

転移修飾を 構文交替として特徴づける

Characterizing the transferred epithet as as an alternation

黒田 航

京都工芸繊維大学(非常勤)

早稲田大学総合研究機構 情報教育研究所 (客員研究員)

第166会メビウス研究会

2011/02/19, 京都外国語大学

背景

- 世話人の渋谷くんから“量的的言語学研究的講演を”と依頼
 - 白状すると、この依頼には少し困った。
- 理由は
 - 京大のY梨研究室の院生の指導ではそういうことをしてきたけれど、私自身はいわゆる「量的言語学研究」したことがない
- “さて、どうしよう?”としばらく考えた結果 ...
 - 研究会の聴衆の関心のあり方を考えると、(日本語よりは)英語の研究がいいだろう

背景

- 過去7年のNICTでの勤務の経験で、大規模データを利用した情報処理 (別名 Super Crunching (SC: 絶対計算)) の威力をまざまざと見てきた
 - Super Crunching の精神とは
 - 大規模データ処理は専門家の直観を越える
 - ということ
 - SC については エアーズ (2007) を参照
 - 言語学者には些か耳が痛い話
- それを言語研究に活かす方法を模索する

発表の概要

- 素朴な実例主義を超えて: データ駆動言語学の本当の目的とは何か?
 - コーパス事例の利用と作例の利用は排他的ではない
- 転移修飾構文 (transferred epithet construction: TEC) の特徴づけ
 - 先行研究の整理
- 他の構文との連関
 - 支援動詞構文 (Support Verb Construction: SVC) との関連
 - 同族目的語構文 (Cognate Object Construction: COC) との関連
 - 同族目的語構文モドキ (Pseudo-Cognate Object Construction: PCOC) との関連
 - TE, SVC, (P)COC の関係

素朴な実例主義を超えて 1/6

- 論点

- 伝統的なコーパス言語学は“正例” (positive evidence) しか扱わない
- 伝統的なコーパス言語学は“負例” (negative evidence) を扱わない

- 問題点

- “実例” (instances, examples) と “正例” は概念的に別のもの
 - “正例” は “負例” の逆概念
- コーパスは実例 (の部分) 集合であるが, 正例 (の部分) 集合とは限らない
 - たいていのコーパスは “誤用例” を数多く含む

素朴な実例主義を超えて 2/6

- 問題

- 伝統的なコーパス言語学は暗黙の内にすべての実例を正例として扱う
- つまり実例集合 (observables) と正例集合 (positive examples) を同一視している
 - 詳しくは黒田 (in press)を
- “負例をどう扱うのか?” という理論的問題がある
 - コーパスに現われていないものがすべて負例なら、コーパスの規模で負例集合が変わる。これはコーパスが十分に大きくない場合には理不尽
 - データをどんなに増やしても、すべての正例が得られるわけではない

素朴な実例主義を超えて 3/6

- 負例は存在しないのか？
 - “負例は存在しない” という立場を採るのは、単なるナンセンス
 - 言語学の一派 (e.g., 認知言語学) は負例は存在しないと主張するが、そのような言語学は実用性を失う

素朴な実例主義を超えて 4/6

- データ駆動言語学にとっての真の課題
 - 言語 L の正例集合 P (の部分集合 $p \subset P$) のみを与えられた状態で、どう負例集合 N (かその部分集合 $n \subset N$) を特定するか?
- データ駆動言語学は伝統的なコーパス言語学と同一でなくてもよい

素朴な実例主義を超えて 5/6

- 前提
 - 負例の“実例”をコーパスに探すことは原理的に不可能
- 対応策
 - コーパス C 中の実例 $P = \{p_1, p_2, \dots, p_n\}$ を下に一定の規則 $R = \{r_1, r_2, \dots, r_n\}$ で生成した仮想例を $Q = \{q_1, q_2, \dots, q_n\}$ とする
 - Q の要素 q_i ($i = 1, 2, \dots, n$) が C 中で期待事例数 $E[q_i]$ より有意に少ない実例数しかない場合, q_i を(暫定的に) 負例と見なす
- 未解決の課題
 - どうやって C 中の $E[q_i]$ を求めるか?

