



2種のコーパスを対象とした 係り先動詞のエントロピーに基づく オノマトペの特徴分析



Hello!

内田ゆず（北海学園大学）です

オノマトペを語義の観点から
研究しています

詳しくはこちら

<http://ono-collo.com/>

1.

オノコロ

プロジェクト

私たちは何を目指しているのか

オノマトペの語義はあいまい

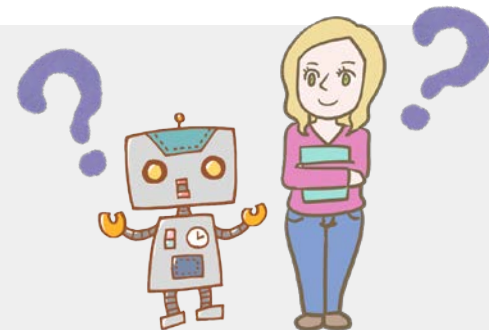
例えば…「ぶらぶら」を辞書で引くと

語義① 宙に釣り下がって揺れ動くさま

語義② あてもなく時間をかけて歩くさま

語義③ 何の目的もなく怠惰に暮らしているさま

…



どんな意味？

足をぶらぶらさせる

語義推定

どんな文脈？

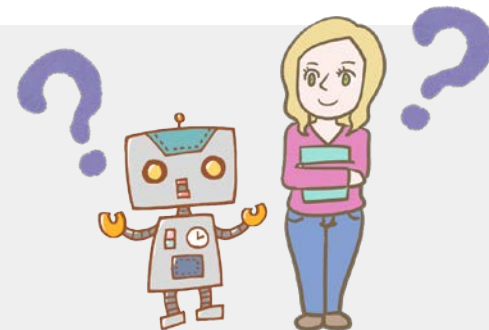
をぶらぶら

日本語学習者支援

オノマトペの品詞・使い方もあいまい

副詞

近所をぶらぶら歩く



サ変名詞

家でぶらぶらする

連体詞

ぶらぶらのイヤリング
ぶらぶらした腕

名詞

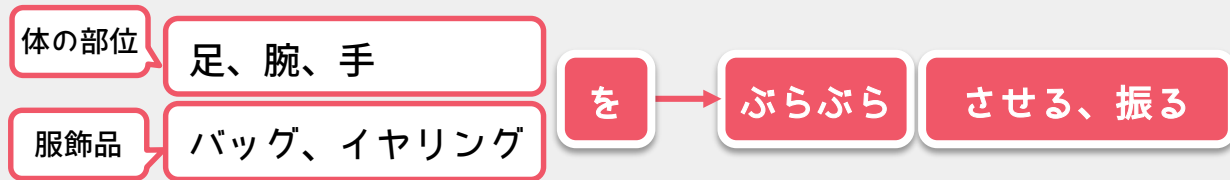
旅はぶらぶらが楽しい



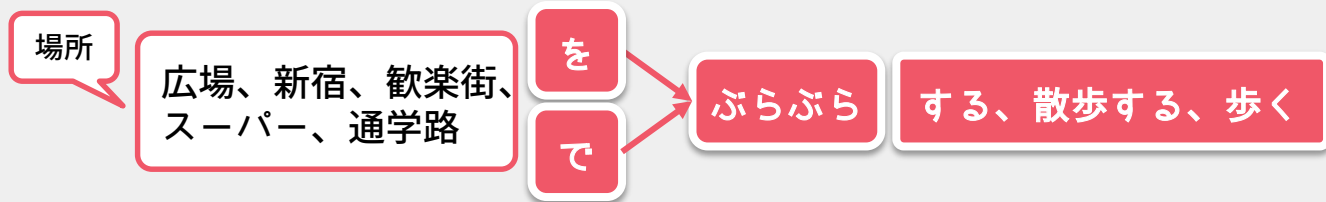
実例から語義とコロケーションを網羅したい

『文脈情報付きオノマトペデータベース』

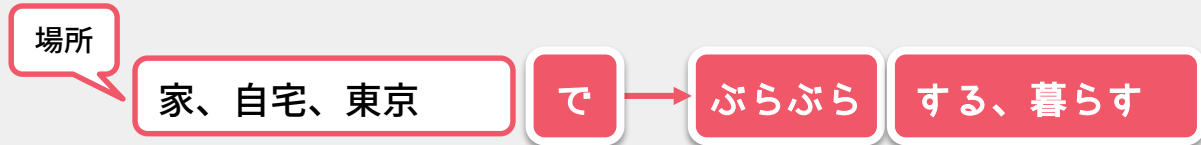
語義① 宙に釣り下がって揺れ動くさま



語義② あてもなく時間をかけて歩くさま



語義③ 何の目的もなく怠惰に暮らしているさま



2.

