

『分類語彙表』の見出し語 95,071 語に対する単語親密度付与 知っている(KNOW), 書く(WRITE), 読む(READ), 話す(SPEAK), 聞く(LISTEN)の5観点を Yahoo クラウドソーシングにより収集し、ベイジアン線形混合モデルで推定

調査方法

分類語彙表 [国立国語研究所 2004]
語を意味によって分類・整理したシソーラス (類義語集)

分類番号の構造 例: よう (分類番号: 3.1300)

類	部門	中項目	分類項目
相 (3)	抽象的関係 (1)	様相 (13)	様相・情勢 (1300)

Yahoo! クラウドソーシングによる調査



2018年11月に実施
3,392人の実験協力者から
1見出し語あたり16人以上のデータ
(5観点)を収集
データポイント数 1,617,184件

5段階評価

- 1 全く知らない/全く出現しない
- 2 あまり知らない/あまり出現しない
- 3 どちらともいえない
- 4 何となく知っている/たまに出現する
- 5 よく知っている/よく出現する

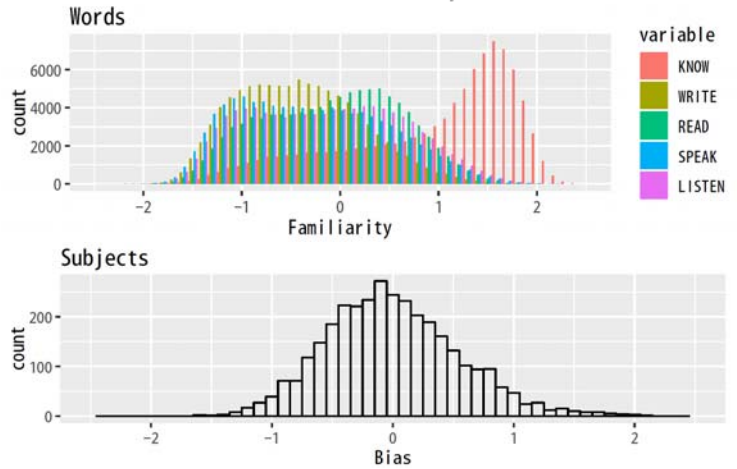
ベイジアン線形混合モデルによるモデル化

N_{word} は調査する単語 (と観点) の数 (= 100,830 × 5), N_{subj} は調査協力者の数 (= 3,392), $i: 1 \dots N_{word}$ が単語に対するインデックスで, $j: 1 \dots N_{subj}$ が調査協力者に対するインデックスである。 $y^{(i)(j)}$ は単語親密度 (KNOW, WRITE, READ, SPEAK, LISTEN) の値で, 次の正規分布としてモデル化する:

$$y^{(i)(j)} \sim Normal(\mu^{(i)(j)}, \sigma)$$

ここで σ は標準偏差である。平均 $\mu^{(i)(j)}$ は, 切片 α と調査協力者のランダムスロープ $\gamma_{word}^{(i)}$ と単語のランダムスロープ $\gamma_{subj}^{(j)}$ からなる次の線形式でモデル化する:

$$\mu^{(i)(j)} = \alpha + \gamma_{word}^{(i)} + \gamma_{subj}^{(j)}$$



推定結果 (各観点の上位 or 下位 10位まで)

