



帝塚山大学
TEZUKAYAMA UNIVERSITY

分科会3(齋藤ひろみ科研)

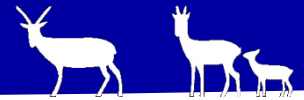
出来事作文における動詞のバリエーションとコロケーション

多様な言語文化背景を持つ子どもたちの
リテラシーフォーラム3

「子どもたちの日本語の発達
を可視化する—語彙・文法
の力に焦点を当てて—」
(於:お茶の水女子大学)
2016.02.28(日)

TEZUKAYAMA UNIVERSITY
50th
Anniversary
since 1964

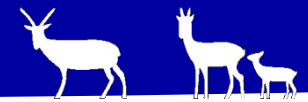
森篤嗣(帝塚山大学)



はじめに

- 本発表の構成は以下の通りです
 - 7年分の全作文の概要を確認する
 - 延べ語数と異なり語数について確認する
 - 動詞に注目して和語動詞とサ変動詞の比率などをみる
 - サ変動詞のバリエーションをみる
 - 「語彙密度」から書き言葉らしさについて分析する
 - サ変動詞のコロケーションパターンをみる





分析対象（作文数）

- 2008～2014年に収集した2～6年生までの7年間分の遠足についての作文（出来事作文）

	外国人児童	日本人児童	計
2008年度	82	73	155
2009年度	97	59	156
2010年度	81	33	114
2011年度	107	44	151
2012年度	103	40	143
2013年度	83	33	116
2014年度	119	122	241
計	672	404	1,076

－外国人児童は80%以上が日本生育





延べ語数と異なり語数

- UniDic2.1.2とMeCab0.996を使用
 - 品詞に「空白」「補助記号」「記号」の文字列を含むものを削除

	外国人児童	外国人児童(調整)	日本人児童
延べ語数	147,379	87,561	87,315
異なり語数	4,234	3,365	3,216
1作文あたりの延べ語数	219.31	—	216.13
TTR (type-token ratio)	0.0287	0.0384	0.0368

- 1作文あたりの分量には、ほとんど差がない
- 分量を調整した場合、TTRもほとんど差がない





和語動詞とサ変動詞

- 今回は**動詞**に注目（括弧内は異なり語数）

	外国人児童	日本人児童
和語動詞	10,095 (617)	6,000 (506)
サ変動詞	501 (97)	274 (69)
サ変／全動詞	4.73%	4.37%

- 和語動詞については「動詞-一般」のみを集計し、「動詞-非自立可能（「する」「いる」「いく」「なる」など）」は含めていない
- サ変動詞については、「名詞-普通名詞-サ変可能」に「する」が後接したものを抽出





和語動詞とサ変動詞

- 今回は動詞に注目

	外国人児童	日本人児童
和語動詞のTTR	$617 \div 10,095 = 0.0611$	$506 \div 6,000 = 0.0843$
サ変動詞のTTR	$97 \div 501 = 0.1936$	$69 \div 274 = 0.2518$

- 和語動詞もサ変動詞もTTRすなわちバリエーションの豊富さに差がある
- 差が大きいサ変動詞に注目してみる





サ変動詞上位10語＋1

相関
0.891

	外国人児童		日本人児童	
びっくりする	94	6.38	36	4.12
抱っこする	48	3.26	19	2.18
交換する	37	2.51	34	3.89
協力する	37	2.51	33	3.78
出発する	25	1.70	9	1.03
正解する	23	1.56	8	0.92
休憩する	15	1.02	5	0.57
移動する	12	0.81	8	0.92
注意する	11	0.75	4	0.46
集合する	9	0.61	2	0.23
行動する	4	0.27	10	1.15

