

数量化 類

人間・環境学研究科 言語科学講座

D1 増田 将伸

数量化 類とは

- 質的データの分析に用いる。
- 項目間の関係を、少数の因子の相関関係を用いて表す。

量的データと質的データ

量的データ:

間隔尺度値(100点満点)による女優評価

女優 視聴者	竹内 結子	伊東 美咲	矢田 亜希子
A	90	70	40
B	30	80	70
C	75	50	40
D	80	40	85

主成分分析

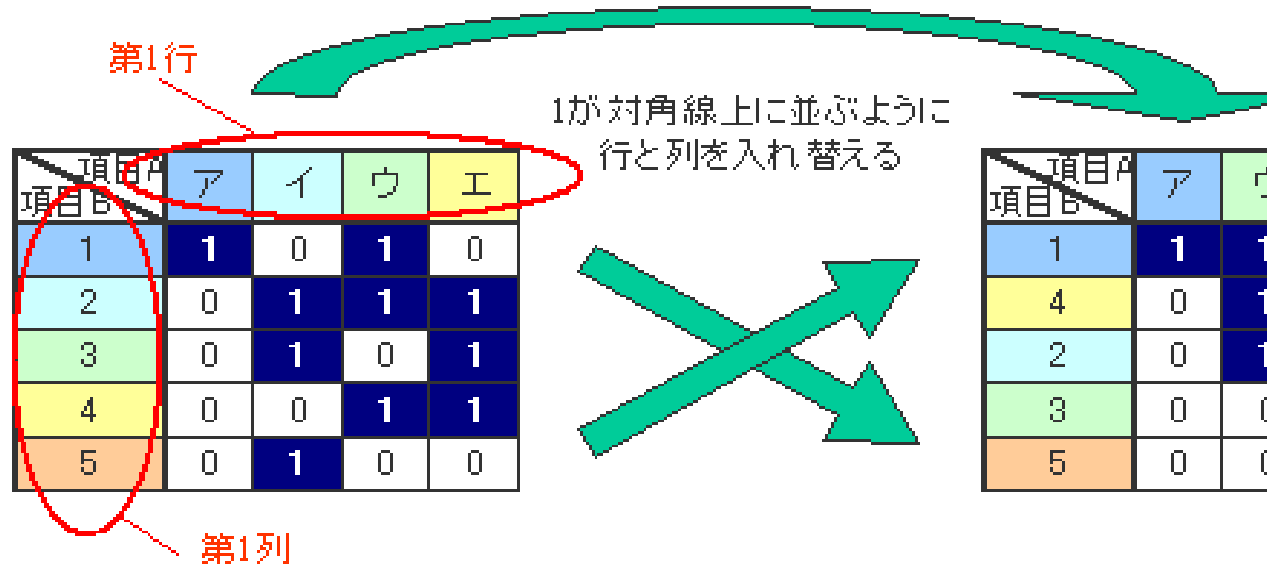
質的データ:

名義尺度値(好き嫌い)による女優評価

女優 視聴者	竹内 結子	伊東 美咲	矢田 亜希子
A			
B			
C			
D			

数量化 類

手順



- ・クロス集計結果を用いて、行の要素と列の要素の相関関係が最大になるように並べ替える。
- ・通常の表やグラフから簡単には読み取れないデータの傾向を把握することが可能になる。
...反応の似たサンプルやカテゴリーの分類、特性調査

事例への適用

質問表現の相互行為的機能

- ・質問表現の基本的な機能は情報要求
- ・しかし、対話では参与者の相互行為を促進する機能を持つ
 - ・気まずい沈黙の回避、話題の転換 etc.
- ・対話例における質問表現の用例を記述的に分析：「対話において、質問表現を用いることで何かなされうるのか？」(増田 2005)

発話単位タグ (荒木他 1999)

質問表現

- ・確認

(Yes/No疑問; 回答内容について予測あり)

- ・真偽情報要求

(Yes/No疑問; 回答内容について予測なし)

- ・未知情報要求 (WH疑問)

談話計画タグ(1)

- 話題導出

0193 L: 大阪のお勧めはどこですか

談話計画タグ(2)

- 内容明確化

0275 L: じゃ本当の大阪を知るには その

0276 L: どこでしょう 心齋橋

0277 R: もうこてこては心齋橋難波の間
それからアメリカ村

0278 L: ふうん

0279 L: *何かよく聞きますよね それは一体
どこにあるんですか*

談話計画タグ(3)

- 妥当性確認 (「確認」)
0527 L: あじゃ ちょっと行ってもいいですか
0528 R: ちょっとうん
0529 R: 一回限りの人いっぱいいますから
0530 L: **本当ですか**

