

「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）

1 はじめに

1 はじめに

容認性判断

想定：「容認」 vs. 「非容認」 の二値的対立

実際：段階性

標準的な立場

文法性の二項対立を保ったまま、観察されるデータの段階性はノイズである（語用論；文処理の不可など）。 Kluender (1998)

本研究の対象

「の／こと」選択を扱う。

1 はじめに

「具体性」

先行研究では、「具体性」という概念が提唱されてきている。

- 正しそうではあるものの、
実際に、各事例において命題内容が具体的かは明確に決まらず、
研究がしづらい

本研究の目標

- 方法論的な限界が、研究対象の限定を引き起こさないために、
本質的に段階性のある現象に対し、客観的に議論を行う土台の整備

2 背景

2 背景：先行研究と課題

「具体性」に注目した先行研究

- 久野 (1973)
- Josephs (1976)
- 影山 (1977)
- 野田 (1995)

“「コト」と「ノ」の違いは、前者が、抽象化された概念を表すのに対して、後者が、五感によって直接体験される具体的動作、状態、出来事を表すことである（久野 1973: 140）”

例

“命題の中には、元々抽象的な概念であって、具体的な動作、出来事を表さないものがある。（久野1973:140）”

(1) 太郎は 人間が羽のない二本足の動物である{こと／*の} を知らなかった。
(容認度判断は久野による)

2 背景：先行研究と課題

限界

(限界1) テスト不在

「具体性」に対応する言語テストが見つからない。

例：stage/individual-level predicatesや特定の時を表す副詞

- (2) 太郎は[花子が去年の7月7日に東京で結婚したこと／の]を知っている。
- (3) 太郎は[花子がアメリカ人である こと／の]を知っている。

(限界2) 明確な意味の使い分けのない事例

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

2 背景：先行研究と課題

コーパス調査からの傾向

「こと」、「の」に偏った動詞ばかりではなく、中立的なふるまいを見せる動詞も多い。

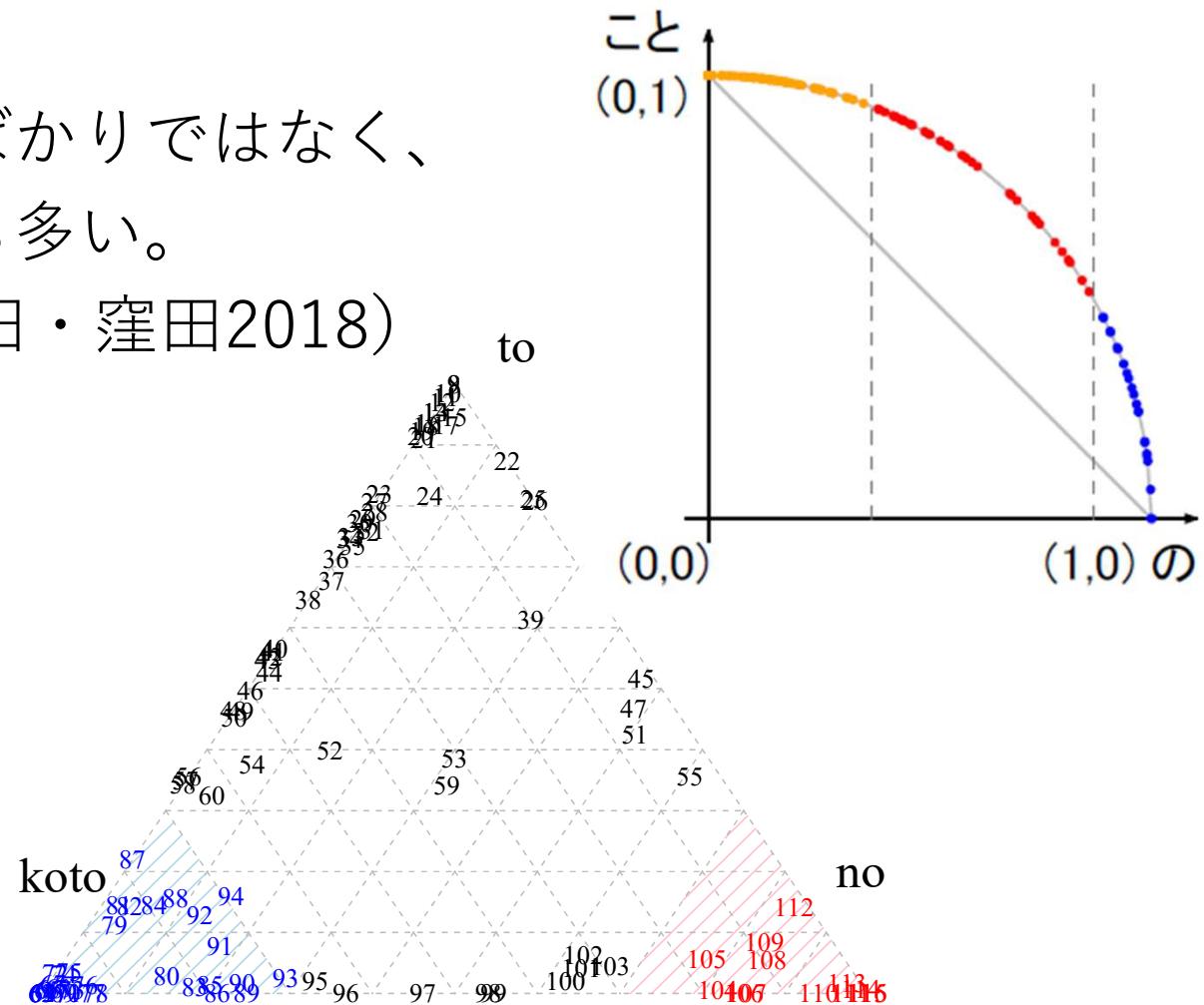
(Yamada 2018; 山田・窪田2018)