この発表の
あらすじ

どんな分析をするのか

オノマトペはいろいろなジャンルの文書に現れる



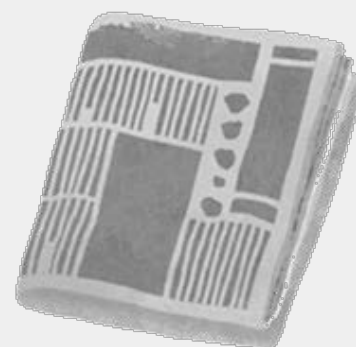
Webデータ



地方議会会議録



2種のコーパス中の
オノマトペを比較しよう



『2種のコーパス』ってどんなもの？



Webデータ

書き言葉が多め
ややカジュアル
多様なジャンル

55語のオノマトペを
検索ワードとして得た
11,186ページ



会議録

話し言葉
フォーマル
狭いジャンル

全都道府県議会会議録の
発言1,300万文



注目

係り先動詞の広がり

動詞が語義を決める第一のカギ

動詞に係るオノマトペってどのくらいあるの？



Webデータ

56,078例

1回以上：602語

100回以上：78語

オノマトペ-係り先TOP10		頻度
じっくり	する	623
がっかり	する	550
ぴったり	する	486
ぎりぎり	する	473
ちゃんと	する	382
あわわ	する	300
どんどん	聞く	286
じっくり	する	272
つつい	する	268
こっそり	する	262



会議録

10,482例

1回以上：273語

30回以上：107語

オノマトペ-係り先TOP10		頻度
ゆったり	する	125
いらいら	する	121
うろうろ	する	109
のんびり	する	108
ばたばた	する	108
ぎくしゃく	する	101
ずらっ	並ぶ	93
きらり	光る	85
ごちゃごちゃ	なる	77
ごっちゃ	なる	77

係り先動詞の「偏り具合」を比較しよう

エントロピーを指標とする

オノマトペ O が動詞 v に係る確率

$$H(O) = - \sum_{v \in V(O)} P(O, v) \log P(O, v)$$

オノマトペ O の
係り先動詞の集合

エントロピーの値が大きい＝幅広い動詞に係りやすい
エントロピーの値が小さい＝特定の動詞に係りやすい

3.