知っている		知らない		書く		書かない		読む		読まない	
全員	うずみひ	一か月・箇月 (かげつ)	玉章 (たまざさ)	聞く	痛罵	聞く	玉章 (たまざさ)	聞く	玉章 (たまざさ)	読む	あおのけ
恋人	玉章 (たまざさ)	仕事する/ををする	御稜威 (みいつ)	話す	暴戻 (ぼうれい)	書く	暴戻 (ぼうれい)	書く	鞆掌 (おうしょう) する	登録する	撃撲
翌朝 (よくあさ)	御稜威 (みいつ)	話す	繞 (にょう)	見る	鞆掌 (おうしょう) する	自分	鞆掌 (おうしょう) する	登録する	鞆掌 (おうしょう) する	自分	席題
退社する	繞 (にょう)	見る	鞆掌 (おうしょう) する	自分	鞆掌 (おうしょう) する	見る	鞆掌 (おうしょう) する	自分	鞆掌 (おうしょう) する	見る	席題
再会	鞆掌 (おうしょう) する	力 (ちから)	スフ	口	鞆掌 (おうしょう) する	明日 (あした)	鞆掌 (おうしょう) する	明日 (あした)	鞆掌 (おうしょう) する	明日 (あした)	暖国
本社	鞆掌 (おうしょう) する	作る	馳名 (おいずり)	作る	鞆掌 (おうしょう) する	人 (ひと)	鞆掌 (おうしょう) する	人 (ひと)	鞆掌 (おうしょう) する	人 (ひと)	退京
入社	馳名 (おいずり)	食料品	宇内 (うだい)	食料品	あおのけ	終わる	鞆掌 (おうしょう) する	終わる	鞆掌 (おうしょう) する	終わる	御稜威 (みいつ)
人見知りする	宇内 (うだい)	今度	賢察	今度	みず垣	今度	鞆掌 (おうしょう) する	今度	鞆掌 (おうしょう) する	今度	劫を経る
持ち帰る	賢察	今度	今度	今度	今度	今度	鞆掌 (おうしょう) する	今度	鞆掌 (おうしょう) する	今度	馳名
ストロー	今度	今度	今度	今度	今度	今度	鞆掌 (おうしょう) する	今度	鞆掌 (おうしょう) する	今度	馳名

生産上位		受容上位		話す		話さない		聞く		聞かない	
聞く	聞く	聞く	聞く	聞く	みず垣	聞く	みず垣	聞く	玉章 (たまざさ)	聞く	玉章 (たまざさ)
終わる	終わる	終わる	終わる	終わる	玉章 (たまざさ)	終わる	玉章 (たまざさ)	終わる	劫を経る	終わる	劫を経る
一か月・箇月 (かげつ)	休む	一か月・箇月 (かげつ)	休む	野菜	痛罵	野菜	痛罵	野菜	あおのけ	野菜	あおのけ
今度	書く	今度	書く	今度	劫を経る	今度	劫を経る	食べ物	鞆掌 (おうしょう) する	今度	鞆掌 (おうしょう) する
今度	今度	今度	今度	母親	薫染	今度	薫染	会う	微賤	今度	微賤
休む	見る	休む	見る	食べ物	使臣	休む	使臣	休む	少時	休む	少時
話す	作る	話す	作る	会う	吏道	話す	吏道	書く	吏道	話す	吏道
見る	明日 (あした)	見る	明日 (あした)	年を取る	微賤	見る	微賤	今度	放念	見る	放念
自分	食べ物	自分	食べ物	書く	馬手 (めて)	自分	馬手 (めて)	今度	席題	自分	席題
書く	今日 (きょう)	書く	今日 (きょう)	今度	地映	書く	地映	寝る	鞆掌 (おうしょう) する	書く	鞆掌 (おうしょう) する

書記上位		音声上位		生産-受容上位		受容-生産上位		書記-音声上位		音声-書記上位	
聞く	聞く	聞く	聞く	毛管	送検する	聞く	送検する	上記	レジ袋	聞く	レジ袋
一か月・箇月 (かげつ)	終わる	一か月・箇月 (かげつ)	終わる	物心 (ぶっしん)	右翼	一か月・箇月 (かげつ)	右翼	追伸	先っちょ	一か月・箇月 (かげつ)	先っちょ
自分	野菜	自分	野菜	消却する	書類送検	自分	書類送検	前述する	ちょうまかす	自分	ちょうまかす
休む	今度	休む	今度	絆創膏	巡業する	休む	巡業する	後述	パイパイ	休む	パイパイ
作る	食べ物	作る	食べ物	ふたとせ	西郷隆盛	作る	西郷隆盛	記	ヨーグルト	作る	ヨーグルト
話す	会う	話す	会う	揚げなべ	殺害 (さつがい・せつがい)	話す	殺害 (さつがい・せつがい)	前略	ドライヤー	話す	ドライヤー
終わる	書く	終わる	書く	吟詠する	革命児	終わる	革命児	在中	まんま [その~]	終わる	まんま [その~]
今度	今度	今度	今度	だるい	護衛する	今度	護衛する	アンバサンド [&]	それではまた	今度	それではまた
私 (わたし)	休む	私 (わたし)	休む	上辺 (うわべ)	識者	私 (わたし)	識者	句読点	鼻水	私 (わたし)	鼻水
登録する	年を取る	登録する	年を取る	幽寂	再審	登録する	再審	下記	どっこいしょ	登録する	どっこいしょ