課題

これら中立的動詞は
どのような選択傾向を示すのか？

例：認識動詞

「分かる」「知る」など



2 背景：研究の目的

方法論におけるジレンマ

一方で、「具体性」という概念とノ・コト選択に関連がありそう

他方で、それを言語テストという「客観」的な形で捉えられない

「構成概念」のモデル化

直接観察することは出来ないが、学問上の目的のために定義され使用される「構成概念 (construct)」を検証する。

- (1) 「直接性」という構成概念を想定し、それを測定する尺度を整備
- (2) 「直接性」という構成概念がノ・コト選択と相関することを示す
⇒ ポリコリック相関係数の算出

3 仮説と尺度化

3 仮説と尺度化

仮説

一方で、「具体性」という概念とノ・コト選択に関連がありそう

他方で、それを言語テストという「客観」的な形で捉えられない

① 直接観察 かつ 「の節」

(5) 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)

② 中間的な事例

(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)

③ 抽象的 かつ 「こと節」

(7) [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]が分かる。(OW3X_00831)

3 仮説と尺度化

尺度化：「具体性」と呼ばれる概念を尺度可能化

一方で、「具体性」という概念とノ・コト選択に関連がありそう

他方で、それを言語テストという 「客観」的な形で捉えられないか

① 直接観察 かつ 「の節」

(5) 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)

② 中間的な事例

(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)

③ 抽象的 かつ 「こと節」

(7) [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]が分かる。(OW3X_00831)

