

F-7 「曖昧」な潜在意味概念の分析にむけて: ノ・コトの間のバリエーションについての統計的アプローチ

山田彬堯 (ジョージタウン大学 [院] / 駿河台大学)、窪田悠介 (国立国語研究所)

要旨

理論言語学で用いられる容認性判断の言語テストは、多くの場合 (文法的対非文法的という) 二値的な判断を前提としており、一つの基準で明確な線引きをすることが難しい連続的な (ないし「曖昧」な) 概念の研究には不向きなことが多い。この結果、直観的にはある程度の妥当性があると考えられるにも関わらず、客観的に判定する基準がないために理論研究における位置づけを明確にするのが難しい概念が存在し、そのような概念が関与すると考えられる言語現象に関しては理論的なアプローチからの研究が立ち遅れている。本研究では、そのような概念の代表例として、補文標識ノとコトの選択に関わるとされる「具体性」の概念を事例としてとりあげ、統計的な手法を援用することで言語テストに立脚するだけでは捉えがたい現象を明確かつ科学的に分析できることを示す。

キーワード: 言語テスト、容認性判断、補文標識、ノとコト、直接性、ポリコリック相関係数、潜在概念

1 導入

統語論などの文法研究におけるデータとして用いられる容認性判断は、(容認可対不可の) 二値的な対立を前提としているものの、実際の容認性判断のタスクにおいては二値的な判断を行うことが困難であるため、様々な形で段階性を認めた基準で判断を行うことが通例となっている。

実際に観察されるデータである段階性をもった容認性判断と、多くの場合暗黙に前提とされている二値的な分類との間の関係をどう考えるかという問題に関しては様々な議論がある (Fanselow et al. 2006 など参照)。一般的な (そして、多くの場合ひとまず妥当と考えてよい) 立場としては、理論的前提としての文法性の概念の二値性は保ったまま、観察されるデータにおける段階的な容認性判断はすべてノイズであるとするものがある。この場合、典型的には、非文と正文の分類を保ったまま、正文ではあるが文法性以外の要因によって容認性が落ちる可能性を認め、中間的な容認性を持つ文はすべてそのようなノイズ (これには語用論的適切性や、文処理における負荷の大小などが含まれる) により容認度が落ちていていると考える。ノイズの性質やノイズ間の相互干渉により容認度の落ち具合は様々であるため、観察されるデータにおける容認度の段階性が説明できることになる。このような観点は、複雑な問題を整理し、データと理論の対応関係を明確にするために非常に有効であり、時として通説の信憑性に大きな疑義を投げかける重要な研究結果をもたらすことがある。例えば島の制約に関する研究の大きな転換点となった Kluender (1998) は、一般的には統語的な制約と考えられている島の制約の一つである Complex NP Constraint に違反する文に関して、文処理の問題と考えられている中央埋め込みと同種の要因によって容認度が段階的に変化することを示し、その後の島の制約に関する文処理や語用論的要因からのアプローチの嚆矢となった。

本論文では、容認性判断に関する段階性 (あるいは非二値性) が、上で概観した典型的ケースとは異なる要因で生じる可能性を認めた上で、文法現象の記述・分析に本質的に段階的な判断が要因として関与していると考えられる現象のケーススタディを行う。具体的な言語現象としては、補文標識ノとコトの選択を扱う。ノとコトの選択に関連する補文事態の「具体性」の概念が本質的に段階的なものであるという立場に立ち、そのような段階的な判断を客観的に収集する方法を提案し、提案手法の言語分析における有用性を示すことを目的とする。

本稿における我々の提案の背後には以下のような問題意識がある。理論言語学や記述言語学で用いられるいわゆる「言語テスト」は、多くの場合、背後に二値的な対立を前提とするため、一つの基準で明確な線引きをすることが難しい連続的な (ないし「曖昧」な) 概念の研究には不向きなことが多く、このため、言語テストを明確に作る事が困難な現象については、研究が立ち遅れている。手法の方法論的限界により研究対象が過度に限定されるという「言語テスト」の潜在的な問題を補うため、本研究は、統計的なアプローチを提案する。統計的な手法を援用することで言語テストに立脚するだけでは捉えがたい現象を明確かつ科学的に分析できることを示す。