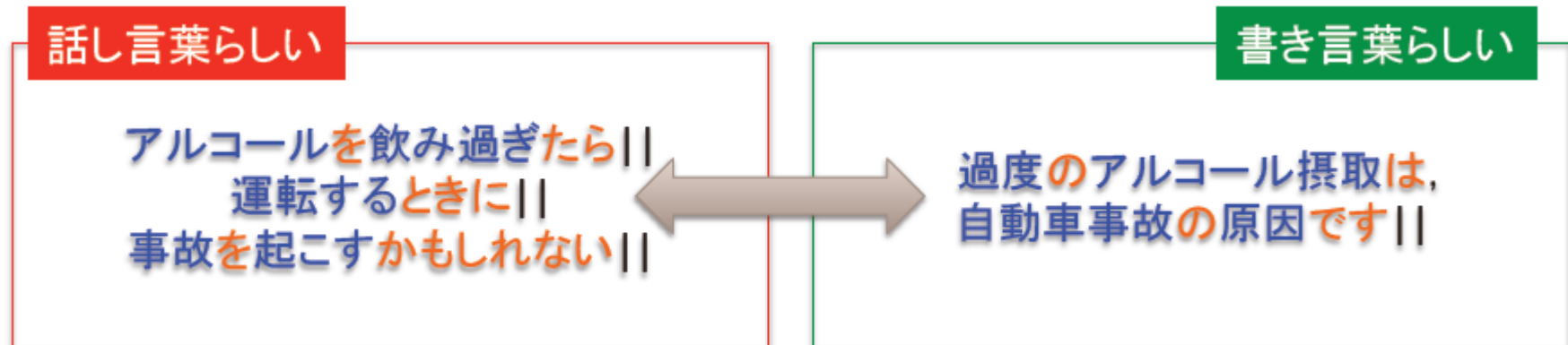
1万語あ
たり出現
数





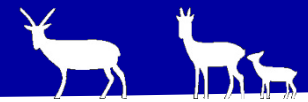
語彙密度

- 書き言葉らしさ, 話し言葉らしさ (Halliday 1985)
 - 書き言葉らしさ = 情報の詰め込みの程度が高い
 - 話し言葉らしさ = 文法的には複雑になり得るが, 情報詰め込みの程度は低い



佐野大樹 2014 「語彙密度から見た語彙シラバス構築に向けて」 国立国語研究所共同研究プロジェクト「学習者コーパスから見た日本語習得の難易度に基づく語彙・文法シラバスの構築」 研究発表会 (於: 国立国語研究所) 2014年6月28日より





語彙密度

$$\text{語彙密度} = \frac{\text{テキストに含まれる内容語の総数}}{\text{テキストに含まれる節 (ranking clause) の総数}}$$

平均でどのぐらいの内容語が1つの節に埋め込まれるかで、情報の詰め込みの程度を表す

語彙密度:低

アルコールを飲み過ぎたら||
運転するときに||
事故を起こすかもしれない||

語彙密度(1.67)=内容語数(5)/節数(3)

語彙密度:高

過度(1)のアルコール(2)摂取(3)は、
自動車(4)事故(5)の原因(6)です||

語彙密度(6)=内容語数(6)/節数(1)

佐野大樹2014「語彙密度から見た語彙シラバス構築に向けて」国立国語研究所共同研究プロジェクト「学習者コーパスから見た日本語習得の難易度に基づく語彙・文法シラバスの構築」研究発表会(於:国立国語研究所)2014年6月28日より



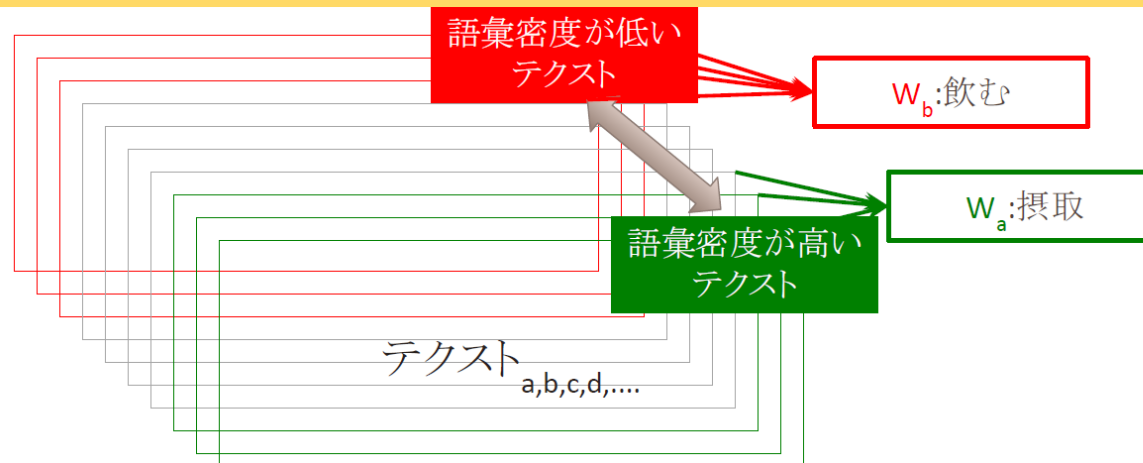


語彙密度平均値

語彙密度と語彙の“重み付け”

テキストの性質をあらわす語彙密度を使って、どう語彙の“重み付け”をする？

$Word_a$ が $Word_b$ に比べて情報の埋め込みの程度が高い(書き言葉らしい)テキストで利用される傾向があるのであれば、 W_a が使用されるテキストの語彙密度の平均値は W_b のそれに比べて高くなる。であれば、語彙密度平均値は、語の書き言葉らしさを“重み付け”できるのでは？