3 仮説と尺度化

尺度化：「具体性」と呼ばれる概念を尺度可能化

一方で、「具体性」という概念とノ・コト選択に関連がありそう

他方で、それを言語テストという 「客観」的な形で捉えられないか

「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、ノと共に起する傾向が高い」という仮説

(8) 1: 補文が表す事態は、直接知覚することができる

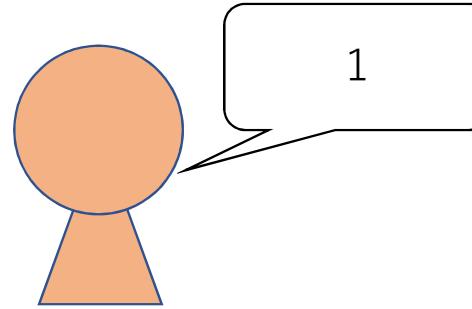
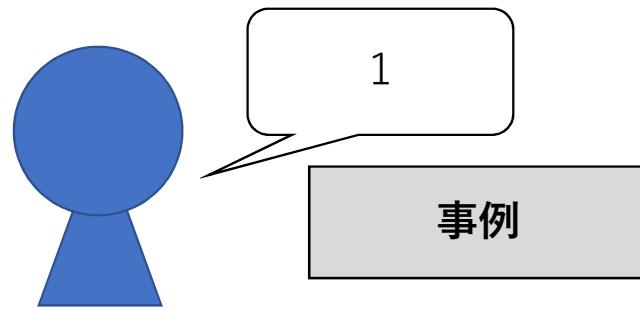
2: 補文の内容は、直接知覚することは出来ないが、直接知覚した事態から解釈／推論を行うことで得られる

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がたく、思弁的、抽象的な内容を表す

4 手法と分析

4 手法と分析

測定：事例に対し一定の手続きに沿って数値を割り当てる



BCCWJ からランダムに例文を抽出
(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。

例

評定者 1 評定者 2

エメットがとなりで身をかたくする【koto/no】がわかった。	
聞き覚えのある声が近付いて来る【koto/no】が分かる。	
見覚えのある座敷に、自分は転がされている【koto/no】を知った。	
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【koto/no】を知った。	
耳の付け根まで赤くなる【koto/no】が分かった。	
顔が赤くなる【koto/no】がわかった。	