2 背景

2.1 先行研究と課題

ノとコトの選択に関しては、補文事態の「具体性」が主要な要因であるという指摘がある（久野 1973, Josephs 1976, 影山 1977, 野田 1995）。この説は直観的である一方で、いくつかの問題を抱えている。第一に、「具体性」に完全に対応する言語テストが見つからない。もちろん、「具体性」にかかわりそうな概念が存在しないわけではない。例えば「場面レベル述語 (stage-level predicates)」と「個体レベル述語 (individual-level predicates)」の区別や、具体的な時を示唆する副詞との共起の可否などは、補文の具体性を判定するテストとして使えそうに思われる。しかし、以下の例が示すように、いずれの基準もノ・コトの選択とは連動せず、ノ・コトの選択にかかわる「具体性」とこれらのテストが関わる「具体性」が異なる概念であることを示唆している。

(1) 太郎は [花子が去年の7月7日に東京で結婚した {の/こと}] を知っている。

(2) 太郎は [花子がアメリカ人である {の/こと}] を知っている。

第二に、ノとコトの使い分けで明確な意味の差が生じない事例の存在が指摘されている（井上 1976 など）。

(3) [誰かが部屋に入ってきた {の/こと}] に気づいた。 (井上 1976)

また、山田・窪田 (2018) は、コーパス分析と従来の言語テストによる内省分析を組み合わせることでコトとの結びつきが強い述語 (主に信念述語) とノとの結びつきが強い述語 (主に信念主体と外界の事態との関係性を規定する述語群) の意味的性質の違いを指摘しているが、ノとコト両方と共起できる述語に関して、(a) これらの述語がどのような意味的性質を持つか、(b) 述語の意味的性質とノ・コトの選択に関連性があるか、の二点を未解決の課題として残している。

2.2 研究の目的

これまでの研究で異なる研究者がたびたび結論づけてきたように「具体性」という概念とノ・コトの選択には関連性があると考えられる。しかし、それを言語テストという「客観的」な形で取り出すことは難しい。だが、言語テストによって取り出せないからと言ってこの「具体性」という概念そのものが誤った、あるいは当該の言語現象の分析にとって意味のない概念であるという証拠にはならない。むしろ、求められるのは、「言語テストには乗りにくい、ネイティブスピーカーの間で共有されている概念」を、「客観的」に調査、議論する方法論を確立することである。

このように直接観測することは難しいものの学問上の目的のために定義され使用される概念は、「構成概念 (construct)」と呼ばれる。本研究では、まず第一に、このノ・コト選択にかかわる「具体性」という構成概念を測定するための尺度を提案し、第二に、実際にこの尺度によって測定された値がノ・コト選択と強い相関関係を持つことを示す。具体的には、山田・窪田 (2018) で未解決の問題として残されたノ・コト両方を取る述語である認識動詞「分かる」「知る」に着目し、要因間の相関を計る尺度の一つであるポリコリック相関係数を算出すると、ノ・コトの選択と提案された尺度に高い相関が示唆されたことを報告する。認識動詞は、意味的にそもそも「具体的」な出来事としか共起不可能な知覚動詞や、同様にそもそも「抽象的」な事柄しか補語としてとりえない信念動詞に比べて、共起可能な補文事態の種類幅が広いと考えられる。そのため、仮にノとコトの選択が主文動詞の意味によって一意に決まるのではなく、それ自体「具体性」の程度のようなものに影響されているとすると、これらの動詞の補文におけるノとコトの選択は、補文事態そのものの性質をより直接的に反映している可能性がある。この仮説の妥当性を検証するために、以下ではノ・コト両方を取る述語である認識動詞に研究対象を絞る。

3 仮説と尺度化

3.1 仮説

まず、「具体性」の概念がノ・コトの選択に関わっているという言語直観をもう少し掘り下げて考えてみたい。先行研究などで繰り返し指摘されている点として、直接観察された出来事はノでマークされやすいという傾向がある。典型的な例としては(4)のようなものがあり、母語話者の直観に照らし合わせても、この例のような文においてはノがコトより自然に感じられる。