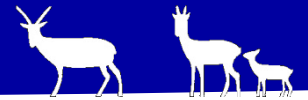
佐野大樹2015「語彙密度から見た語彙シラバス」国立国語研究所共同研究プロジェクト「シラバス作成を科学にする—日本語教育に役立つ多面的な語彙シラバスの作成—」研究発表会(於: 国立国語研究所) 2015年2月22日より





サ変動詞上位10語+1の語彙密度平均値

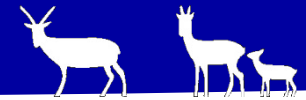
	外国人児童		日本人児童		語彙密度
びっくりする	94	6.38	36	4.12	—
抱っこする	48	3.26	19	2.18	2.97
交換する	37	2.51	34	3.89	4.83
協力する	37	2.51	33	3.78	4.95
出発する	25	1.70	9	1.03	4.30
正解する	23	1.56	8	0.92	3.66
休憩する	15	1.02	5	0.57	4.25
移動する	12	0.81	8	0.92	4.42
注意する	11	0.75	4	0.46	4.34
集合する	9	0.61	2	0.23	5.29
行動する	4	0.27	10	1.15	4.57



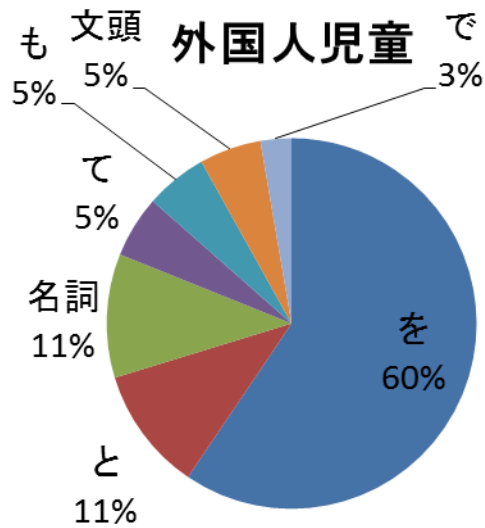
サ変動詞全体の語彙密度

- 各語の1万語あたり出現数 × 語彙密度平均値を求める
 - 外国人児童：異なり97語（延べ501語）
 - 日本人児童：異なり69語（延べ274語）
- 上記の平均値は・・・
 - 外国人児童1.44
 - 日本人児童1.89
- 全体としても差が見られる



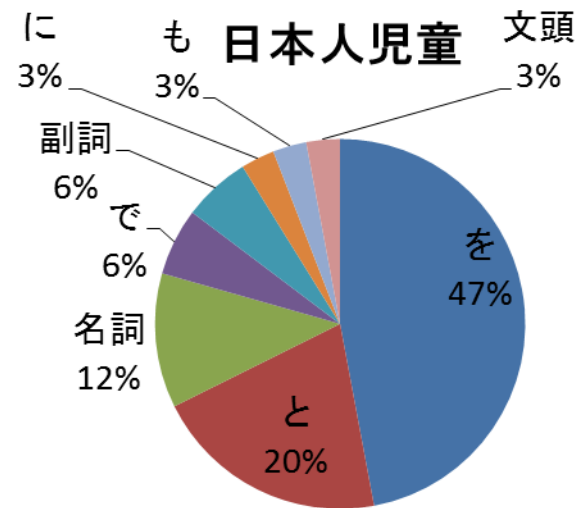


「交換する」のコロケーション



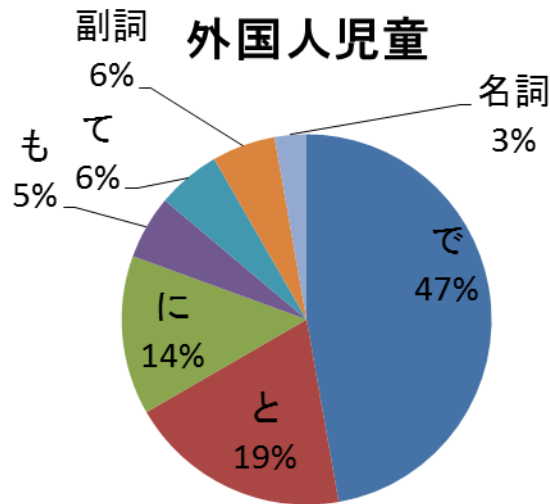
- ・3年生と場所を交かんして行きました。(を)
- ・私は、グミとこうかんしました。(と)
- ・みんなでこうかんして食べました。(で)
- ・おかずちょっと交かんしよう。(副詞)