分析その1 手順と結果

係り先動詞の広がり？

分析の目的と手順

1. まずは全体の傾向をみたい

平均エントロピーを比較する

Webデータ 全602語／高頻度78語

会議録 全273語／高頻度107語

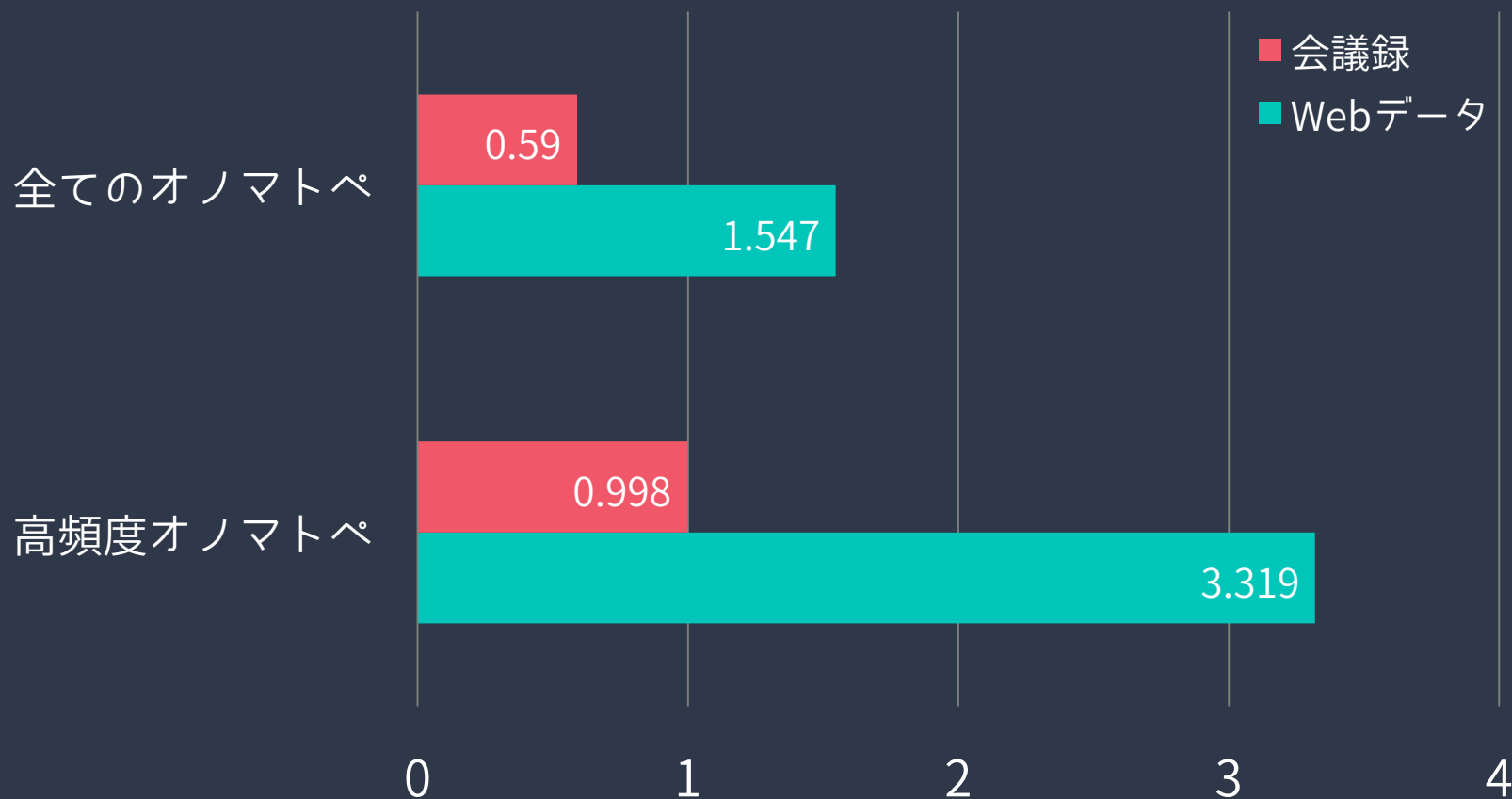
2. 具体的なオノマトペ-係り先動詞をみたい

高エントロピーなオノマトペを比較

低エントロピーなオノマトペを比較

エントロピーが0のオノマトペを比較

平均エントロピーの比較



Webデータ：幅広い文脈

会議録：決まった文脈

高エントロピーな高頻度オノマトペ



Webデータ

オノマトペ	H(O)	頻度	語義数
しっかり	4.58	697	4
ゆっくり	4.58	903	4
じっくり	4.54	2678	2
ぐっすり	4.46	613	4
じろじろ	4.37	633	1
がっかり	4.28	1732	2
ぼろぼろ	4.24	663	3
わくわく	4.11	280	3
きらきら	4.04	389	3
どンドン	3.98	1519	3



会議録

オノマトペ	H(O)	頻度	語義数
ばんばん	1.67	161	2
ちょこちょこ	1.65	236	4
ぽんぽん	1.59	88	4
ぼんぼん	1.59	97	3
どンドン	1.52	71	3
ぼつぼつ	1.46	76	5
じわじわ	1.45	85	5
ぴたっ	1.43	79	4
すんなり	1.43	186	2
たっぷり	1.42	75	3

動作そのものより動作を修飾するオノマトペ

低エントロピーな高頻度オノマトペ



Webデータ

オノマトペ	H(O)	頻度	語義数
ぎりぎり	1.43	614	4
さくさく	1.50	225	6
ふっくら	1.69	254	1
さっぱり	2.71	105	3
だんだん	2.73	155	2
しっとり	2.84	126	3
ぴったり	2.95	962	4
どきどき	3.03	160	1
わんわん	3.03	1,077	3
のんびり	3.04	315	1



会議録

オノマトペ	H(O)	頻度	語義数
やきもき	0.00	30	1
はらはら	0.13	35	6
もたもた	0.19	50	2
どきどき	0.21	79	1
ぎくしゃく	0.22	113	2
どきっ	0.25	44	1
いらいら	0.31	140	2
あたふた	0.32	34	1
うろうろ	0.37	131	3
ふらふら	0.42	88	5

