

反 応 語 の 系 列 効 果

— 評定尺度の研究 —

織 田 揮 準*

問 題

評定尺度における反応語（カテゴリー用語とも評語とも呼ばれる）の主な目的は、カテゴリー間の順序（上位下位）関係を言語的に明確化することである。従って、評定尺度を構成する反応語群は、研究者は無論のこと被験者群においても十分に意味化していなければならない。さらに、反応語群の配列順序（以下、反応語系列）に関しても、研究者と被験者群の理解に一致がなければならないとされている（田中，1973；柿崎，1974；織田，1970，1976b）。

織田は、評定尺度の収集と分析（1967，1975b），および、反応語の収集・分類を行い、さらに、一対比較法による反応語の意味強度（尺度値）を測定した（1967，1970，1976b）。収集された評定尺度の反応語系列と一対比較法で求めた反応語の意味強度の順序との間には、強い対応関係があり、従来の評定尺度の妥当性が確認された。しかし、1つの例外として、反応語「どちらともいえない」と「あまり……」との上位下位関係に関して不一致がみられた。従来の評定尺度では、「どちらともいえない」を「あまり……」の上位カテゴリーにあてはめるのが一般的な配列である（FIG. 1の尺度事例 5B₃₁を参照のこと）。これに対して、一対比較法による調査結果（織田，1967，1970，1976b）によると、両語の意味強度に有意差はない（ χ^2 -検定， $p \geq 0.05$ ）。たとえば、被験者を大学生（327名）とした一対比較の結果をみると、「どちらともいえない」が「あまり大きくない」よりも、上位（大きい）の関係（以下、「どちらともいえない」>「あまり大きくない」）にあると応答したものは55%であり、両語の意味強度には有意差がない（ $\chi^2=3.12$ ， $p > 0.05$ ）。

「どちらともいえない」と「あまり……」のように、意味強度の類似した反応語群が、同一の評定尺度（順序尺

度）に組み込まれると、尺度の反応語系列と被験者の理解する順序関係との間に不一致が生じ、判断過程に重大な影響が生じるであろう。そこで、本研究においては、意味強度の類似した反応語群が、同一の尺度に組み込まれたときに生ずる判断過程への影響を検討する。

実 験 I

目的

意味強度の類似した反応語「どちらともいえない」と「あまり……」の配列順序の違いによる判断過程への影響を検討する。「どちらともいえない（以下、記号D）」を「あまり……（以下、記号A）」の上位カテゴリーにあてはめる従来の配列（以下、DA配列）に対して、「あまり……」を上位カテゴリーとする反応語系列をAD配列とする。

織田（1976a）は、41種類の5段階評定尺度（順序尺度）を用いて、カテゴリー幅の等間隔性仮定の検討を試みた。その結果、カテゴリー幅の近似的等間隔性が確認された。ところで、41種類の評定尺度には、「どちらともいえない」と「あまり……」を含む尺度が18種類あり、いずれもDA配列型の尺度であった。DA配列型の尺度（順序尺度）が、近似的にはあるが、等間隔尺度の1種とみなして使用されたという結果から、次の推論が演繹される。すなわち、意味強度の類似した、換言すれば、意味の未分化な反応語群（操作的には、一対比較判断に偏りのない反応語群）が、同一尺度に組み込まれた場合、被験者は被験者自身の言語的な理解に基づく順序よりも、与えられた尺度の反応語系列を優先させるであろう。従って、「あまり……」を「どちらともいえない」の上位カテゴリーとするAD配列型の尺度による判断では、「あまり……」で判断される刺激値は「どちらともいえない」で判断される刺激値よりも大となるはずである。そこで、実験Iでは、次の仮説を検証する。すなわち、意味強度の類似した反応語群が、同一の評定尺度に組み込まれたとき、被験者は被験者自身の言語的理

* 三重大学

解に従うというよりも、尺度の反応語系列に従って判断する。

方法

実験方法は、織田 (1975 a) と同じである。そこで、詳細は先の論文にゆずり、ここでは簡単な説明にとどめる。

1. 評定尺度の構成と分類

程度量表現用語の調査をもとに、15語を選び、それらの組み合わせによって、中性カテゴリー (どちらともいえない) を含む41種類の5段階評定尺度が構成された (織田, 1975 a)。41種類の尺度のなかには、「どちらともいえない」と「あまり……」を同時に含む18種類の尺度がある。本実験では、これら18種類の尺度のうち、「長い—長くない」次元の9種類の尺度を用いることにする。これらの尺度は、いずれもDA配列型尺度である。そこで、「どちらともいえない」と「あまり長くない」

の配列順序を入れ替えてできる9種類のAD配列型尺度を用いる (TABLE 1)。

2. 刺激図形

水平線を変化刺激、垂直線を標準刺激とする水平垂直線図形 (以下、相対判断図形) と水平線 (変化刺激) のみからなる絶対判断図形を用いる (織田, 1975 a, FIG. 2)。変化刺激は、1単位から14単位まで1単位きざみで変化する14種類、標準刺激は7単位の長さである。スライド・プロジェクターで投影された刺激図形の1単位あたりの長さは、約7—9 cm である。なお、各被験者には、1つの尺度につき、14 (刺激図形数) × 2 (くり返し回数) 回の判断を求めた。

3. 被験者

各尺度とも5名の実験結果が、分析にかけられた。従って、AD配列型尺度の数が9、判断事態が2 (相対判断と絶対判断) であるから、被験者総数は90名である。

4. 手続

9種類のAD配列型尺度ごとに、5つの反応語が反応語系列順に印刷された6枚つづりのブックレットが作られた (織田, 1975 a, FIG. 3)。被験者には、9種類のブックレットの一冊がランダムに配布され、次の教示が与えられた。

“いまから水平線 (または、水平線と垂直線) の描かれている28枚のスライドをお見せいたします。水平線の長さがいろいろ変化します。(垂直線の長さはいつも同じです)28枚のスライドとは次のようなものです。(28枚のスライドを1枚当たり約3秒間の呈示速度で一通り見せる) では、次にもう一度、今のスライドを順番に出しますから、お手もとのブックレットに印刷されている5個のコトバでもって、水平線の長さを判断して下さい。水平線の長さにびたりあてはまるコトバがないかもしれませんが、そのときは最も近いと思われるコトバにチェックして下さい。では1枚目です”。