見出し語上位		見出し語下位		分類番号上位		分類番号下位		分類番号(中項目)上位		分類番号(中項目)下位			
知っている	知らない	知っている	知らない	知っている	知らない	知っている	知らない	知っている	知らない	知っている	知らない		
全員 2.44	うずみび -1.86	2.1530	用-関係-作用-出入り	1.60	1.3112	体-活動-音読-音讀	0.05	3.53	相-自然-生物	1.41	3.52	相-自然-天地	0.13
芝人 2.44	玉童(たまござ) -1.86	4.3160	他-接続-書き	1.60	3.5230	相-自然-天地-地	0.13	3.17	相-関係-空間	1.41	1.54	体-自然-植物	0.40
遊園(よくあそ) 2.44	御縁(みいつ) -1.85	2.3240	他-接続-演劇-映画	1.58	4.3320	他-挨拶-手紙用語など	0.14	2.10	用-関係-真偽	1.35	1.55	体-自然-動物	0.64
通社する 2.38	観(みよう) -1.85	4.1160	他-接続-補充	1.57	1.3104	体-活動-音読-叙述	0.19	3.56	相-自然-身体	1.34	1.31	体-活動-音讀	0.66
再訪する 2.38	観(おうしよう)する -1.84	2.1331	用-関係-事情-こそあど、他	1.57	1.5502	体-自然-動物-鳥類	0.21	2.56	用-関係-真偽	1.32	1.23	体-主体-人物	0.67
本社 2.38	スフ -1.82	2.1332	用-関係-様相-良不良、不適	1.56	1.4201	体-生産物-衣料-布、布地、織物	0.21	2.14	用-関係-力	1.32	1.42	体-生産物-衣料	0.68
入社 2.37	観名 -1.79	2.1010	用-関係-真偽-こそあど、他	1.55	1.3136	体-活動-音読-説明	0.23	3.35	相-活動-交わり	1.32	1.52	体-自然-天地	0.70
人見知りする 2.36	驚慌(おひざり) -1.79	4.3310	他-接続-挨拶	1.55	1.4161	体-生産物-資材-絹	0.24	4.32	他-呼び掛け	1.31	1.32	体-活動-芸術	0.71
持ち帰る 2.36	芋内(うだい) -1.76	2.5720	用-自然-生命-障害、けが	1.54	1.5401	体-自然-植物-木本	0.26	4.31	他-判断	1.29	4.50	他-動物の鳴き声	0.72
ストロー 2.36	賢察 -1.75	3.1643	相-関係-時間-未来	1.51	1.5402	体-自然-植物-草本	0.27	3.57	相-自然-生命	1.26	1.51	体-自然-物質	0.76
書く	書かない	書く	書かない	書く	書かない	書く	書かない	書く	書かない	書く	書かない		
一か月・月間(かげつ) 1.97	玉童(たまござ) -2.15	4.1160	他-接続-補充	0.65	1.2390	体-主体-人物-個人人名	-0.94	2.10	用-関係-真偽	0.16	3.52	相-自然-天地	-0.74
聞く 1.90	痛風 -2.03	2.1010	用-関係-真偽-こそあど、他	0.49	1.4161	体-主体-成員-農林水産業	-0.92	4.11	他-接続	0.15	1.55	体-自然-植物	-0.73
仕事する/をす 1.86	暴息(ぼうし) -1.97	1.1632	相-関係-時間-週、週日	0.43	1.3840	体-活動-事業-観望	-0.90	4.31	他-判断	0.11	1.54	体-自然-植物	-0.73
話す 1.85	観(おうしよう)する -1.97	1.1010	体-関係-事情-こそあど、他	0.36	1.4161	体-生産物-衣料-絹	-0.90	4.32	他-呼び掛け	0.09	1.23	体-主体-人物	-0.72
見る 1.82	観死(いし) -1.93	4.1130	他-接続-反対	0.33	1.3112	体-活動-音読-音讀	-0.90	3.10	相-関係-真偽	0.05	1.42	体-生産物-衣料	-0.65
自分 1.81	驚議 -1.92	3.1643	相-関係-時間-未来	0.27	1.5502	体-自然-動物-鳥類	-0.89	2.14	用-関係-力	0.04	1.52	体-生産物-衣料	-0.65