- ・ウエンちゃんとおやつを、こうかんしました。(を)
- ・おやつは、友達とこうかんしたりしました。(と)
- ・ねえおかしこうかんしようよ。(名詞)

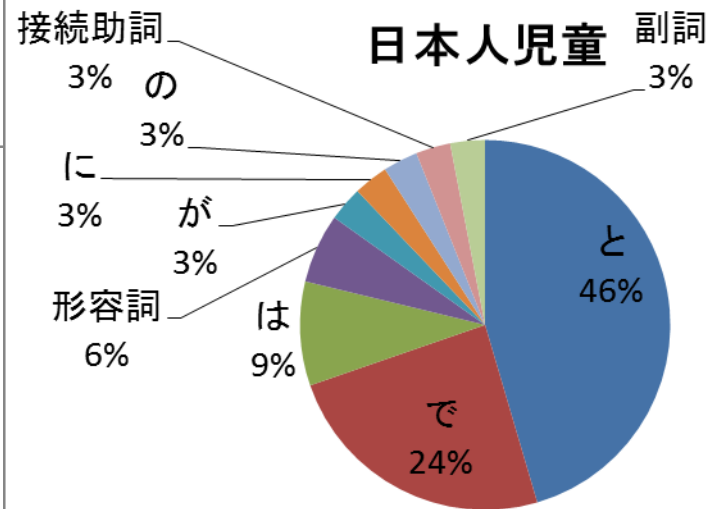




「協力する」のコロケーション



- ・はんのみんなと協力して作ります。(と)
- ・グループで協力してスタンプをぜんぶ集めたからです。(で)
- ・けど班のみんなが協力して分かりました。(が)



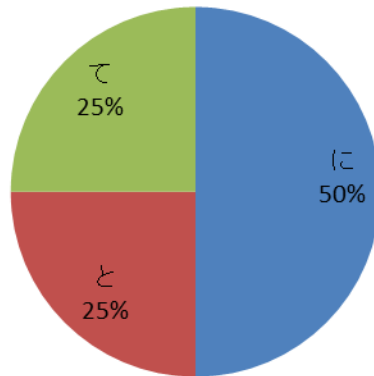
- ・みんなで協力して...(で)
「で」のうち「みんなで」が15/17(日本人も6/8)
- ・みんなと協力して...(と)
「と」のうち「みんなと」が3/7(日本人は7/15)





「行動する」のアンケート

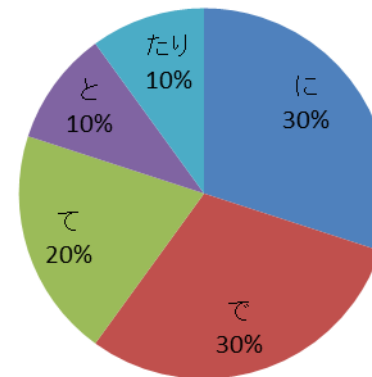
外国人児童

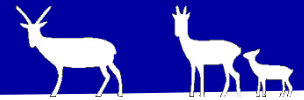


- ・一年生といっしょに、行動して、私は...(に)
- ・はんとか色で四、五人ぐらいで行動していたけれど...(で)
- ・遠足中ほとんど手をはなさないようにして行動しました。(て)

- ・三年生といっしょに行動しました。(に)
- ・8班と会ったので8班と行動して...(と)
- ・三年生とペアになって行動するときに...(て)

日本人児童





まとめ1

- 本科研の作文データは縦断調査なので「変化」を見ることが本来の目的
- しかし、全データを横断的に検索することにより「量」が確保できるメリットもある。
- 1作文あたりの分量も、TTR(語彙の豊富さ)もほとんど差はない
- サ変動詞の使用率は外国人児童の方がほんの少し高い
- 一方でTTR(バリエーション)で差が見られる





まとめ2

- 特にサ変動詞の使用において、「書き言葉らしさ(かたい表現)」で差が見られる
- 本発表では語彙密度平均値から「交換する」「協力する」「行動する」に着目(これらは「社会性の発達」にも関わる語群)
- コロケーションの観点からもバリエーションに若干の差が見られる
- 外国人児童は「みんなで協力」など特定のチャンクを多用する傾向ありか？





参考文献

- 齋藤ひろみ・森篤嗣・北澤尚・菅原雅枝・畠田陽子・工藤聖子・阿部志野歩2014「第32回研究大会ワークショップ 日本成育外国人児童のリテラシー発達を追う—作文縦断調査の多面的分析—」『社会言語科学』16(2), 90-98
- 佐野大樹(2016予定)「語彙密度から見た語彙シラバス」山内博之(監修)現場に役立つ日本語教育研究2, 森篤嗣(編)『ニーズを踏まえた語彙シラバス』くろしお出版