テスト試行での刺激呈示時間は、1枚あたり10—13秒で、予め決められたランダムな順序で呈示された。1回の実験では、尺度の種類をかえて最大2試行まで行った。ただし、同一被験者が、相対判断と絶対判断の両判断事態で判断することはなかった。10名から30名の集団実験である。

結果

DA配列型尺度の実験 (織田, 1975 a) と同様に、各AD配列型尺度とも信頼性の高い5名分の資料の得られたところで、実験を終了した。すなわち、水平線分の長さを異にする14種類の変化刺激に対するそれぞれ2回の判断間の不一致数が、14分の4以下の資料を分析の対象

TABLE 1 DA, AD 配列型尺度とその構成*

尺度の配列型	尺度番号	Category Words						
		全然長くない	長くない	あまり長くない (A)	どちらともいえない (D)	どちえらばか長い	長い	非常に長い
DA配列型尺度	5B ₇			①	②	③	④	⑤
	5B ₂₉	①		②	③	④	⑤	
	5B ₃₀	①		②	③	④		⑤
	5B ₃₁	①		②	③		④	⑤
	5B ₃₂		①	②	③	④	⑤	
	5B ₃₃		①	②	③	④		⑤
	5B ₃₄		①	②	③		④	⑤
	5B ₃₉ 5B ₄₀	① ② ③	④	⑤				⑤
AD配列型尺度	5B ₇			②	①	③	④	⑤
	5B ₂₉	①		③	②	④	⑤	
	5B ₃₀	①		③	②	④		⑤
	5B ₃₁	①		③	②		④	⑤
	5B ₃₂		①	③	②	④	⑤	
	5B ₃₃		①	③	②	④		⑤
	5B ₃₄		①	③	②		④	⑤
	5B ₃₉ 5B ₄₀	① ② ④	③	⑤				⑤

* ○つき数字はカテゴリー番号を示す。
 ** DA配列型尺度の構成については、織田 (1975 a, TABLE 1) を参照のこと。なお、Dは「どちらともいえない」、Aは「あまり長くない」を表わす。

TABLE 2 「どちらともいえない (D)」と「あまり長くない (A)」の平均刺激値 $m_i(D)$, $m_j(A)$ の差 (d_{ij} , d_{ji}) の検定 (t-検定)†

判断事 態	DA 配列型尺度 (尺度数 9††)				AD 配列型尺度 (尺度数 9††)			
	平均刺激値の差 d_{ij}	有意差 ($p \leq .01$) のある尺度数		有意差のない 尺度数 $d_{ji} \neq 0$	平均刺激値の差 d_{ji}	有意差 ($p \leq .05$) のある尺度数		有意差のない 尺度数 $d_{ji} \neq 0$
		$d_{ij} > 0$	$d_{ij} < 0$			$d_{ji} > 0$	$d_{ji} < 0$	
相 対 判 断	M=2.3 S ² =0.21 R=1.7	9	0	0	M=0.1 S ² =2.99 R=6.3	2	2	5
絶 対 判 断	M=2.9 S ² =0.25 R=1.7	9	0	0	M=1.6 S ² =0.91 R=2.4	4	0	5

† i は「どちらともいえない」、 j は「あまり長くない」のあてはめられたカテゴリー番号を示す。平均刺激値の差 $d_{ij} = m_i(D) - m_j(A)$ ($i = j + 1$), $d_{ji} = m_j(A) - m_i(D)$ ($j = i + 1$), M は差の平均, S^2 は差の分散, R は差の範囲を表わす。

†† 1 尺度あたりの被験者数は 5 名である。

とした (以下, 75% 基準をみたま資料, 織田, 1975 a, p.305)。

次に, DA 配列型尺度 (織田, 1975 b, 附表 1, 2) と本実験の AD 配列型尺度による判断結果を比較検討する (TABLE 2, 3)。

1. 「どちらともいえない」と「あまり長くない」を生起する平均刺激値におよぼす配列効果の検討

「どちらともいえない (D)」(カテゴリー番号 i) を生起した平均刺激値を $m_i(D)$, 「あまり長くない (A)」(カテゴリー番号 j) を生起した平均刺激値を $m_j(A)$ であらわす。ここで, DA 配列型尺度群では, $i = j + 1$, AD 配列型尺度群では $j = i + 1$ の関係がある。

まず, DA 配列型尺度群では, すべての尺度において, $m_i(D) > m_j(A)$ の関係が認められる (t-検定, $p \leq .01$)。次に, AD 配列型尺度群についてみると, 相対判断では $m_j(A) > m_i(D)$ と仮説で期待される関係を支持するもの 2 に対して, 仮説を否定するもの ($m_j(A) < m_i(D)$) が 2, 残りの 5 種類の尺度には有意差がない (t-検定, $p \leq .05$)。また, 絶対判断では, 平均刺激値が $m_j(A) > m_i(D)$ と仮説を支持するもの 4, 残りの 5 種類の尺度には有意差がない (t-検定, $p \leq .05$)。

2. 「どちらともいえない」と「あまり長くない」を生起する平均刺激値の分散におよぼす配列効果の検討

相対判断では, 「どちらともいえない」と「あまり長くない」を生起する刺激値の分散におよぼす配列効果は, ほとんど認められない (F-検定, $p \leq .05$, TABLE 3)。しかし, 絶対判断では, AD 配列型尺度群における反応語「あまり長くない」を生起する刺激値の分散が, 増大する傾向がある。従って, 反応語「どちらともい

TABLE 3 反応語を生起する刺激値の分散におよぼす配列効果の検定 (F-検定)

判断事 態	反 応 語	F-検定の結果		
		有意差 ($p \leq .05$) のある尺度数		有意差 ($p \leq .05$) のない尺 度数
		DA 配列 の分散大	AD 配列 の分散大	
相 対 判 断	どちらともいえない	1	1	7
	あまり長くない	0	2	7
絶 対 判 断	どちらともいえない	1	3	5
	あまり長くない	0	5	4