力(ちから) 1.81	痛風 -1.92	3.1641	相-関係-時間-現在	0.23	1.2420	体-主体-成員-軍人	-0.89	3.19	相-関係-量	0.01	1.32	体-活動-芸術	-0.64
口 1.80	観名 -1.92	4.3100	他-判断-判断	0.23	1.3821	体-活動-事業-印刷、製本	-0.89	3.16	相-関係-時間	-0.01	4.50	他-動物の鳴き声	-0.64
作る 1.80	あおのけ -1.92	3.5170	相-自然-物質-熱	0.22	1.4201	体-生産物-衣料-布、布地、織物	-0.84	3.17	相-関係-時間	-0.01	1.45	体-生産物-道具	-0.59
食料品 1.80	みずぼ 1.91	4.1140	他-接続-選択	0.20	1.3047	体-活動-心-信仰、宗教	-0.83	3.11	相-関係-類	-0.05	1.24	体-主体-成員	-0.58
読む	読まない	読む	読まない	読む	読まない	読む	読まない	読む	読まない	読む	読まない		
聞く 2.20	玉童(たまござ) -2.09	4.1160	他-接続-補充	0.87	1.3112	体-活動-音読-音讀	-0.71	4.11	他-接続	0.42	3.52	相-自然-天地	-0.52
読む 2.00	あおのけ -2.03	2.1010	用-関係-真偽-こそあど、他	0.73	1.5502	体-自然-動物-鳥類	-0.68	2.10	用-関係-真偽	0.40	1.54	相-自然-植物	-0.51
書く 2.00	観(おうしよう)する -1.96	1.1632	相-関係-時間-週、週日	0.64	1.3840	体-活動-事業-観望	-0.67	3.17	相-関係-空間	0.40	1.55	体-自然-動物	-0.47
今度 1.98	驚議 -1.92	2.1010	用-関係-真偽-こそあど、他	0.59	1.4161	体-主体-成員-農林水産業	-0.67	4.31	他-判断	0.38	4.50	他-動物の鳴き声	-0.45
自分 1.98	痛風 -1.90	4.1130	他-接続-反対	0.58	1.2390	体-主体-人物-個人人名	-0.63	3.10	相-関係-真偽	0.35	1.42	体-生産物-衣料	-0.39
見る 1.97	観死(いし) -1.89	3.1643	相-関係-時間-未来	0.55	1.4201	体-生産物-衣料-布、布地、織物	-0.63	2.14	用-関係-力	0.35	1.23	体-主体-人物	-0.35
明日(あした) 1.96	遠京 -1.89	3.1651	相-関係-時間-終始	0.51	1.4540	体-生産物-道具-農具など	-0.59	4.32	他-呼び掛け	0.33	1.52	体-自然-天地	-0.32
人(ひと) 1.94	御縁(みいつ) -1.86	4.3100	他-判断-判断	0.51	1.5402	体-自然-植物-草本	-0.59	3.35	相-活動-交わり	0.29	1.32	体-活動-芸術	-0.32
終わる 1.94	観(おうしよう)する -1.86	2.1332	用-関係-様相-良不良、不適	0.50	1.5401	体-自然-植物-木本	-0.59	3.19	相-関係-量	0.29	1.45	体-生産物-道具	-0.30
今度 1.93	観名 -1.86	4.1140	他-接続-選択	0.50	1.3821	体-活動-事業-印刷、製本	-0.58	3.16	相-関係-時間	0.28	1.31	体-活動-音讀	-0.28
聞く	話さない	話す	話さない	話す	話さない	話す	話さない	話す	話さない	話す	話さない		
聞く 2.20	みずぼ -2.16	2.1010	用-関係-真偽-こそあど、他	0.87	1.2390	体-主体-人物-個人人名	-0.93	4.32	他-呼び掛け	0.48	3.52	相-自然-天地	-0.72
終わる 2.18	玉童(たまござ) -2.16	4.1160	他-接続-補充	0.69	1.2420	体-主体-成員-軍人	-0.92	2.10	用-関係-真偽	0.38	1.23	体-主体-人物	-0.66
野菜 2.16	痛風 -2.04	1.1632	相-関係-時間-週、週日	0.66	1.3104	体-活動-音読-叙述	-0.92	4.31	他-判断	0.34	1.55	体-自然-動物	-0.64