(4) サングラス越しに [イズムの目が真っ赤になっているの] がわかった。(PM32_00102)

反対に、「抽象的」な概念はコトでマークされる傾向があるということも古くから知られており、これに関しては(5)のようなものが典型例である。

(5) [郵便と電話はそれぞれのメディアとしての長所・欠点を補う関係にあること] 分かる。(OW3X_00831)

この二つのタイプの文の間に様々な境界的な例が存在するのだが、実例の観察から一つ特徴的なケースとして括り出せそうなクラスを見出すことができる。典型例な例としては以下のようなものである。

(6) 夢二を訪ねた時、家の様子を一目見て、私は [誰も看病する者の居ないの] を知った。(LBk7_00014)

ここでは、補文事態は直接的な観察の対象になっていないが、補文事態が表す命題の証拠として、何らかの直接的に知覚された事態が存在する。直観に照らし合わせて考えてみると、このタイプの例においても(4)と同様、ノがコトより自然という傾向があるように感じられる。

以上の観察は、いまだ印象レベルのものに過ぎず、分析のための手がかりとはなるが、分析を支える客観的な証拠とはなりえない。そこで以下では、上の三つのクラスを、厳密な境界が曖昧な「具体性の高さ」という連続的なスケールの異なった区間に対応するクラスであると仮定して、そのようなクラス分けが異なる母語話者間でどの程度一致するかということを統計的なアプローチを用いて検証する。

3.2 尺度化

上の議論を踏まえて、先行研究で「具体性」と呼ばれている概念を、本稿では以下の定義で尺度化／操作可能化することを提案する（以下、この尺度を「直接性」と呼ぶ）。これは、「補文命題が直接的に知覚された事態を描写する命題、または直接的に知覚された事態を証拠として持つ命題である場合、ノと共に起る傾向が高い」という仮説を反映している。(7)の尺度は数字が低いほど直接性が高くなっている。

(7) 1: 補文が表す事態は、直接知覚することができる

2: 補文の内容は、直接知覚することは出来ないが、直接知覚した事態から解釈／推論を行うことで得られる

3: 補文の内容は、具体的な直接経験と結びついているとはいいがたく、思弁的、抽象的な内容を表す

この基準は、定義を行うことを主目的にしているというより、1から3という順序を持った値を測定するための尺度化を行うために用意された指針である。重要な点は、「直接知覚」のような鍵となる概念にあえて厳密な定義を与えることをせずに「直接性」の概念を規定していることであり、ここでは個別例に関して評定者間の判断の揺れが生じることを、むしろ、想定している。

「直接性」の尺度が日本語話者にとって共有されやすい概念であり、かつ、評定者が理性的に判断を下しているかぎり、評定結果に強い相関が見出されるはずである。さらに、先行研究の直観、また、上記の仮説が妥当であれば、この操作可能化された「直接性」という潜在概念と「ノ・コト」選択の間にも強い相関が認められるはずである。

4 手法・分析

4.1 分析手順

「こと／の」の補文選択と「補文命題の抽象度」の関係を調べるために、次の手順で調査を行った。第一に、ノ・コトを含んだ例文を用意する必要がある。本研究では自然な環境で産出された用例を使用するために、BCCWJ からランダムに例文を抽出し、(i) ノと共に起した「知る」、(ii) コトと共に起した「知る」、(iii) ノと共に起した「分かる」、(iv) コトと共に起した「分かる」を、それぞれ25件ずつ、計100例を採取した。

第二に、実際に補文標識がノであるかコトであるかが評定者の判断に影響を与えることがないように、(8)に示すように補文標識部分を空欄としたうえで、複数の評定者に補文命題（すなわち、(8)の例で言えば下線が引かれた部分）の直接性を(7)に記された1-3の尺度で判断させた。^{*1}

(8) そのとき初めて、あのとき聞いた直江の声が幻ではなかった こと・の] を知った。

第三に、アノテーションを施す過程で、例文間の交換可能性を保証するために以下の3項目に該当する事例を考察

^{*1} 試験的な段階である本研究では、発表者二名が評定者を務めたが、将来の研究では、調査対象が何であるかを知らない母語話者を実験協力者として募集することが望ましい。