ない」の使用に際して, 混乱が生じているといえよう。

考察

本実験では, 「どちらともいえない」と「あまり長くない」を生起する平均刺激値におよぼす配列効果を検討した。分析の結果, 「どちらともいえない」を上位カテゴリーとする DA 配列型尺度群で認められた反応傾向 (仮説) が, 「あまり長くない」を上位カテゴリーとする AD 配列型尺度群では認められなかった。また, 「あまり長くない」を生起する刺激値の分散は, DA 配列型尺度群に比べて, AD 配列型尺度群で増大する。これらの結果は, 「どちらともいえない」と「あまり長くない」の上下位関係に関して, 被験者の言語的理解と AD 配列型尺度の反応語系列との間に不一致が生じ, 判断過程に混乱が生じたことを示す。従って, AD 配列型尺度群に関する限り, 仮説は否定された。

AD 配列型尺度群で仮説が否定された原因として, 次の可能性が考えられる。第 1 は, 「どちらともいえない」と「あまり長くない」の意味強度に関する被験者の言語的理解に関連するものである。一対比較調査によれば,

両語の意味強度は類似し、意味分化は低い。しかし、この結果はあくまでも被験者集団の平均的な反応傾向であり、両語の意味強度が個々の被験者においても未分化であることを示すものでは必ずしもない。たまたま、DA配列型尺度群のすべての被験者が、「どちらともいえない」を「あまり長くない」よりも意味強度の強いことばであると判断し、AD配列型尺度群の被験者群は、「どちらともいえない」をより意味強度の強いことばと判断するものとその逆の判断をするものの混在する集団であったとも考えられる。従って、仮説は、両語の意味強度に関する被験者の言語的理解との関連において再検討されなければならない。

第2は、表現形式からみた反応語系列に関するものである。実験Iで使用した尺度は、刺激図形の線分の長さを、「長い—長くない」次元で判断することを目的にして作られた。その場合、意味強度の類似した「あまり長くない」と「どちらともいえない」のいずれを上位カテゴリーにあてはめるのが、表現形式からみて妥当な反応語系列かという問題がある。DA配列型の尺度事例5B₃₁ (FIG. 1) を例にとれば、「どちらともいえない」よりも上位(より長い)のカテゴリーは、すべて肯定的表現用語で統一され、「どちらともいえない」よりも下位カテゴリーは、すべて否定的表現である。一方、AD配列型の尺度事例5B'₃₁は、「どちらともいえない」の上位カテゴリーに肯定的表現と否定的表現の反応語が混在し、表現上の一貫性に欠ける。

われわれは、一連の刺激系列を「長い—長くない」次元で判断するとき、肯定的判断から否定的判断、あるいは、その逆の構えでもって判断する。DA配列型尺度群では、「肯定的表現から否定的表現」へという表現形式上の一貫性が確保されており、被験者の判断の構えと矛盾しない。従って、意味強度の未分化な反応語「どちらともいえない」と「あまり長くない」が同一の尺度に用いられたとしても、DA配列型尺度では、表現形式の一貫性が確保されているために、反応語系列に関する矛盾を被験者に感じさせることが少ない。ところが、AD

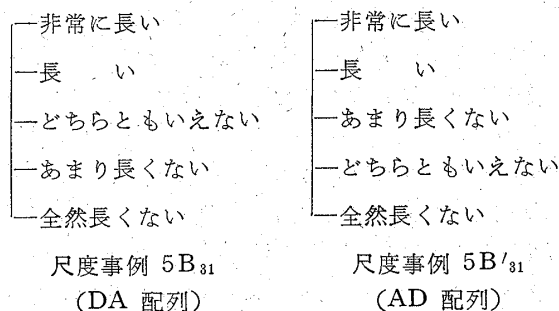


FIG. 1 DA, AD 配列型尺度構成例

配列型尺度では、①意味強度の未分化に加えて、②表現形式からみた反応語系列の一貫性の欠如によって、被験者に尺度の反応語系列に関する疑問を抱かせ、判断過程を混乱させることになる。

実験Iは、評定尺度構成上の条件として、従来からいわれている「意味強度の分化した反応語群を意味強度の順に系列化する」ことに加え、「反応語系列における表現形式の一貫性の確保」が、重要な条件であることを示唆するものである。

実験 II

目的

実験Iの目的は、反応語系列の妥当性を意味強度の面から検討することであり、表現形式における一貫性について配慮しなかった。そこで、実験IIでは、表現形式を一定にして、実験Iの仮説を再検討する。

表現形式が等しく、意味強度の類似した反応語として、「やや長い」と「どちらかといえば長い」を用いる。これらの反応語を用いる理由は、①両語とも評定尺度の反応語として使用頻度の高いことばであること、②一対比較調査の結果、意味強度が類似していることと(「やや長い」>「どちらかといえば長い」と応答したもの75%、織田, 1976b, 表2)、③程度量表現用語(34語)の5段階、7段階分類調査で、両語とも同一の段階値が得られたこと(織田, 1976b, 表3)である。

なお、実験Iの仮説は、意味強度の類似した反応語群が、同一の尺度に組み込まれたときの配列効果に関するものであるが、実験IIでは、意味分化が明確な反応語群の配列効果をもあわせて検討する。具体的には、意味分化の明確な反応語として、①「やや長い」と「かなり長い」(「かなり長い」>「やや長い」とするもの99%)、②「長い」と「非常に長い」(「非常に長い」>「長い」とするもの99%)の配列効果を検討する(織田, 1976b, 表2)。

方法

実験方法は、評定尺度の構成、判断実験後の一対比較調査の実施などの一部を除き、実験Iと同じである。

1. 評定尺度の構成

意味強度の類似した「やや長い(Y)」と「どちらかといえば長い(D)」、意味強度の大きく異なる「かなり長い(K)」と「やや長い(Y)」、および、「非常に長い(短い)(H)」と「長い(短い)(N)」の配列をかえて、全部で6種類の尺度が作られた(TABLE 4, 5)。なお、尺度ごとに作られたブックレットの印刷形式は、反応語系列を強調するために FIG. 2 のごとく修正した。