連体詞的用法をとるオノマトペ、感情に関するオノマトペ

エントロピーが0のオノマトペ



Webデータ

オノマトペ	係り先	頻度
しんなり	する	21
きよとん	する	13
おどおど	する	10
もたもた	する	8
ぎくしゃく	する	6
やきもき	する	6
うようよ	する	5
ほっそり	する	4
きゃーきゃー	言う	3
ぴんぴん	する	3
ぽたり	落ちる	3
むしゃくしゃ	する	3
わなわな	震える	3
むかむか	する	2
めきめき	する	2



会議録

オノマトペ	係り先	頻度
やきもき	する	30
びっくり	する	24
がっかり	する	6
つくづく	思う	4
ひよこ	出る	2
ごっそり	持つ	2
ごそごそ	繰り入れる	2
じっ	とく	2

係り先動詞が
1種類しかない！



注目

「する」

たくさん登場しているぞ…!

4.

分析その2 手順と結果

「する」の影響は？

分析の目的と手順

3. サ変動詞「する」の影響をみたい

「する」の有無による平均エントロピーの差を算出

Webデータ 高頻度78語

会議録 高頻度107語

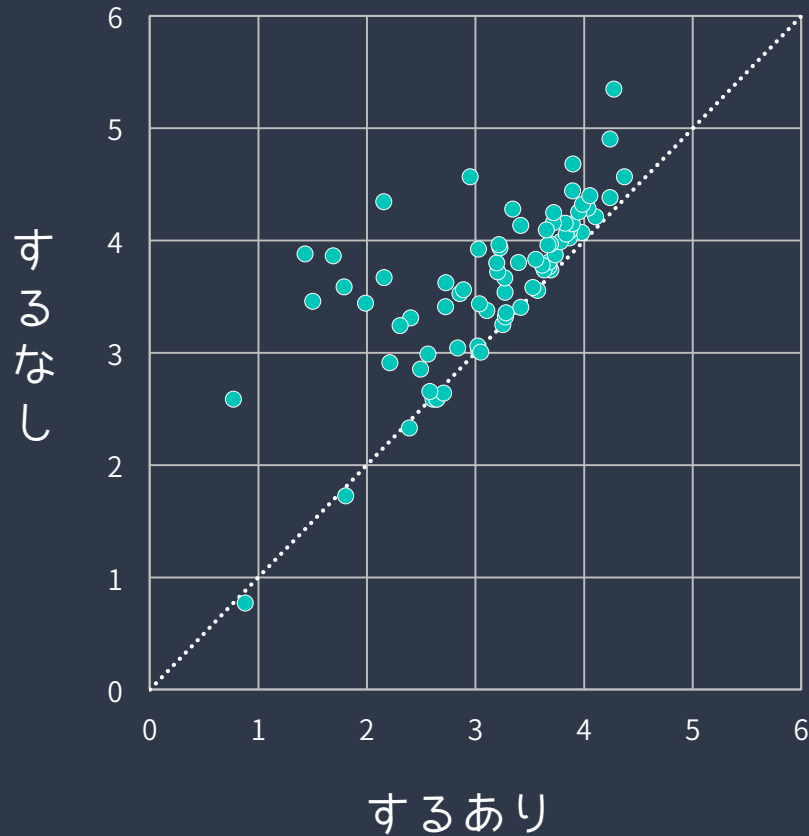
4. 具体的なオノマトペ-係り先動詞をみたい

「する」の影響が強いオノマトペを観察

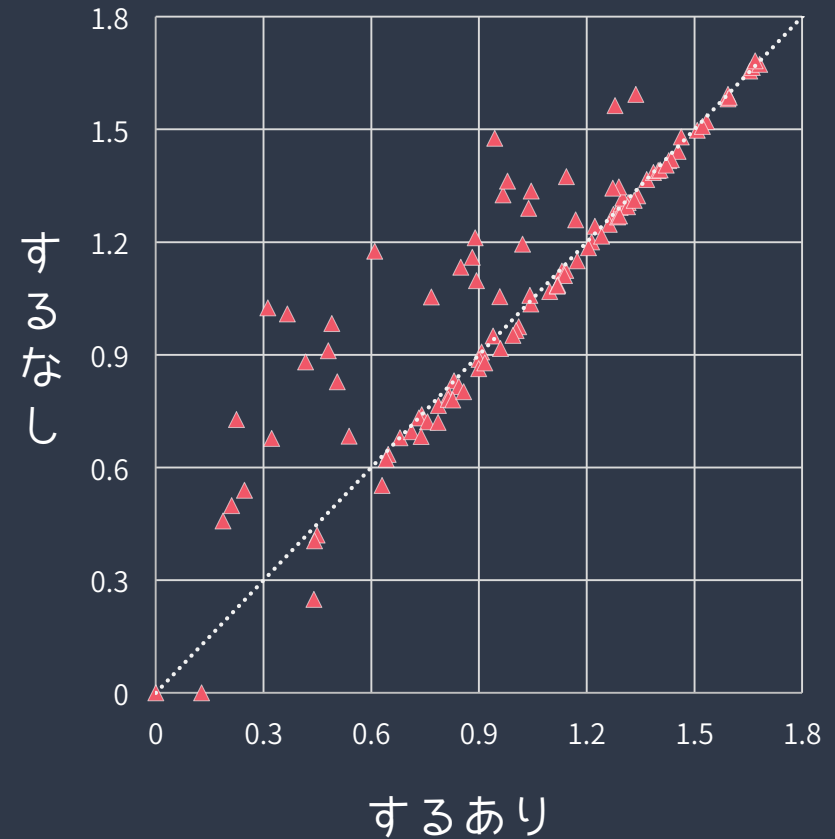
「する」の影響が弱いオノマトペを観察

「する」の有無で平均エントロピーを比較する

Webデータ



会議録



「する」の影響が大きいオノマトペ



Webデータ

オノマトペ	$\Delta H(O)$	頻度 (する)	語義数
うんざり	2.45	605 (180)	1
がっかり	2.19	1732 (550)	2
いらいら	2.17	146 (36)	2
ドキドキ	1.96	160 (64)	1
びっくり	1.79	443 (88)	2
わくわく	1.61	280 (29)	3
うきうき	1.51	123 (9)	1
きょろきょろ	1.46	211 (26)	1
しっかり	1.07	697 (129)	4
ごろごろ	0.94	220 (129)	4



会議録

オノマトペ	$\Delta H(O)$	頻度 (する)	語義数
いらいら	0.71	140 (121)	2
うろうろ	0.64	131 (109)	3
ぼんやり	0.57	51 (36)	4
ゆったり	0.53	222 (125)	2
ぎくしゃく	0.50	113 (101)	2
もやもや	0.49	72 (54)	5
ふらふら	0.46	88 (69)	5
ほのぼの	0.43	43 (32)	3
ばたばた	0.38	215 (108)	7