素朴な実例主義を超えて 6/6

- よい知らせ
 - 実用的には $\mathbb{E}[q_i]$ の定義が不十分でも深刻ではない
 - もちろん, 将来的には解決する必要がある
- 理由
 - \mathbb{R} で生成された候補 $Q = \{q_1, q_2, \dots, q_n\}$ のほとんどは実例数=0
 - 少なくとも後述の転移修飾の case study ではそう
 - 期待実例数の条件は, $\mathbb{E}[q_i] > 1$ で実用上は十分だろう
 - 精度を上げたいなら, $\mathbb{E}[q_i] > k$ として k を現実的な数値にすればよい

Case Study

- 課題: コーパスに基づいて転移修飾表現 (1) が容認可能か判定せよ
 - (1) He was now smoking a sad cigarette.
- 直接法による判定
 - sad cigarette の類例数 N_0 を調べ, その数によって(1)の容認度を判定する
- 間接法による判定
 - he is sad の類例数 N_1 と he smoked a cigarette sadly の類例数 N_2 を調べ
 - N_1 , と N_2 から (1) の容認度を判定する
 - これは後述の神澤 (2011) の手法と同じ

転移修飾(構文)とは?

- 転移修飾 (transferred epithet: TE) は Hall (1973) で取り上げられ、言語学的に分析された修辞技法の一つで、次の例が有名:

(1) He was now smoking a **sad** cigarette and waiting for the blow to fail.

(2) I balanced a **thoughtful** lump of sugar on the teaspoon.

- Hall (1973) の挙げた例でいずれも P. G. Wodehouse 作

(3) The ploughman homeward plods his **weary** way. —Thomas Gray

- explained at http://www.answerbag.com/q_view/192567

転移修飾の特徴

- Hall (1973) が述べる TE の認定条件は次の二つ
 - S V (P) A O X で
 - C1: [Transferred epithets], in their turn, involve a reference to the subject of the verb in each instance.
 - C2: [W]e might interpret the adjective, in the construction of Adjective+Noun [= A O], as equivalent to an adverb transferred from its position modifying the verb of a clause.
- C1はヒト主語指向条件 (解釈条件), C2は構文交替条件
 - C1とC2の条件は独立で作用が異なる

先行研究の難点

- Hall (1973) に続く先行研究の大半はC2を無視している
 - Shibuya (2008) は 重要な例外
- つまり Loose な条件で TEC を認定し分析, 説明している:
 - Strict: C1 とC2を同時に満足する A N 構文
 - Loose: C2のみを満足する A N 構文
- 問題点
 - 本来とは別の現象が転移修飾として分析, 説明されている(可能性がある)
- 詳細は神澤 (2011)を参照

本研究の方向性

- TECは C1 と C2の両方が真である場合に限ると想定し
 - C1とC2を独立に調査する
- その一方で
 - C1とC2の成立条件の相関の有無を調べる
 - 最終的に、支援動詞構文，同族目的語構文，同族目的語構文モドキとの対比から，C2がC1に無関係に起こり得ることを指摘

類例検索

- (1) の類例検索は
 - 品詞タグつきコーパス (BNC) があり,
 - その上でターンマッチ言語 (e.g., Corpus Query Language (CQL)) が使用できるのであれば
- 簡単に実装できる
 - CQL の詳細は
 - <https://trac.sketchengine.co.uk/wiki/SkE/CorpusQuerying>
 - 投野由紀夫 (東京外国語大学) の解説も便利
 - <http://lexicon.tufs.ac.jp/~tono/pdf/Jaecs35WorkshopTono.pdf>

直接法 1/3

- CQL
 - [lemma="sad"] [tag="CJ.* | AJ.* | N.']* [lemma="cigarette"]
- の実例数=0
 - 因みに CQL
 - [lemma="cigarette"]
 - の実例数=3332