2. 刺激図形

6種類の尺度のなかには、中性判断（どちらともいえない）を含まない尺度があるために、実験Iの水平線図形（絶対判断図形）を刺激図形として用いる。

3. 被験者

被験者は、大学生135名（実験A：81名、実験B：54名）である。

4. 実験手続

刺激図形の判断実験（集団実験）の終了後、「やや長い—どちらかといえば長い」、「かなり長い—やや長い」、および、「非常に長い—長い」の刺激語対を含む一対比較調査が実施された。

結果と考察

実験IIは、AとBの実験に分かれる。実験Aでは、「やや長い」と「どちらかといえば長い」の配列効果、実験Bでは「かなり長い」と「やや長い」、および、「非常に

長い（短い）」と「長い（短い）」の配列効果を検討する。

1. 実験A（「やや長い」と「どちらかといえば長い」の配列効果）の結果と考察

1.1 一対比較調査の結果

81名の被験者のうち、「やや長い（Y）>どちらかといえば長い（D）」と一対比較判断したものは47名（58%）（以下、YD群）、逆の判断をしたものは34名（42%）（以下、DY群）であった。度数の検定の結果、両群間の被験者数に有意差はない（ $\chi^2=1.96, p>.10$ ）。従って、両語は意味強度の未分化なことばであることが確認された。

1.2 判断実験の結果

YD群（47名）とDY群（34名）にYD配列尺度とDY配列尺度による刺激図形の判断を求めた。なお、YD配列尺度とDY配列尺度で、各群の被験者数が異なるのは、信頼性の低い資料（75%基準をみたさない資料）を除いて集計したためである。TABLE 4とFIG. 3から、次の点が指摘できる。

(1) YD配列尺度とDY配列尺度とも、また、YD群とDY群に共通して、カテゴリー番号*i*を生起した平均刺激値を*m_i*であらわせば、 $m_i > m_{i-1}$ (z -検定, $p \leq .01$)の関係が認められる。これは、実験Iの仮説を支持する。

(2) 「やや長い」と「どちらかといえば長い」があてはめられたカテゴリー4と3の平均刺激値の差 ($d_{43} = m_4$

尺度	スライド番号			
	1	2	...	28
—長い			...	✓
—やや長い	✓		...	
—どちらかといえば長い			...	
—あまり長くない		✓	...	
—長くない			...	

FIG. 2 YD 配列尺度と応答形式例

TABLE 4 「やや長い（Y）」と「どちらかといえば長い（D）」の配列効果

尺度†	群†† (判断総数)	Y D 群 (47人×28回)				D Y 群 (34人×28回)					
		度数	平均	分散	差 <i>d</i> と <i>z</i> -検定†††	度数	平均	分散	差 <i>d</i> と <i>z</i> -検定†††		
YD 配列 尺度	—⑤長い	311	12.7	1.2	$d=3.1^{***}$	214	12.8	1.2	$d=3.2^{***}$		
	—④やや長い	267	9.6	2.0		$d=2.2^{***}$	180	9.6		1.7	$d=1.5^{***}$
	—③どちらかといえば長い	208	7.4	2.0		$d=2.7^{***}$	148	8.1		2.6	$d=3.2^{***}$
	—②あまり長くない	288	4.7	1.7		$d=2.8^{***}$	232	4.9		1.7	$d=3.0^{***}$
	—①長くない	242	1.9	0.8			178	1.9		1.0	
DY 配列 尺度	—⑤長い	317	12.6	2.0	$d=3.6^{***}$	218	12.9	1.0	$d=3.1^{***}$		
	—④どちらかといえば長い	202	9.0	2.6		$d=0.9^{***}$	154	9.8		2.0	$d=1.7^{***}$
	—③やや長い	234	8.1	2.6		$d=3.5^{***}$	195	8.1		2.0	$d=3.2^{***}$
	—②あまり長くない	285	4.6	1.4		$d=2.7^{***}$	212	4.9		1.4	$d=2.8^{***}$
	—①長くない	222	1.9	0.6			201	2.1		1.0	

† ○つき数字は、カテゴリー番号を示す。

†† YD群とは、「やや長い>どちらかといえば長い」、DY群とは、「やや長い<どちらかといえば長い」と一対比較判断したものである。

††† ***; $p \leq .001$

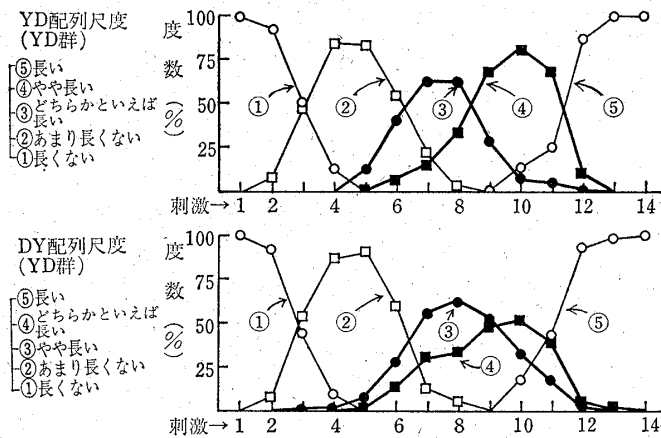


FIG. 3 YD, DY 配列尺度による判断結果 (度数分布)

- m_3) をみると、YD群のYD配列尺度で最大値 ($d_{43} = 2.2$) が得られるのに対して、DY配列尺度では $d_{43} = 0.9$ と小さい。一方、DY群では、 $d_{43} = 1.5$ と 1.7 と近似した値が得られた。被験者の理解する意味強度の順序と尺度の反応語系列とが一致するときの平均刺激値の差 d_{43} と一致しないときの差 d_{43} を比較すると、一致するときの差の値が大きい。この傾向は、YD群において顕著である。