「本当ですか」「そうですか」に類する表現ばかりである。

談話計画タグ(4)

- まとめ・言い換え

(大阪・南堀江についての話)

0259 R: 急にお洒落になっちゃった

0260 L: 今じゃえらくお洒落 表参道みたいな

⋮

0264 R: 歩いてる人もお洒落ですし

0265 L: うーん

0266 R: うん 私一回しか行ったことないんですよ

0267 L: あ そう

0268 L: *じゃ あれが本当の大阪じゃないんですね*

談話計画タグ(5)

- 認識提示

0552 R: うん 楽しいですよ もう本当一日限りの体験でもお勧めですよ

0553 L: ふーん

0554 R: 遊びとしても

0555 L: *でもゴスペルって凄い大変じゃないですか 歌うの*

談話計画タグ(6)

- 話題提示(**本発表では省略**)

0294 L: *あいのり*という番組知ってます

談話計画タグ(7)

- 同意要求

0321 R: 東京もおもしろ 東京の方がでも規模が

0322 L: うーん

0323 R: 違いますもん

⋮

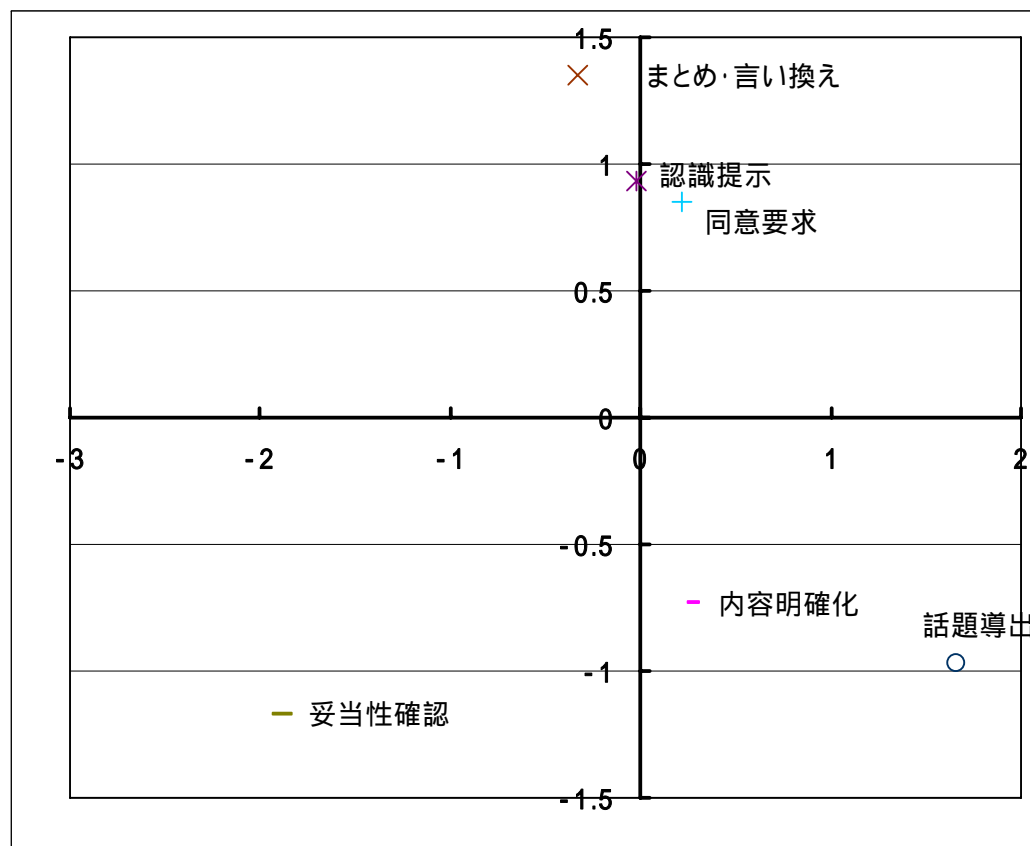
0326 L: *でもおんなじ顔してないですか東京の人 みんなおんなじ洋服着て*

提示される認識と先行発話との関連は弱く、聞き手に同意を求める意図が強く見て取れる。

実践

Black-Box (青木) を利用
Excelで散布図作成

カテゴリースコア(談話計画タグ)



大(自分の話題を導入)