対象から除外した。

(9) 「という」という表現が前接するもの

[「闇金」での多重債務の成れの果てが、TVドラマでは無くて、実際の出来事と言う **の**] が分かりました。

(10) 補文ではなく関係節が「こと」を修飾している事例

[親のやるべき **こと**] が分かりました。

(11) 補文に表された内容が、純粋な知識ではなく術・技術・やり方を表す、「知る」の周辺的な事例

[子供たちは一度勢いに乗ってしまうと止まる **こと**] を知らない。

これらの調整の結果、計 84 例を分析対象とした。

第四に、二週間以上の間において、二名の評定者が 3 回評定を下し、これを本研究で扱うデータとした。複数回測定を行ったのは、同一評定者の判断のばらつきを考慮するためである。同じ評定者であったとしても、場合によっては、複数回測定を行うと結果が揺らぐ可能性がある。しかし、評定者が理性的な判断に基づき回答をしている場合、そして、「直接性」という概念が日本人母語話者にとってなじみ深い概念である場合、評定結果は大きく変動することはないと予想される。逆に、どの程度変動するのかを評価することで、「直接性」という構成概念のロバストネスが検証される。

4.2 記述統計：クロス集計表

評定者間の判定結果は、下に見るようにクロス集計表の形で表現することができる。例えば、この表1は、評定者 A と評定者 B の第一回目における実際のコーディングの結果を示している。

B1/A1	「1」	「2」	「3」
「3」	0	2	49
「2」	1	3	0
「1」	18	6	5

表1: 評価者の判断結果

B1/A1	「1」	「2」	「3」
「3」	2	0	49
「2」	1	3	0
「1」	18	6	5

表2: 架空の判断結果

4.3 推測統計

4.3.1 完全一致率／質的相関

問題となるのは、この記述統計（クロス集計表）を前に、どのような分析を行うかということである。例えば、次のような指標を算出することもできるが、データの性質を鑑みれば、これらの指標には不十分な点がある。

- **指標 1（完全一致率）**：第一は、完全一致率である。これは、表の対角線の部分の和が全体に占める割合であり、今回は $(49 + 3 + 18) / 84 * 100 = 83.3\%$ となる。表3はこの結果をまとめたものである（表中の A1、B2 などは、それぞれ評定者 A の第一回目の評定結果、評定者 B の第二回目の評定結果を意味している）。解釈のしやすい指標である反面、この指標は対角線以外の部分の違いを捨象するという欠点を持つ。表2の架空データは完全一致率の上では表1と同じだが、A が「1」と付けたものを B が「3」と付けたものが多くなっているため、合意の度合いは実質的には低いはずである。しかし完全一致率はこの点を看過してしまう。

	A1	A2	A3	B1	B2
A2	0.9				
A3	0.77	0.8			
B1	0.83	0.77	0.7		
B2	0.82	0.82	0.7	0.77	
B3	0.76	0.76	0.67	0.71	0.94

表3: 完全一致率

- **指標 2（質的相関）**：第二は、 χ^2 検定であるが、これは「二つの質的変数の間に連関が全くない状況を想定し、その状況下で期待される結果と実際のデータとの差異を評価する有意性検定」であり、両者の連関の程度については分からない。Cramer の連関係数や ϕ 係数も表1と表2の差を捉えられず不十分である。