以上の分析結果は、実験Iの仮説を支持するものである。しかし、FIG. 3 から明らかなように、意味強度が十分に分化した反応語があてはめられたカテゴリーの平均刺激値の差 ($d_{i,i-1}$, $i=2, 3, 5$) に比べて、意味強度の類似したカテゴリーの平均刺激値の差 d_{43} は小さい。この結果は、意味強度の類似した反応語群のあてはめられたカテゴリーは、相互に独立したカテゴリーとして十分に機能しないことを示す。従って、実験Iの仮説で述べられた尺度の使用傾向があるとはいえ、意味強度の未分化な反応語群を同一の尺度に組み込むべきではない。

2. 実験B (「かなり長い」と「やや長い」, および、「非常に長い(短い)」と「長い(短い)」の配列効果)の結果と考察

2.1 一対比較調査の結果

54名の被験者のすべてが、「かなり長い(K) > やや長い(Y)」, 「非常に長い(H) > 長い(N)」と一対比較判断をしている。従って、両語の意味強度は十分に分化している。

2.2 判断実験の結果

54名の被験者をランダムに2分し、KY群(27名)とHN群(27名)の2群を作る。KY群には、最初にYK配列尺度による判断を求め、次にKY配列尺度による判断を求めた(TABLE 5の尺度を参照)。また、HN群

TABLE 5 「かなり長い(K)」と「やや長い(Y)」および「非常に長い(短い)(H)」と「長い(短い)(N)」の配列効果

尺 度 †	度数	平均	分散	差dとZ-検定 ††
KY 配列 尺度	⑤非常に長い	139	13.0	1.0)d=2.5 ***
	④かなり長い	151	10.5	1.7)d=2.9 ***
	③やや長い	164	7.6	1.7)d=3.0 ***
	②あまり長くない	170	4.6	1.4)d=2.4 ***
	①全然長くない	132	2.2	0.6
判断総数		27人×28回		
YK 配列 尺度	⑤非常に長い	142	12.9	1.0)d=4.5 ***
	④やや長い	159	8.4	2.6)d=-1.6 ***
	③かなり長い	118	10.0	3.2)d=5.2 ***
	②あまり長くない	177	4.8	1.7)d=3.6 ***
	①全然長くない	132	1.9	0.9)d=2.9 ***
判断総数		26人×28回 †††		
HN 配列 尺度	⑤非常に長い	167	12.8	1.2)d=3.2 ***
	④長い	191	9.6	2.0)d=3.3 ***
	③どちらともいえない	171	6.3	2.0)d=2.7 ***
	②短い	131	3.6	1.0)d*2.1 ***
	①非常に短い	96	1.5	0.4
判断総数		27人×28回		
NH 配列 尺度	⑤長い	165	10.0	2.3)d=-2.6 ***
	④非常に長い	159	12.6	1.7)d=6.0 ***
	③どちらともいえない	170	6.6	2.0)d=5.0 ***
	②非常に短い	94	1.6	0.6)d=-2.1 ***
	①短い	140	3.7	1.4
判断総数		26人×28回 †††		

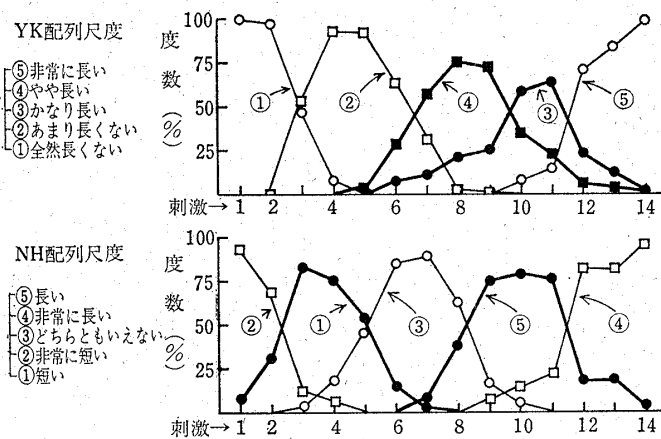
† ○つき数字はカテゴリー番号を示す。

†† ***; $p \leq .001$

††† 被験者数が26名と1名減少しているのは、75%基準をみたさない資料が1名分あり、これを除いたためである。

は、最初にNH配列尺度による判断を、次にHN配列尺度による判断を求めた。TABLE 5 と FIG. 4 から、次の点が指摘できる。

(1) 被験者の言語的理解と尺度の反応語系列とが一致するKY配列尺度では、隣り合ったカテゴリーの平均刺激値の差 $d_{i,i-1}$ は $3.0-2.4$ で、その分散は小さい。一方、被験者の言語的理解と尺度の反応語系列とが一致しないYK配列尺度では、差 $d_{43} = -1.6$ となり、カテゴリー番号と平均刺激値の大きさの順序が逆転している。ところが、カテゴリー3と4の順序を入れ替えてKY配列尺度と同じ配列にすると、平均刺激値の差 ($d_{i,i-1}$) は $3.6-1.6$ となり、その分散は減少する。しかし、FIG. 4 から明らかなように、YK配列尺度のカテゴリー3と4の度数分布の重複度が高い。従って、カテゴリー3と4が、相互に独立したカテゴリーとして十分に機能してい



るとは認めがたい。

(2) 被験者の言語的理解と尺度の反応語系列とが一致するHN配列尺度では、平均刺激値の差 ($d_{i \cdot i-1}$) は 3.3—2.1 である。一方、一致しないNH配列尺度では、6.0—2.6 となり、分散が大きい。しかし、カテゴリー 4 と 5、および、1 と 2 の順序を入れ替えると、差 ($d_{i \cdot i-1}$) は 3.4—2.1 となり、分散が減少する。

実験Bの結果は、意味強度が十分に分化した反応語群で構成されている尺度の使用に際して、尺度の反応語系列と被験者の言語的理解との間に不一致が生ずれば、被験者は被験者自身の理解する意味強度の順序に反応語群を再系列化して使用することを示唆するものである。