直接法 2/3

- smoked a sad cigarette が容認可能かどうかは BNC を調べてもわからない
- 因みに Sketch Engine の enTenTen での CQL
 - [lemma="sad"] [tag="C.*|A.*|N.*"]* [lemma="cigarette"]
 - enTenTen は PennTreeBank (PTB) のタグが前提で BNC の query は流用不可
 - の実例数=0
- enTenTen (3,268,798,627 tokens) の規模は BNC (112,181,015 tokens) の30倍

直接法 2/3

- Google で “a sad cigarette” の該当数 1,000ほど
(2011/2/18)
- (1)が実例であることを保証するには非常に規模の大きなデータが必要
- というわけで、コーパスの規模が十分に大きくない場合には直接法の有効性には限界がある
- より厄介な点はどれだけの token数があれば十分に大きなコーパスと言えるのかわからないという点

発展的問題 1/2

- コーパスが十分に大きいという保証がない場合に、誰がどうやって (1) が (少なくとも一部の人にとって) 容認可能な表現だと保証するのか?
- 実例しか利用しない (つまり作例を極力用いない) 伝統的なコーパス言語学が取り得る方法は次:
 - cigarette と 文脈 sad __ に生起する名詞 (e.g., reflection, tale) との類似度を計算し,
 - 類似度が十分に高い場合には sad cigarette も容認可能だと考える

sad *(adjective)* British National Corpus freq = 3321

<u>and/or</u>	<u>644</u>	<u>2.0</u>
novi	<u>8</u>	8.63
lonely	<u>21</u>	8.29
sad	<u>19</u>	7.38
pathetic	<u>5</u>	7.04
weary	<u>4</u>	6.62
ironic	<u>4</u>	6.56
tired	<u>10</u>	6.54
disappointed	<u>5</u>	6.32
funny	<u>9</u>	6.28
tragic	<u>4</u>	6.18
depressed	<u>4</u>	6.04
bitter	<u>6</u>	5.97
happy	<u>21</u>	5.96
sick	<u>7</u>	5.81
angry	<u>7</u>	5.75
mad	<u>5</u>	5.72
strange	<u>10</u>	5.56
worried	<u>5</u>	5.5
gentle	<u>5</u>	5.49
sweet	<u>5</u>	5.4
anxious	<u>5</u>	5.25
proud	<u>4</u>	5.24
little	<u>39</u>	5.17
empty	<u>5</u>	4.85
pale	<u>4</u>	4.76

<u>unary rels</u>		
it+	<u>86</u>	11.5

<u>adj comp of</u>	<u>198</u>	<u>19.4</u>
feel	<u>94</u>	5.55
look	<u>60</u>	4.48
sound	<u>6</u>	4.41
seem	<u>19</u>	3.1
grow	<u>4</u>	2.66

<u>np adj comp of</u>	<u>56</u>	<u>15.2</u>
find	<u>16</u>	1.8
make	<u>32</u>	1.53

<u>adj subject</u>	<u>132</u>	<u>4.7</u>
face	<u>7</u>	1.95
eye	<u>7</u>	1.6

<u>modifies</u>	<u>1442</u>	<u>2.5</u>
reflection	<u>31</u>	7.56
farewell	<u>10</u>	7.05
indictment	<u>7</u>	6.78
tale	<u>19</u>	6.42
ending	<u>8</u>	6.36
clown	<u>5</u>	6.28
git	<u>4</u>	6.23
smile	<u>29</u>	6.19
plight	<u>5</u>	5.9
truth	<u>18</u>	5.77
story	<u>50</u>	5.77
spectacle	<u>4</u>	5.76
ruin	<u>5</u>	5.75
irony	<u>4</u>	5.6
demise	<u>4</u>	5.58
news	<u>29</u>	5.39
bastard	<u>4</u>	5.25
blow	<u>6</u>	5.2
fact	<u>53</u>	5.18
disappointment	<u>4</u>	5.08
ghost	<u>4</u>	5.08
reality	<u>11</u>	5.04
sight	<u>10</u>	4.94
decline	<u>8</u>	4.87
note	<u>17</u>	4.78