評定者 1	評定者 2
1	1
1	1
1	2
1	2
1	2
1	2

4 手法と分析

除外した事例

(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] がわかりました。

(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。

例

評定者 1 評定者 2

エメットがとなりで身をかたくする【koto/no】がわかった。

1 1

聞き覚えのある声が近付いて来る【koto/no】が分かる。

1 1

見覚えのある座敷に、自分は転がされている【koto/no】を知った。

1 2

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【koto/no】を知った。

1 2

耳の付け根まで赤くなる【koto/no】が分かった。

1 2

顔が赤くなる【koto/no】がわかった。

1 2

4 手法と分析

(1) 記述統計

クロス集計表

評定者間の一致／不一致をまとめる記述としてクロス集計表を用いることができる。

限界

どの程度背後に相関関係があるのか、不透明。

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

評定者1 評定者2



エメットがとなりで身をかたくする【koto/no】がわかった。

聞き覚えのある声が近付いて来る【koto/no】が分かる。

見覚えのある座敷に、自分は転がされている【koto/no】を知った。

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【koto/no】を知った。

耳の付け根まで赤くなる【koto/no】が分かった。

顔が赤くなる【koto/no】がわかった。

1	1
1	1
1	2
1	2
1	2
1	2

4 手法と分析

(1) 記述統計

クロス集計表

評定者間の一致／不一致をまとめる記述としてクロス集計表を用いることができる。

限界

どの程度背後に相関関係があるのか、不透明。

A. 完全一致率

B. χ^2 二乗検定

C. Cramerの連関係数/ファイ係数

表1と表2の違いを区別できない。