3. 実験IIの結論

実験IIの結果は、実験Iの仮説を支持するものである。すなわち、意味強度の類似した反応語群が同一の尺度に用いられ、その反応語系列が被験者の理解する順序と矛盾する場合、被験者は尺度の反応語系列に自分の言語的理解を合わせて判断する傾向がある。しかし、意味強度の類似した反応語群のあてはめられたカテゴリーの平均刺激値の差 ($d_{i \cdot i-1}$) は、意味強度の十分に分化した反応語群のあてはめられたときの差 ($d_{i \cdot i-1}$) に比べて小さい。従って、意味強度の類似した反応語群で構成された尺度では、判断過程に混乱が生じ易い。

一方、意味強度が十分に分化した反応語群の使用される場合は、尺度の反応語系列と被験者の言語的理解が一致しないとき、被験者は自分の言語的理解に合わせて反応語群を再系列化して判断する。なお、意味強度が十分に分化した反応語群であれば、尺度の反応語系列と被験者の理解する順序の間に不一致が生じて、それが原因で判断過程に混乱が生ずることは比較的少ない。

以上の結果から、実験Iの仮説は次のように修正される。すなわち、表現形式の一貫性が確保されているなら

ば、①被験者にとって意味強度の未分化な反応語群が含まれた尺度では、被験者は尺度の反応語系列に従って判断するが、②意味強度の十分に分化した反応語群で構成されているが、その反応語系列と被験者の理解する順序とが一致しないときは、被験者は被験者自身の理解する反応語系列に従って判断する。

実験 III

目的

実験IIで、次の点が明らかにされた。すなわち、意味強度の類似した、従って意味強度の未分化な反応語群が同一の尺度に組み込まれた場合、たとえその反応語系列が被験者の理解する順序と異なるとしても、表現形式の一貫性が確保されている反応語系列ならば、被験者は尺度の反応語系列に従って判断する。本研究では、尺度の反応語系列に表現形式の一貫性が欠如する場合の応答傾向をさぐる。なお、表現形式は異なるが、意味強度の未分化な反応語として、実験Iで用いた「どちらともいえない」と「あまり長くない」を用いる。

方法

実験方法は、実験IIとほぼ同じである。

1. 評定尺度の構成

「どちらともいえない(D)」と「あまり長くない(A)」の配列順序をかえて、DA配列尺度とAD配列尺度の2種類の尺度が作られた (TABLE 6)。

2. 刺激図形

刺激図形として、実験Iで用いられた水平垂直線図形(相対判断図形)が用いられた。

3. 被験者

被験者は、79名の大学生である。

4. 実験手続

刺激図形の判断実験の終了後、「どちらともいえない—あまり長くない」と「あまり長くない—どちらともいえない」の比較対を含む一対比較調査が行われた。

結果と考察

1. 一対比較調査の結果

「どちらともいえない(D)」と「あまり長くない(A)」の一対比較判断を2回求めたところ、一貫して「どちらともいえない>あまり長くない」と応答したものは79名中、43名 (54%) (以下、DA群)、一貫して逆の反応をしたもの12名 (15%) (以下、AD群)、応答に一貫性の欠けるもの24名 (31%) であった。次に、DA群(43名)とAD群(12名)の実験結果を分析する。

2. 判断実験の結果

実験結果が、TABLE 6 と FIG. 5 に示されている。

TABLE 6 「どちらともいえない (D)」と「あまり長くない (A)」の配列効果

尺度†	群††	D A 群				A D 群			
		度数	平均	分散	差 d と Z-検定†††	度数	平均	分散	差 d と Z-検定†††
D A 配列尺度	⑤非常に長い	249	12.8	1.2) d=2.7 ***) d=2.9 ***) d=2.5 ***) d=2.5 ***	58	13.1	1.1) d=2.7 ***) d=2.9 ***) d=2.5 ***) d=2.8 ***
	④長い	277	10.1	2.0		80	10.4	1.6	
	③どちらともいえない	212	7.2	1.2		59	7.5	1.8	
	②あまり長くない	203	4.2	1.2		63	5.0	1.9	
	①全然長くない	263	2.2	1.2		76	2.2	1.2	
判断総数		43人×28回				12人×28回			
A D 配列尺度	⑤非常に長い	202	13.0	1.1)d=2.7 ***)d=4.1 ***)d=-.5 ***)d=4.1 ***	60	13.0	1.1) d=2.5 ***) d=3.5 ***) d=0.4) d=4.0 ***
	④長い	267	10.3	2.0		70	10.5	1.8	
	③あまり長くない	149	6.2	5.7		50	7.0	6.3	
	②どちらともいえない	201	6.7	1.2		64	6.6	0.9	
	①全然長くない	301	2.6	1.8		92	2.6	1.6	
判断総数		40人×28回				12人×28回			