<u>modifier</u>	<u>737</u>	<u>1.1</u>
terribly	<u>13</u>	7.9
very	<u>290</u>	6.85
rather	<u>41</u>	6.82
profoundly	<u>4</u>	6.69
bit	<u>24</u>	6.67
strangely	<u>4</u>	6.61
desperately	<u>6</u>	6.41
so	<u>112</u>	6.08
little	<u>16</u>	5.85
extremely	<u>9</u>	5.41
really	<u>30</u>	5.31
somewhat	<u>5</u>	5.29
slightly	<u>8</u>	5.19
particularly	<u>12</u>	5.12
pretty	<u>4</u>	4.84
quite	<u>14</u>	4.3
always	<u>14</u>	4.14
almost	<u>6</u>	3.83
too	<u>11</u>	3.44
both	<u>4</u>	2.94
even	<u>4</u>	2.42
as	<u>4</u>	2.14
just	<u>6</u>	1.87
most	<u>5</u>	1.78
more	<u>7</u>	1.4

発展的問題 2/2

- だが, 待て
 - sad __ に現われやすい次の名詞は cigarette に似ているか??
 - reflection, farewell, indictment, tale, ending, clown, git, smile, plight, truth, story, spectacle, ruin, irony, demise, news, bastard, blow, fact, disappointment, ghost, reality, sight, decline, note, ...
- 結論
 - このアプローチは見こみ薄
- 理由
 - 転移修飾が詩的な表現であるので, 典型性を利用した判定は無意味

間接法の概要

- 課題 (再掲): コーパスに基づいて (1) の表現が容認されるか判定する

(1) He was smoking a sad cigarette.

- ヒト主語指向条件 (C1): 次の類例の個数を N_1 とする
 - he {is, was, feels, felt, turned out to be, ...} sad
- 交替条件 (C2): 次のいずれかのパターンの類例の個数を N_2 とする
 - he smoked a cigarette sadly
 - he smoked sadly a cigarette
 - he sadly smoked a cigarette
- (1) の容認度の推定
 - N_1 の個数が十分ならば, N_2 が多いほど (1) は容認度が上がる

間接法 1/3

- ヒト主語指向条件の CQL
 - [tag= "PNP" & lemma!="it" & lemma!="this" & lemma!="that" & lemma!="they" & lemma!="these" & lemma!="those"]
[tag="V.*"]* [lemma="be"] [tag="CJ.* |AV.* |AJ.*"]*
[word="sad"]
- のBNC中の実例数=236
- この結果は sad がヒト主語指向形容詞 (Human-Oriented A: HOA) だと示唆

間接法 2/3

- 交替条件 `pattern1` の CQL
 - `[lemma="cigarette"] [tag="AV.*"]* [lemma="sadly"]`
- の実例数 = 0
- 交替条件 `pattern2` or `pattern3` の CQL
 - `[lemma="sadly"] [tag!="PU.*"]* [lemma="cigarette"]`
- のBNC中の実例数 = 0
- 因みに
 - enTenTenで同様の検索を行なっても正例数=0

間接法 3/3

- コーパス規模が十分でない場合には、間接法を用いても (1) が容認可能だという証拠は得られない
- ただし
 - 理由の一部は `sad cigarette` の組み合わせが奇抜すぎるため
- 実際
 - 後述の間接法の実行例から、ヒト主語指向条件と交替条件の両方を満足する、もっと普通の転移修飾は数多く見つかる

本発表の狙い

- 神澤 (2011)の手法で英語のTEの実態を調査する
 - 神澤は日本語の調査をしているだけなので、本家本元の英語で同じことをやってみる
- TE認定の手順の大部分が自動化可能であることを示す
 - $G: S V (P) O A\text{-ly} X \Rightarrow F: S V (P) A O X$ の変形は自動化可能
 - $F: S V (P) A O X$ の容認度は“AO”の事例数を数えることで近似可能
 - $F: S V (P) A O X$ でAがSの述語になるかどうかは、 $S \{is, was, \dots, become, \dots\}$ A の事例数を数えることで近似可能

最終チェックの例 (V D ... grateful N)

Corpus: British National Corpus

Hits: 111

Page 1 of 6 [Go](#) [Next](#) | [Last](#)

