも指摘した.

References

- 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.]
- Bacharach, M. (1976). *Economics and the Theory of Games*. London: Macmillan. [翻訳: 『経済学のためのゲーム理論』(訳: 鈴木 光男 and 是枝 正啓). 東洋経済新報社, 1981.]
- Blackwell, D. and M. A. Girshick (1954). *Theory of Games and Statistical Decisions*. John Wiley & Sons.
- Blutner, R., H. D. Hoop, and P. Hendriks (2006). *Optimal Communication*. Stanford, CA: CSLI.
- Blutner, R. and H. Zeevat (Eds.) (2004). *Optimality Theory and Pragmatics*. Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.
- Cubitt, R. and R. Sugden (2003). Common knowledge, salience and convention: A reconstruction of David Lewis’s game theory. *Economics and Philosophy* 19, 175–210.
- Kuroda, K., K. Nakamoto, and H. Isahara (2007). When nonce words behave like “real” words: A case study of the Japanese verb *oso(warer)u*. In P. Bouillon and K. Kanzaki (Eds.), *Proceedings of the 4th International Conference on Generative Approaches to the Lexicon, Paris, 2007*.
- Lewis, D. K. (1969). *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.
- Maynard-Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge University Press. [翻訳: J. メイナード-スミス. 進化とゲーム理論: 闘争の論理 (寺本英・梯 正之 (訳)). 産業図書, 1985.]
- Nakamoto, K. and K. Kuroda (2008). Representing selectional restrictions in terms of semantic frames equated with situational schemas: A case study of the Japanese verb *osou*. *Studies in Language Science* 7, 265–282.
- Sugden, R. (2004). *The Economics of Rights, Co-operation and Welfare* (2nd ed.). Palgrave Macmillan. [First edition was published in 1986.]
- von Neumann, J. and O. Morgenstern (1972). *Theory of Games and Economic Behavior*. Princeton University Press. [日本語訳: 『ゲームの理論と経済行動 I, II, III』(ちくま学術文庫)].
- 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].
- 中本 敬子, 黒田 航, and 野澤 元 (2005). 素性を利用した文の意味の心内表現の探索法. *認知心理学研究* 3 (1), 65–81.
- 黒田 航 (2009). (容認度の低い例の) 作例を効果的にこなうためのツールの紹介と方法論の提示. In 日本認知言語学会発表論文集第 9 巻, pp. 458–464. 原典版: <http://clsl.hi.h.kyoto-u.ac.jp/~kkuroda/papers/paper-on-exgen.pdf>.
- 黒田 航 (2011). 自作例を使った研究の基礎. In 中本 敬子 and 李 在鎬 (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.
- 黒田 航 and 寺崎 知之 (2010). 言語の「自然態」を捉える言語理論の必要性. In *言語処理学会第 16 回年次大会発表論文集*.