† ○つき数字はカテゴリー番号を示す。

†† DA 群とは、一対比較調査で「どちらともいえない>あまり長くない」、AD 群は「どちらともいえない<あまり長くない」と判断したものである。

††† *: $p \leq .05$, ***: $\leq p .001$

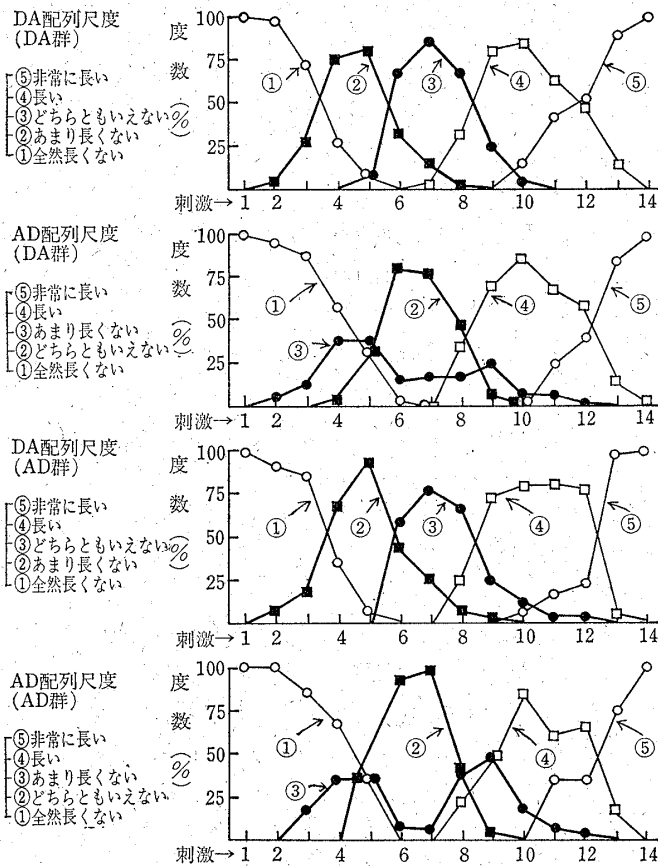


FIG. 5 DA, AD 配列尺度による DA, AD 群別の判断結果

なお、DA 群の DA 配列尺度と AD 配列尺度で被験者数が異なるのは、信頼性の低い (75% 基準をみたさない) 資料を除外したためである。

2.1 DA 群における「どちらともいえない」と「あまり長くない」の配列効果

DA 群の理解と一致する反応語系列をもつ DA 配列尺度の平均刺激値の差 ($d_{i,i-1} = m_i - m_{i-1}$) は 2.9-2.5 であり、近似的等間隔性が認められる。一方、DA 群の理解と矛盾する AD 配列尺度では、平均刺激値の差 ($d_{i,i-1}$) が 4.1-1.5 となり、不等間隔性が認められる。とくに、差 d_{32} は -1.5 で、カテゴリー番号と平均刺激値の大きさとの関係が逆転している。

2.2 AD 群における「どちらともいえない」と「あまり長くない」の配列効果

AD 群の理解と一致する反応系列をもつ尺度は、AD 配列尺度である。ところが、AD 配列尺度の平均刺激値の差 ($d_{i,i-1}$) は、0.4-4.0 であり、カテゴリー幅の不等間隔性が目立つ。とくに、 $d_{34} = 0.4$ であり、差の検定の結果有意差はない (α -検定, $p \leq .10$)。逆に、AD 群の理解とは異なる反応語系列をもつ DA 配列尺度では、平均刺激値の差 ($d_{i,i-1}$) は 2.5-2.9 となり、カテゴリー幅の等間隔性が強い。

次に、FIG. 5 から次の点が指摘できる。

2.3 DA 配列尺度のカテゴリー別度数分布をみると、DA 群、AD 群ともに単峰的度数分布が認められる。

2.4 AD 配列尺度の反応語「あまり長くない」のあて

はめられたカテゴリー3の度数分布は、DA群、AD群ともに2峰の度数分布である。

3. 実験Ⅲの結論

一般に、意味強度の十分に分化した、換言すれば、上位下位関係のはっきりした反応語群で構成された反応語系列をもつ尺度を用いて、一連の刺激系列を判断したとき、隣り合ったカテゴリーの平均刺激値の差は、ほぼ等しい。いわゆるカテゴリー幅の等間隔性が認められる。また、このような尺度においては、各カテゴリーごとの度数分布は、単峰的分布をする。これに対して、意味強度の未分化な反応語群を含む反応語系列の尺度による判断では、判断過程に混乱が生じ、カテゴリー幅の等間隔性がくずれる。すなわち、隣り合ったカテゴリーの度数分布曲線は重複 (FIG. 3) したり、あるいは、単峰的分布がくずれて複峰的分布をする。

DA配列尺度では、カテゴリー幅の等間隔性 ($d_{i+1,i} = d_{i,i-1}$) と単峰的度数分布が得られた。一方、AD配列尺度では、反応語「あまり長くない」のあてはめられたカテゴリーの度数分布は、2峰の分布であった。

実験Ⅲの以上の結果は、評定尺度の構成にあたり、次の点の重要性を示すものである。すなわち、①尺度に使用される反応語群の意味強度は十分に分化し、従って、上位下位関係が明確であること (意味強度的妥当性)、②尺度の反応語系列は意味強度の順序に従って系列化されること (系列的妥当性)、および③反応語系列には表現形式上の一貫性が確保されること (表現的妥当性)、の重要性を示すものである。とくに、意味強度の類似した反応語群の含まれる反応語系列では、表現的妥当性が重要な要件である。

結 論

評定尺度に用いる反応語群は、意味強度が十分に分化していること、および、意味強度の順序に従って系列化されることが必要であるとされてきた。そして、このような条件を満たす評定尺度 (厳密には順序尺度) のカテゴリー幅は、近似的に等間隔であり、その度数分布は単峰的分布をすることが、経験的にいわれ、また、実験的にも明らかにされた (織田, 1976 a)。

織田 (1967, 1970, 1976 b) は、心理学研究で使用された評定尺度の収集と分析を行うとともに、程度量表現用語の対比較調査を実施した。その結果、収集された評定尺度の反応語系列は、対比較調査で求められた意味強度の順序とよく一致することが判明した。ところが、1つの例外として、比較的使用頻度の高い評定尺度、具体的には、反応語「どちらともいえない」と「あ

まり……」を含む尺度の反応語系列に矛盾がみられた。従来の評定尺度では、「どちらともいえない」が「あまり長くない」の上位カテゴリー用語として使用されてきた。ところが、程度量表現用語の対比較調査では、両語の意味強度は近似的に等しいことが判明した。従って、意味強度の未分化な反応語「どちらともいえない」と「あまり……」を含む評定尺度は、従来の経験から判断すれば、不適切な尺度である。ところが、「どちらともいえない」を「あまり……」の上位カテゴリーにあてはめた従来の評定尺度では、カテゴリー幅の近似的等間隔性と単峰的度数分布が認められ、経験的な予測と実験結果との間に矛盾が生じた (織田, 1976 a)。

判断過程におけるこの経験的な予測と実験結果との不一致の原因を解明するために、次の仮説が実験Ⅰで検討された。すなわち、意味強度の類似した反応語群が、同一の尺度に組み込まれたとき、被験者は被験者自身の言語的理解に従うというよりも、尺度の反応語系列に従って判断する。反応語「あまり……」を「どちらともいえない」の上位カテゴリーにあてはめた従来の配列法とは異なる尺度 (AD配列型尺度) による実験結果は、仮説を否定するものであった。AD配列型尺度では、意味強度の未分化な反応語群が用いられたことに加えて、反応語系列に表現形式の一貫性が欠如しているが、これが判断過程を混乱させる原因と解釈された。そこで、この解釈の妥当性を検討するために、実験ⅡとⅢが計画された。