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 手法と分析

(2) 推測統計

ポリコリック相関係数

順序という情報を最大限生かした連関の指標

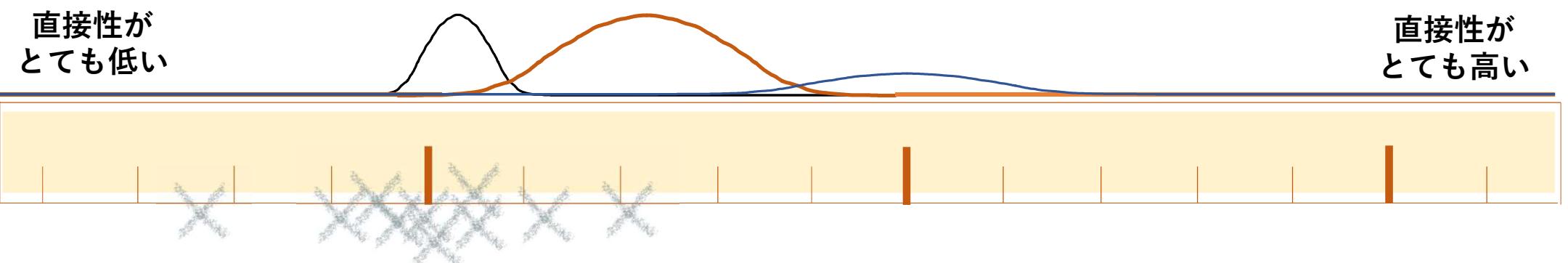
Example

エメットがとなりで身をかたくする【koto/no】がわかった。

顔が赤くなる【koto/no】がわかった。

Step 1: 潜在変数の想定

離散的な順序変数の背後に連続的な潜在変数



4 手法と分析

(2) 推測統計

ポリコリック相関係数

順序という情報を最大限生かした連関の指標

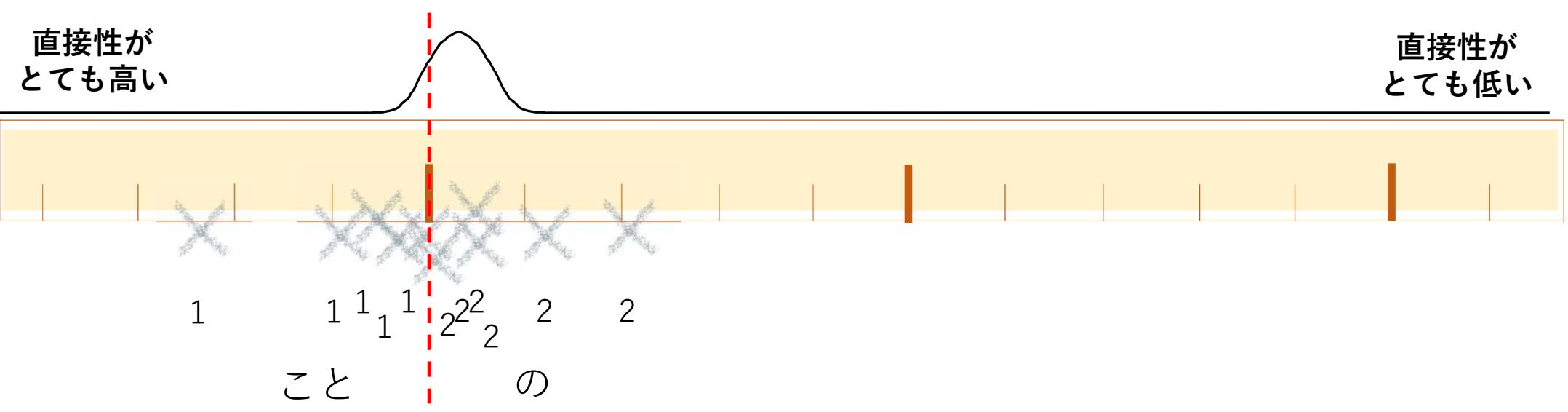
Example

エメットがとなりで身をかたくする【koto/no】がわかった。

顔が赤くなる【koto/no】がわかった。

Step 2: 離散化

離散的な順序変数の背後に連続的な潜在変数
連続的な値は閾値によって離散化される



4 手法と分析

(2) 推測統計

ポリコリック相関係数

順序という情報を最大限生かした連関の指標

Example

エメットがとなりで身をかたくする【koto/no】がわかった。

顔が赤くなる【koto/no】がわかった。

私は誰も看病する者の居ない【koto/no】を知った。

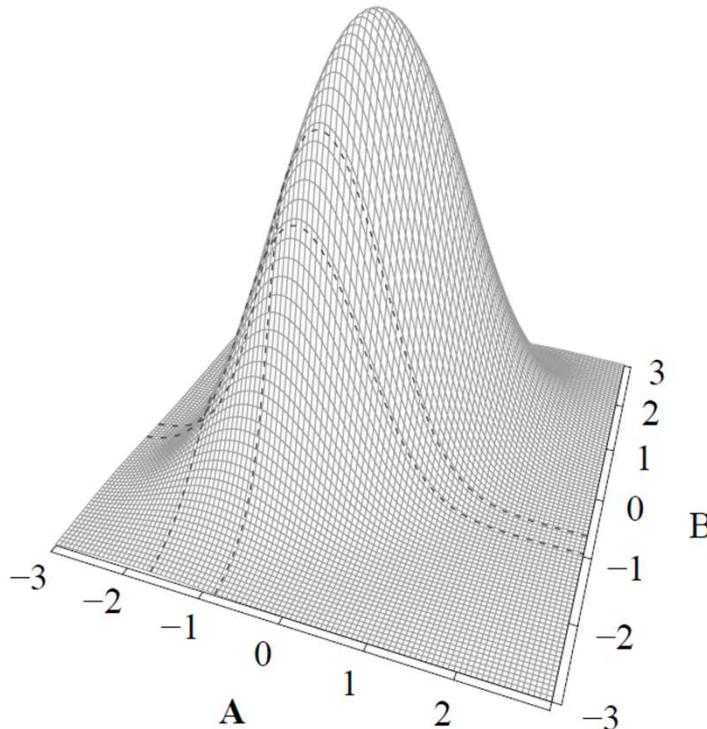
測定 1 (評定者Aの第一回目)		測定 2 (評定者Bの第一回目)	
観測値	直接性	観測値	直接性
1	4.3	2	4.0
2	2.3	2	2.2
3	-0.8	3	-0.5

4 手法と分析

(2) 推測統計

ポリコリック相関係数

順序という情報を最大限生かした連関の指標



測定 1
(評定者Aの第一回目)

観測値

1
2
3

直接性

4.3
2.3
-0.8

測定 2
(評定者Bの第一回目)

観測値

2
2
3

直接性

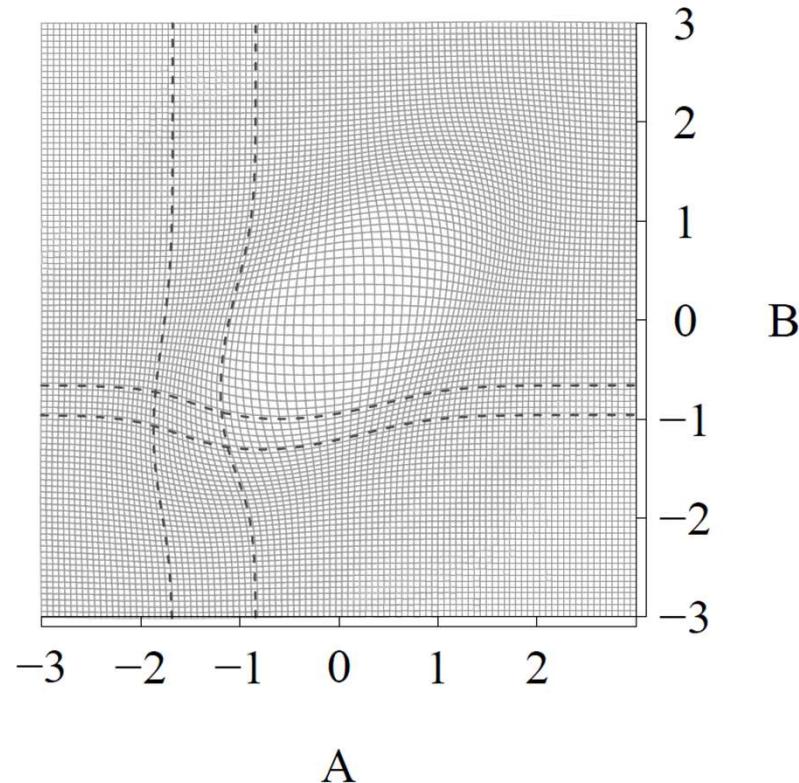
4.0
2.2
-0.5

4 手法と分析

(2) 推測統計

ポリコリック相関係数

順序という情報を最大限生かした連関の指標



測定 1
(評定者Aの第一回目)

観測値

1

2

3

直接性

4.3

2.3

-0.8

測定 2
(評定者Bの第一回目)