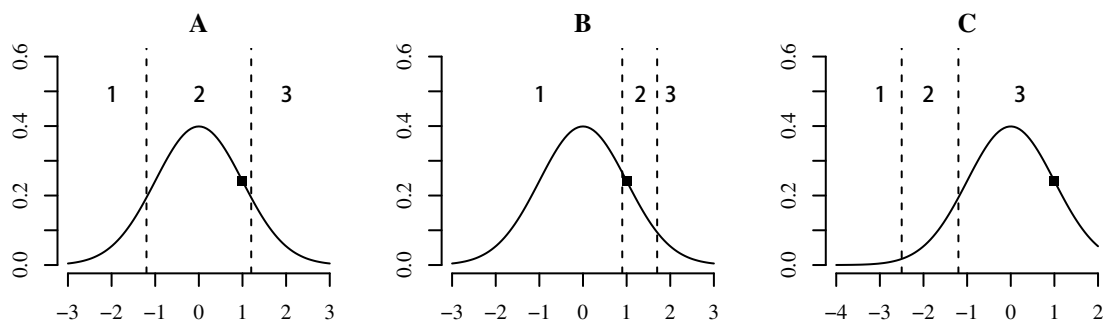


図1: 評価者の印象のモデル化

4.3.2 ポリコリック相関係数

今回のデータは、質的変数であることに加え、変数の大小に順序がある順序変数である。したがって、この順序という情報を最大限生かした連関の指標を求めることが望ましい。このような指標としてポリコリック相関係数が知られている。以下、この指標の基本的な考え方を説明したうえで、本事例における推定結果を報告する。

(1) 潜在変数の想定: この相関係数は、離散的な順序変数の背後に連続的な潜在的関係があることを想定し、この連続的な値が複数の閾値によって離散化されたものが観測されているのだという想定に立つ。例えば、実際には、「1, 2, 3 / A,B,C」のような離散的な回答を行う場合であったとしても、その回答の背後には、評定者の印象に相当する連続的な潜在変数が存在すると考えられる。本事例の場合、「直接性」は、評定者が抱く連続的な値（印象）であり、評定者の回答は、その潜在的には連続的な値を特定の閾値に基づいて離散化したものだというモデルを想定する。

(2) 連続変数の従う分布: 今回は、例文が BCCWJ からランダムに抽出されているため、中には、評定者の直観において「とても直接性が高そうに思われるもの」もあれば、「直接性が高そうだと思うがそこまではっきりは言い切れないもの」なども存在するはずである。極端な事例ほど全体に占める数は少ないはずだという発想から、各事例の「直接性」は、図1に示されたように、正規分布に従っているという簡便な仮定／モデルを採用する。もともと観測されない連続値であるので、ここでは標準正規分布を想定する。もちろん、何らかの理論的な根拠が存在している場合には、正規分布以外の分布を用いることも可能ではあるが、特にそのような事前情報がない本研究では、最も無標な分布である標準正規分布を選択したと考えられたい。

(3) 閾値の独自性: それぞれの評定者がもつ閾値には独自性が存在するはずである。例えば、仮想的な三人の評定者を考える。図1の点線はこれらの評定者の閾値を表している。評定者 B は、評定者 A に比べ、より多くのものに「1」という判断をくだす傾向に、評定者 C は、評定者 A に比べより多くのものに「3」をつける傾向にある。図1の丸と四角で表した同一の事例を判断したとしても、評定者の独自の閾値により、丸はそれぞれ「2」、「1」、「3」と評価され、また、四角はそれぞれ「2」、「2」、「3」のように評価される。

(4) 複数の評価者の相関関係: 複数の評価者の「直接性」についての判断には、例えば、図2の最左端のパネルに示されたような正の相関関係があることが予想される。すぐわかるように、これは評定者 A が高い値を感じるものには評定者 B も高い値を感じるはずだという私たちの直観を反映している（図2は専門的には同時密度分布と呼ばれるものを視覚化したものである）。一つ右隣の図は、これを上空から見た時の鳥観図である。x 軸（評定者 A の判断）が大きくなれば、y 軸（評定者 B）の判断も高まる様子が示唆されている。各事例は、この山からランダムに抽出された一地点であるとモデル化され、この様子を表したものが右から二番目の図である。実際には、評定者 A にも評定者 B にもこの連続値を離散化して報告するため、点線で示されたように、閾値が存在し、この閾値によって、この空間は、 $3 \times 3 = 9$ 個の部分空間に分割されていることになる（すなわち、評定者 A が「3」、評定者 B が「3」の空間；評定者 A が「3」、評定者 B が「2」の空間；...；評定者 A が「1」、評定者 B が「1」の空間である）。現実に観察できるクロス集計表（例えば表1など）はこの部分空間に入り込んだ地点の数を集計したものであるとみなされる（最右端の図）。