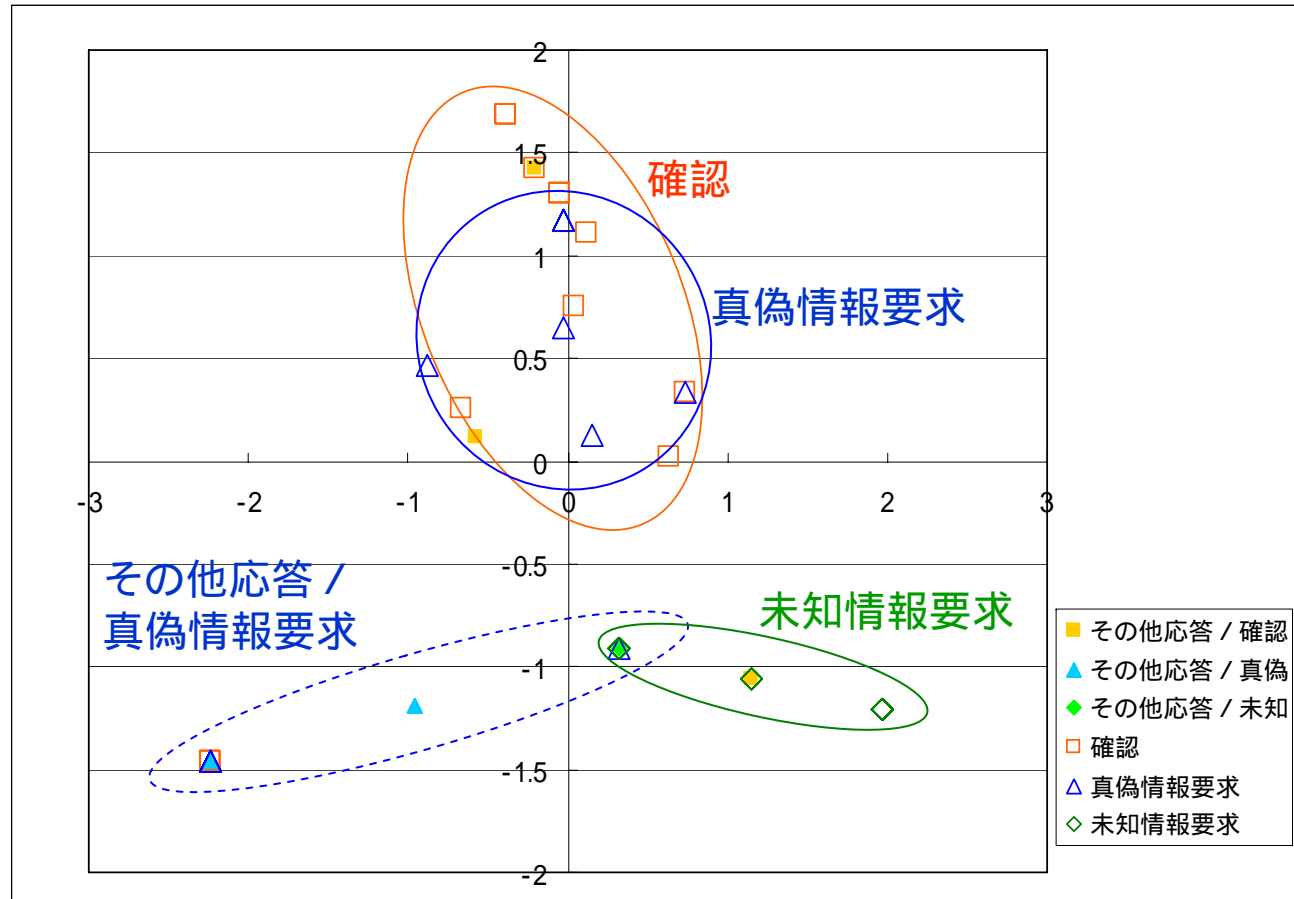
小(相手の話題を理解)
解)