のんびり	0.36	217 (108)	1

サ変動詞になるものが多い＝述語だから重要

「する」の影響が小さいオノマトペ



Webデータ

オノマトペ	$\Delta H(O)$	頻度 (する)	語義数
ぼろぼろ	0.013	663 (75)	3
どっぴり	0.015	517 (84)	2
すやすや	0.023	150 (17)	2
ぐびぐび	0.036	110 (27)	4
すっきり	0.042	366 (111)	4
ざくざく	0.043	374 (96)	3
とっぴり	0.045	155 (31)	2
げらげら	0.057	106 (18)	2
あわわ	0.061	553 (300)	1
しとしと	0.075	222 (32)	4



会議録

オノマトペ	$\Delta H(O)$	頻度 (する)	語義数
やきもき	0.000	30 (30)	1
どっぴり	0.000	43 (0)	2
まざまざ	0.000	55 (0)	3
がぼっ	0.000	32 (0)	2
すいすい	0.000	32 (0)	2
ぶつぶつ	0.000	35 (0)	6
がたっ	0.000	32 (0)	2
ぽんぽん	0.000	88 (0)	4
はっきり	0.000	112 (49)	4
ちょちょ	0.002	236 (15)	4

動作を修飾するものが多い=省略されても大丈夫かも

5.

この発表の まとめ

何がわかったのか

オノマトペの体系化に向けて

Webと会議録でエントロピーを比較

Web : 雑多な内容⇒係り先動詞の広がり大きい


会議録 : ドメインが限定されている⇒係り先動詞に偏り

特徴的なオノマトペはジャンルによって異なる

今後の予定

いろいろな言語資源から実例を集める必要がある

各種コーパスからオノマトペ+コロケーションの抽出&語義分類

An aerial photograph of a small village nestled in a valley surrounded by lush green mountains. The village features a mix of traditional and modern buildings, with a prominent white building that appears to be a temple or shrine. The surrounding landscape is densely forested, and the sky is bright with scattered clouds.

質問・コメントを
お願いします
<http://ono-collo.com/>

山寺へ行って足がガクガク