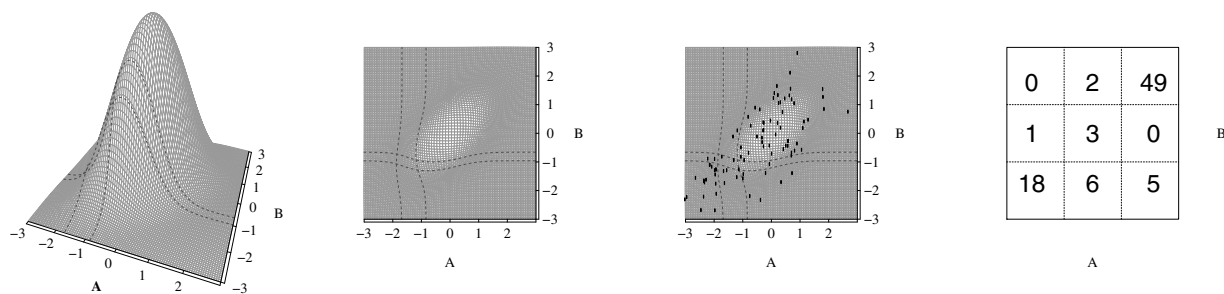


図2: 評定者間の間の関係 (相関係数 0.6)

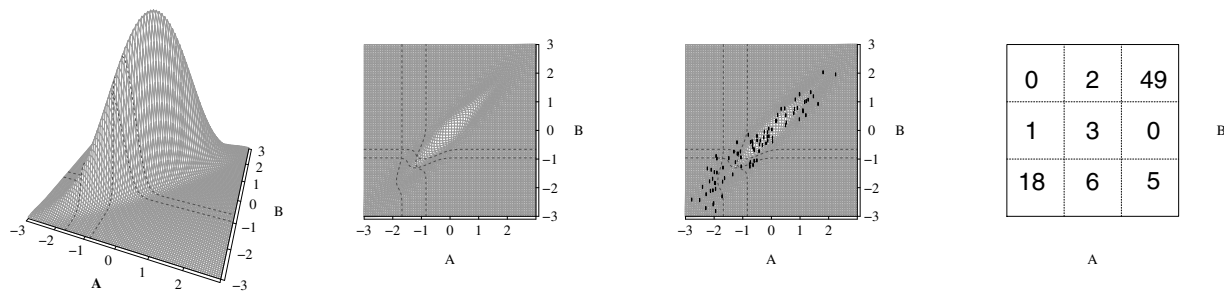


図3: 評定者間の間の関係 (相関係数 0.95)

	A1	A2	A3	B1	B2	B3
A2	0.99 ± 0.01					
A3	0.92 ± 0.04	0.9 ± 0.08				
B1	0.95 ± 0.03	0.94 ± 0.06	0.88 ± 0.11			
B2	0.96 ± 0.02	0.96 ± 0.04	0.85 ± 0.11	0.96 ± 0.05		
B3	0.94 ± 0.03	0.95 ± 0.05	0.84 ± 0.12	0.94 ± 0.06	NA	
koto	0.82 ± 0.08	0.87 ± 0.13	0.86 ± 0.14	0.75 ± 0.19	0.81.15	0.84 ± 0.13

表4: 推定結果 (ポリコリック相関係数)

(5) 推定: 図2は、評価者 A と評価者 B の間の相関関係が 0.6 であるときの関係を示したもののだが、この 0.6 という値は説明の都合上恣意的に選択された値であり、もちろん、相関関係が 0.6 でなくとも、同一のクロス集計表が得られる可能性もある (図3)。そこで、このクロス集計表が得られたときに最も妥当だと考えられる、潜在変数同士の相関係数を推定する。このように、観察された順序尺度から、背景に想定される連続値の真の相関係数を推定したものをポリコリック相関係数 (polychoric correlation) と呼ぶ。表1では、その値は $\hat{\rho} = 0.945$ と点推定され、評定者の間で強い相関関係が存在していることが示唆される (推定には R の polycor パッケージに含まれる hetcor 関数を使用し、推定方法には最尤推定を用いた)。

(6) 解釈: 今回計測した各評定結果から推定されたポリコリック相関係数を示したものが表4である (表中の土に後続する値は標準誤差二つ分の値を示している)。表中の koto は実際のノ・コト選択を示している。^{*2}