今度 2.12	観(おうしよう)する -1.97	3.5170	相-自然-物質-熱	0.59	1.4161	体-生産物-資材-絹	-0.89	4.11	他-接続	0.26	1.54	体-自然-植物	-0.63
母親 2.09	驚議 -1.95	3.1915	用-関係-量-寒暖	0.59	1.3112	体-活動-音読-音讀	-0.88	2.14	用-関係-力	0.23	1.32	体-活動-芸術	-0.60
今度 2.07	痛風 -1.95	4.3210	他-呼び掛け-応答	0.54	1.2413	体-主体-成員-農林水産業	-0.87	3.19	相-関係-量	0.22	1.52	体-自然-天地	-0.60
自分 2.06	観死(いし) -1.95	1.1010	体-関係-事情-こそあど、他	0.53	1.5502	体-自然-動物-鳥類	-0.85	2.10	相-関係-真偽	0.20	1.42	体-生産物-衣料	-0.56
会う 2.06	遠京 -1.95	3.1643	相-関係-時間-未来	0.48	1.2320	体-主体-人物-君主	-0.83	3.17	相-関係-空間	0.19	1.31	体-活動-音讀	-0.55
年を取る 2.06	観議 -1.95	4.3160	他-判断-書き	0.45	1.3047	体-活動-心-信仰、宗教	-0.83	2.19	用-関係-量	0.17	1.24	体-主体-成員	-0.50
書く 2.05	芋内(うだい) -1.95	2.5170	用-自然-物質-熱	0.45	1.3840	体-活動-事業-観望	-0.83	3.51	相-自然-物質	0.15	1.27	体-主体-成員	-0.49
今度 2.04	地味 -1.93	4.3160	他-接続-書き	0.45	1.3840	体-活動-事業-観望	-0.83	3.51	相-自然-物質	0.15	1.27	体-主体-成員	-0.49
聞く	聞かない	聞く	聞かない	聞く	聞かない	聞く	聞かない	聞く	聞かない	聞く	聞かない		
聞く 2.20	玉童(たまござ) -2.09	2.1010	用-関係-真偽-こそあど、他	0.87	1.2420	体-主体-成員-軍人	-0.83	4.32	他-呼び掛け	0.54	3.52	相-自然-天地	-0.61
終わる 2.17	あおのけ -2.03	4.1160	他-接続-補充	0.75	1.3112	体-活動-音読-音讀	-0.82	4.31	他-判断	0.43	1.54	体-自然-植物	-0.55
野菜 2.17	あおのけ -2.03	1.1632	相-関係-時間-週、週日	0.73	1.3104	体-活動-音読-叙述	-0.81	2.10	用-関係-真偽	0.42	1.55	体-自然-動物	-0.54
食へ物 2.07	観(おうしよう)する -1.95	3.1915	用-関係-量-寒暖	0.67	1.2390	体-主体-人物-個人人名	-0.81	2.14	用-関係-力	0.35	1.23	体-主体-人物	-0.52
会う 2.07	観議 -1.95	3.5170	相-自然-物質-熱	0.63	1.5502	体-自然-動物-鳥類	-0.78	4.11	他-接続	0.35	1.32	体-活動-芸術	-0.48
読む 2.06	少時 -1.95	1.1010	体-関係-事情-こそあど、他	0.61	1.4161	体-生産物-資材-絹	-0.78	3.17	相-関係-空間	0.32	1.42	体-生産物-衣料	-0.46
書く 2.06	更迭 -1.95	4.3210	他-呼び掛け-応答	0.59	1.2413	体-主体-成員-農林水産業	-0.76	3.19	相-関係-量	0.31	1.52	体-自然-天地	-0.46
今度 2.05	驚議 -1.95	3.1643	相-関係-時間-未来	0.54	1.3047	体-活動-心-信仰、宗教	-0.74	3.10	相-関係-真偽	0.29	1.31	体-活動-芸術	-0.44
食へ物 2.05	痛風 -1.92	3.5050	相-自然-物質-熱	0.56	1.3840	体-活動-事業-観望	-0.73	2.19	相-関係-時間	0.25	1.45	体-生産物-道具	-0.38
寝る 2.04	観死(いし) -1.92	3.1915	用-関係-量-過不足	0.56	1.4201	体-生産物-衣料-布、布地、織物	-0.71	3.57	相-自然-生命	0.24	4.50	他-動物の鳴き声	-0.37