観測値

2

2

3

直接性

4.0

2.2

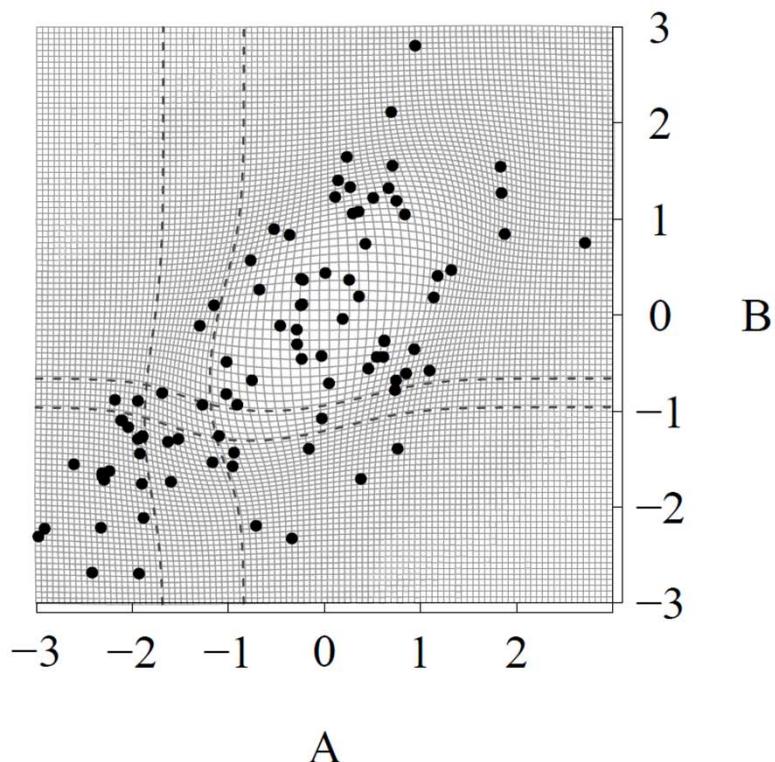
-0.5

4 手法と分析

(2) 推測統計

ポリコリック相関係数

順序という情報を最大限生かした連関の指標



測定 1
(評定者Aの第一回目)

観測値

1
2
3

直接性

4.3
2.3
-0.8

測定 2
(評定者Bの第一回目)

観測値

2
2
3

直接性

4.0
2.2
-0.5

4 手法と分析

(2) 推測統計

ポリコリック相関係数

順序という情報を最大限生かした連関の指標

0	2	49
1	3	0
18	6	5

B

A

$$\hat{\rho} = 0.945$$

測定 1 (評定者Aの第一回目)		測定 2 (評定者Bの第一回目)	
観測値	直接性	観測値	直接性
1	4.3	2	4.0
2	2.3	2	2.2
3	-0.8	3	-0.5

4 手法と分析

(2) 推測統計

ポリコリック相関係数

順序という情報を最大限生かした連関の指標

	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)

4 手法と分析

(2) 推測統計

ポリコリック相関係数

順序という情報を最大限生かした連関の指標

評定者間の相関

最大値 (A1/A2): 0.99

最小値 (A3/B3): 0.84

⇒評定者間で共有しやすい概念

	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)

4 手法と分析

(2) 推測統計

ポリコリック相関係数

順序という情報を最大限生かした連関の指標

評定者と補文選択間の相関

最大値 (A2): 0.87

最小値 (B1): 0.75

⇒ 中から強い相関

	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 結論と今後の課題

5 まとめと今後の課題

言語研究の手法における多様化

言語テスト

離散化された判断に基づく、強い主張

統計的なアプローチ

連続的な判断／潜在的な判断のモデル化により“曖昧な”概念の分析

今後の課題

フォーマリティ、敬意の度合い、帰属度…

5 まとめと今後の課題

言語研究の手法における多様化

言語テスト

離散化された判断に基づく、強い主張

統計的なアプローチ

連続的な判断／潜在的な判断のモデル化により“曖昧な”概念の分析

事例研究

「の」と「こと」の選択

- ・「直接性」という概念は、統計的なモデルとして研究の対象になりうる
- ・完全な相関ではないことから、その他の要因の存在も示唆

今後の課題

- ・複数の要因を加味した統計モデルに基づく推論
- ・フォーマリティ、敬意の度合い、帰属度…

ご清聴ありがとうございました

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。

- 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.