- (評定者間の相関): 表4の破線より上の部分は、評定者間の相関を示している。最も高い相関がみられたのは A1 と A2 の間であり、点推定で $\hat{\rho} = .99$ という極めて強い相関が推定された。最も相関が小さかったものでも 0.84 という高い値を示していることから (A3 と B3 の間)「直接性」という概念は、評定者の間で極めて共有されやすい概念であることが示唆された。

^{*2} 表中の B2 と B3 については、最適化に難が生じた。ただし、このペアは「直接一致率」の上で最も高い値を示していたため、評定者間の相関が高いという一般化自体は損なわれないと考えられる。

- (評定者と補文選択の間の相関)：評定者の持つ「直接性」とノ・コト選択の間のポリコリック相関係数については、破線部の下の行に示されている。ここでは、コトを「1」、ノを「0」とコーディングしているため「直接性」の値が高くなればなるほどコトが強く選択されるという傾向が示されている（厳密に言えば、ここでは Albert and Chib (1993) で論じられているように、ノ・コトの二値選択の背後にも連続的な実数が想定され、その潜在変数と評定者の直観との相関が推定されている)。もっとも強い相関は A2 との相関でおおよそ $\hat{\rho} = 0.87$ 、もっとも低いものが B1 とのものでおおよそ 0.75 程である。ここから、「直接性」とノ・コト選択の間には中程度～高い相関が存在することがわかるが、他方で、それだけでは完全にノ・コト選択を説明できず、他の要因による影響が示唆される。

5 結論と今後の課題

本研究では統計分析により、「直接性」という、従来の言語テストに基づく研究手法では扱いづらい「曖昧」な概念が、具体的な言語現象と連動していることを、客観的なデータに基づき示した。このことにより、ノ・コト選択に「具体性／直接性」が中心的な要因として関わっているという、古くから指摘されていたが検証が困難だった仮説の科学的な裏付けを行った。

将来の研究には「直接性」以外の別の潜在概念（「フォーマリティ」など）をさらにモデルに組み入れ、それらの関係を吟味するという発展が期待される。また、二値的な言語テストに馴染みにくい現象は、「ノ・コト」選択以外にも多数存在する。中でも、敬語構文における敬意の度合いと各種敬語表現の選択関係、事物の話者／聞き手への帰属度 (cf. 情報の縄張り) と指示詞選択の関係などが本手法の適用例として考えられる。

謝辞

本発表は国立国語研究所共同研究プロジェクト「対照言語学的観点から見た日本語の音声と文法」の研究成果の一部である。

参考文献

- Albert, James H. and Siddhartha Chib (1993) “Bayesian Analysis of Binary and Polychotomous Response Data.” *Journal of the American Statistical Association*, Vol. 88, pp. 669–679.
- Fanselow, Gisbert, Caroline Féry, Matthias Schlesewsky and Ralf Vogel (eds.) (2006) *Gradience in Grammar: Generative Perspectives*. Oxford University Press.
- Josephs, Lewis S. (1976) “Complementation,” in Shibatani, Masayoshi ed. *Japanese Generative Grammar, Vol. 5 of Syntax and Semantics*, New York and Tokyo: Academic Press, pp. 307–369.
- 影山太郎 (1977) 「いわゆる日本語の『名詞補文辞』について」『英語教育』25 (11), 66–70.
- Kluender, Robert (1998) On the Distinction between Strong and Weak Islands: A Processing Perspective. In *The Limits of Syntax*, eds. Peter Culicover and Louise McNally. Vol. 29 of *Syntax and semantics*, 241–279. San Diego: Academic Press.
- Kuno, Susumu (1973) *The Structure of the Japanese Language*, Cambridge, Mass.: MIT Press.
- 野田春美 (1995) 「埋め込み節」, 宮島達夫・仁田義雄 (編) 『日本語類義表現の文法 (下) : 複文・連文編』, pp. 419–428, くろしお出版
- 井上和子 (1976) 『変形文法と日本語 (上)』, 大修館.
- 山田彬堯, 窪田悠介 (2018) 「ノとコト再考: 主文述語の新たな意味分類に向けて」, 『日本言語学会第 157 回大会予稿集』, pp. 276–281.