旧

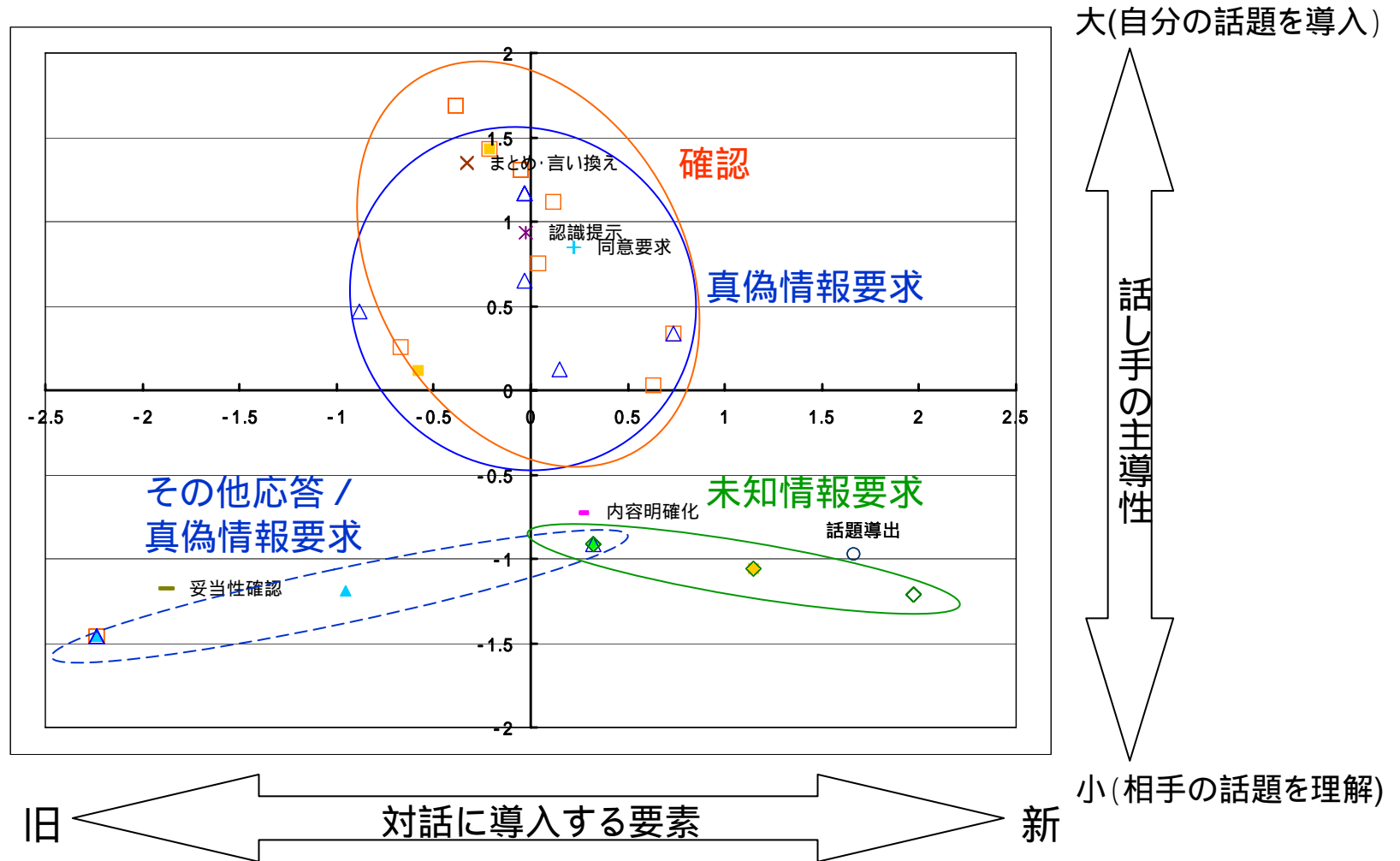


新

サンプルスコア (発話単位タグ)



両グラフの合成



分析結果 (1/2)

- 未知情報要求：
対話に新要素を導入し、話し手の主導性は小さい。

「話題導出」と「内容明確化」にしか用いられていない。

cf. 「話題導出」と「内容明確化」が「確認」「真偽情報要求」(計12例)において用いられているのは1例(その他の応答 / 確認)。

分析結果 (2/2)

- **確認・真偽情報要求:**
非常に類似した分布を示す。
対話に導入する要素の新旧については中立、話し手の主導性は大きい。

**未知情報要求 (話題導出・内容明確化) -
確認・真偽情報要求という二分化**

使用コーパス

国立国語研究所・情報通信研究機構. 日本語話し言葉コーパス. 文部科学省科学技術振興調整費開放的融合研究制度課題「話し言葉の言語的・パラ言語的構造の解明に基づく『話し言葉工学』の構築」, 1999-2003, 総括責任者: 古井貞熙, 2004.

参考文献

- 青木繁伸. おしゃべりな部屋. <http://aoki2.si.gunma-u.ac.jp/index.html>
- 荒木雅弘・伊藤敏彦・熊谷智子・石崎雅人. 1999. 「発話単位タグ標準化案の作成」. 人工知能学会誌. 第14巻第6号.
- 有馬哲・石村貞夫. 1987. 『多変量解析のはなし』. 東京図書.
- 原口正行のホームページ. <http://gucchi24.hp.infoseek.co.jp/index.htm>
- 増田将伸. 2005. 「対話における質問表現の機能に関する予備的検討」. 人工知能学会 第44回言語・音声理解と対話処理研究会.
- 大野高裕. 1998. 『多変量解析入門 自由自在に使いこなすコツ』. 同友館.
- 平知宏. 2004. 「数量化 類・対応分析」. 心理データ解析演習発表資料.

ご清聴ありがとうございました。