実験Ⅱの目的は、表現形式の一貫性が確保された反応語系列における反応語の配列効果を検討することである。実験の結果、次の点が判明した。すなわち、尺度の反応語系列と被験者の理解する意味強度の順序との間に不一致が生じたとき、反応語系列に表現形式の一貫性があるならば、①意味強度の類似した反応語群を含む尺度では、被験者は尺度の反応語系列に基づいて判断する。ところが、②意味強度の十分に分化した反応語系列では、被験者は自己の理解に基づいて尺度の反応語系列を再系列化して判断する。

次に、実験Ⅲでは、意味強度の類似した反応語群を含む反応語系列における表現形式の効果が検討された。実験の結果、対比較調査で意味強度の類似度が高いと判定された反応語群で構成された反応語系列では、表現形式に一貫性があれば、その反応語系列に従って判断されるが、一貫性が欠如する場合には判断過程に混乱の生ずることが判明した。

実験Ⅰ—Ⅲによって、次の結論に達した。理想的な評定尺度の反応語系列とは、①表現形式上の一貫性が確保

されていること（表現的妥当性）、②反応語群の意味強度が十分に分化していること（意味強度的妥当性）、③反応語群は意味強度の順序に従って系列化されること（系列的妥当性）の3条件を満たすものである。従って、これらの3条件を満たさない尺度では、判断過程に混乱が生じやすい。しかし、意味強度の類似した反応語群を含む反応語系列であっても、表現的妥当性のある尺度ならば、被験者は尺度の反応系列に従って判断する傾向が強いために、判断過程に混乱が生ずることは少ない。また、意味強度的妥当性の高い尺度では、反応語系列と被験者の理解する意味強度の順序とが一致しないときは、被験者は自己の言語的理解に従って尺度の反応語系列を再系列化して判断する。この場合、系列的妥当性の高い尺度に比べれば、判断過程の混乱は生じ易いが、その混乱は比較的少ないことが判明した。従って、系列的妥当性に比べて、表現的妥当性と意味的妥当性が、評定尺度構成上より重要な要件であるといえよう。

（附記）

本論文の基礎となった「評定尺度におけるカテゴリ用語の意味強度効果」に対して、昭和51年度の文部省科学研究費補助金（一般研究D課題番号161037）の援助を受けた。本研究にあたり御指導いただいた名古屋大学教育学部内田良男教授、また、実験実施にあたり多大な御

便宜をお与え下さいました松阪女子短期大学柴原恭治教授、名古屋女子大学三輪弘道教授に厚く感謝いたします。

文 献

- 柿崎祐一 1974 知覚判断 八木 晁(監) 現代の心理学1 培風館 p.181
 織田揮準 1967 評定尺度に関する基礎的研究(I) 名古屋大学教育学部紀要(教育心理学科), 14, 7-42
 織田揮準 1970 日本語の程度量表現用語に関する研究 教育心理学研究, 18, 166-176
 織田揮準 1975 a 中性カテゴリの位置効果に関する研究 心理学研究, 45, 300-312
 織田揮準 1975 b 評定尺度に関する研究(III) 三重大学教育学部紀要 26(3), 111-128
 織田揮準 1976 a 反応語のカテゴリ内およびカテゴリ間意味効果 心理学研究, 46, 305-315
 織田揮準 1976 b 判断カテゴリ用語の研究 三重大学教育学部紀要, 27(4), 229-237
 田中良久 1973 尺度構成法 続 有恒・八木 晁(監) 心理学研究法16 東大出版会 p.98

(1977年1月21日受稿)

ABSTRACT

ORDER EFFECTS OF CATEGORY-WORDS

—a study on category scales—

by

Kijun Oda

SUMMARY

It is said that the category-words of a category scale must be differentiated from each other in strength of their semantic meanings and be arranged in order of strength of their semantic meanings. If the category-words not differentiated from each other in strength of their semantic meanings are used to construct a category scale, inconsistency may occur between the category-word order of the given category scale and the category-word order of the semantic meanings understood by the subjects and their judgment process with the category scale may

be confused.

The purpose of this study is to investigate the order effects of category-words on the subjects' judgment process with category scales which include the category-words undifferentiated from each other in strength of the semantic meanings.

From our experiment, we get the following results:

(1) The category-word order of a category scale must fulfill the following conditions:

A. Expressive validity: The category-word order of a category scale must possess an expressive

consistency.

a. Example of the category scale with the high expressive validity.

- very longaffirmative expression
- long.....affirmative expression
- neither long nor short...(neutral category)
- not so longnegative expression
- not longnegative expression

b. Example of the category scale without the validity.

- very longaffirmative expression
- long.....affirmative expression
- not so longnegative expression
- neither long nor short...(neutral category)
- not longnegative expression

B. Validity of strength of semantic meanings :

The category-words of a category scale must be differentiated from each other in strength of the semantic meanings. "Dochiratomoienai (neither long nor short)" and "amarinagakunai (not so long)" should not be included in the same category scale, because the significant difference in strength of their semantic meanings is not found between the words —55% of the 327 students judged "dochiratomoienai (neither long nor short)" is longer than "amari-

nagakunai (not so long)" in paired comparison ($\chi^2=3.12, p>0.05$)—.

C. Validity of category-word order : The category-words of a category scale must be arranged in order of strength of their semantic meanings.

Judgment process with a category scale which does not fulfill these three conditions must be confused.

(2) when a category scale which possesses the high expressive validity but lacks the validity of strength of the semantic meanings is given, the subjects have a strong tendency to judge according to the category-word order of the given category scale.

(3) When the category-word order of a category scale possessing the high validity of strength in the semantic meanings is not identical with the category-word order of strength of the semantic meanings understood by the subjects, the subjects rearrange the category-word order of the category scale according to their semantic understanding. So the expressive validity and the validity of strength of the semantic meanings are more important factors than the validity of the category-word order in order to construct a category scale.