聞く	生産	生産	生産	生産	生産	生産	生産	生産	生産	生産	生産		
聞く 4.10	玉童(たまござ) -4.30	2.1010	用-関係-真偽-こそあど、他	1.36	1.2390	体-主体-人物-個人人名	-1.87	4.32	他-呼び掛け	0.57	3.52	相-自然-天地	-1.45
終わる 3.93	痛風 -4.07	4.1160	他-接続-補充	1.33	1.2420	体-主体-成員-軍人	-1.82	4.10	用-関係-真偽	0.54	1.23	体-主体-人物	-1.38
一か月・月間(かげつ) 3.88	みずぼ -4.07	1.1632	相-関係-時間-週、週日	1.09	1.2413	体-主体-成員-農林水産業	-1.79	2.10	用-関係-真偽	0.45	1.55	体-自然-動物	-1.37
今度 3.87	観(おうしよう)する -3.88	1.1010	体-関係-事情-こそあど、他	0.89	1.4161	体-生産物-資材-絹	-1.78	4.11	他-接続	0.41	1.54	体-自然-植物	-1.36
今度 3.79	地味 -3.85	1.3112	体-活動-音読-音讀	0.81	1.3112	体-活動-音読-音讀	-1.78	2.14	用-関係-力	0.27	1.52	体-自然-天地	-1.24
読む 3.76	観名 -3.84	4.1130	他-接続-反対	0.76	1.5502	体-自然-動物-鳥類	-1.74	3.10	相-関係-真偽	0.25	1.32	体-活動-芸術	-1.24
読む 3.75	驚議 -3.83	3.1643	相-関係-時間-未来	0.75	1.3840	体-活動-事業-観望	-1.72	3.19	相-関係-量	0.23	1.42	体-生産物-衣料	-1.24
見る 3.75	更迭 -3.83	3.1915	用-関係-量-寒暖	0.74	1.2320	体-主体-人物-君主	-1.66	3.17	相-関係-空間	0.18	1.31	体-活動-音讀	-1.13
自分 3.74	痛風 -3.83	4.3100	他-判断-判断	0.64	1.3047	体-活動-心-信仰、宗教	-1.66	3.16	相-関係-時間	0.13	1.24	体-主体-成員	-1.08
書く 3.70	暴息(ぼうし) -3.83	4.3210	他-呼び掛け-応答	0.64	1.4201	体-生産物-衣料-布、布地、織物	-1.65	2.19	用-関係-量	0.11	4.50	他-動物の鳴き声	-1.08
受容	受容	受容	受容	受容	受容	受容	受容	受容	受容	受容	受容		
聞く 4.40	玉童(たまござ) -4.18	4.1160	他-接続-補充	1.62	1.3112	体-活動-音読-音讀	-1.53	4.32	他-呼び掛け	0.87	3.52	相-自然-天地	-1.13
終わる 4.11	あおのけ -4.05	2.1010	用-関係-真偽-こそあど、他	1.60	1.5502	体-自然-動物-鳥類	-1.46	2.10	用-関係-真偽	0.82	1.54	体-自然-植物	-1.06
読む 4.06	観(おうしよう)する -3.92	1.1632	相-関係-時間-週、週日	1.37	1.2390	体-主体-人物-個人人名	-1.44	4.31	他-判断	0.81	1.55	体-自然-動物	-1.02
書く 4.06	地味 -3.89	1.1010	体-関係-事情-こそあど、他	1.19	1.4161	体-生産物-資材-絹	-1.43	4.11	他-接続	0.76	1.23	体-主体-人物	-0.87
今度 3.98	痛風 -3.83	4.1130	他-接続-反対	1.13	1.2420	体-主体-成員-軍人	-1.41	3.17	相-関係-空間	0.72	1.42	体-生産物-衣料	-0.86
見る 3.93	遠京 -3.79	3.1915	用-関係-量-寒暖	1.12	1.3840	体-活動-事業-観望	-1.40	2.14	用-関係-力	0.70	1.52	体-自然-天地	-0.81
作る 3.93	更迭 -3.77	3.1643	相-関係-時間-未来	1.12	1.4201	体-生産物-衣料-布、布地、織物	-1.34	3.10	相-関係-真偽	0.65	1.32	体-活動-芸術	-0.80