JXX	added in a mocking drawl, `you	seemed remarkably grateful and content	in my arms last night! </p><p>
JXV	her with a smile that would have	made a lesser woman fall in a grateful little heap	at his feet. `We know, don't
JY8	she offered, and the young woman	took it with a grateful smile	. </p><p> `Thanks. You must be Paige
JYE	` </p><p> They left the cottage,	leaving a grateful if sad couple	who were quite amazed at the
JYC	Vaughan to stay the night, Leonora	declined with grateful thanks	, hoping her longing to escape
HRJ	study which had been Harold's and	write him a letter of appreciation and grateful thanks	. </p><p> Strangely enough, at no
HRT	Thanks <p> THE Summer Fayre Committee	express grateful thanks	to the following for their generous
HTX	Yussuf shuffled off and Georgiades	waved a grateful hand	. </p><p> `I thought you had powerful
HGM	just at the moment! </p><p> Kate	gave him a grateful smile	. `This idea of a job change has
HGG	's letter, duly come to hand, I	send you my thanks and my grateful sense	of your lordship's kindness and
HGV	is why -' raising her head to	give him a grateful look	` - I had none to spare for a
HGV	for all of them! </p><p> Theda	threw her a grateful look	, and hurried after Miss Merchiston
HGV	not abandon you.' </p><p> Theda	gave him a grateful smile	, but rueful withal. `I am causing
HGE	earlier filled with sweet peas	brought by a grateful patient	who owned a little garden. She
HSA	front of them and Stuart's Mum	parked with a grateful wave	to the driver. </p><p> Standing
HSG	his electro-mechanical gadgets,	dreaming of a future in which a grateful world	uses Bain's invention and blesses
HJ3	with only the keeper to beat,	shot weakly at a grateful Henderson	. </p><p> The Dubliners' were desperately
HHO	looks at me as if I am crazy, and	watches a grateful cat	dragging the carcass to a safe
HHV	its last year in office. May I	extend the grateful thanks	of Mr. and Mrs. Emery of Eye
HHV	which adopted such a policy would	have the grateful thanks	of a vast number of women? </p>

Page 1 of 6 [Go](#) [Next](#) | [Last](#)

以下の流れ

- まず、神澤 (2011) を手法紹介しながら、日本語の事例を扱う
- 次に、同様の手法で英語の実例を扱う

Part II
神澤 (2011)の調査の紹介

神澤 (2011) の解説

- 神澤 (2011) は次の特徴をもった日本語の転移修飾の研究
 - 構文交替条件を転移修飾の必須条件と考えて分析
 - 自作例と内省に全面的に依存した分析を退け、可能な限りコーパス事例に基づいた記述と説明を行なう
 - コーパスの実例を尊重するが、それに拘泥しない
- 特に最期の点で、コーパス基盤のTE研究の先駆である Shibuya (2008)の手法と対比的

Shibuya (2008) の手法

- BNC 中の形容詞 A の事例 (e.g., $A N$) を検索する. 得られた事例を $E = \{e_1, e_2, \dots, e_n\}$ とする
- E で A が目的語をもっている事例 (e.g., $S V A O$) のみを選択する. 結果を $F = \{f_1, f_2, \dots, f_m\}$ ($m < n$) とする.
- F の事例のそれぞれについて, A が S を修飾する述語になるかを判定する (C2の意味)
 - 条件が満足された場合に TE とし, 満足されない場合は $non TE$ とする

神澤 (2011) の手法

- Sketch Engine の JpWaC コーパス 中の副詞形 (e.g., $S V A\text{-ly}$, S が A (し)く V) を検索する. 得られた事例を $E = \{e_1, e_2, \dots, e_n\}$ とする
- E で目的語をもっている事例 (e.g., $S V O A\text{-ly}$, S が A (し)く O を V) のみを選択する. 結果を $F = \{f_1, f_2, \dots, f_m\}$ ($m < n$) とする.
- F の個々の事例を (e.g., $S V A O$, S が A (し)い O を V) という形容詞構文に変形する. この結果を $G = \{g_1, g_2, \dots, g_m\}$ とする
- G の事例のそれぞれについて, A が S を修飾する述語になっているかを判定する (C2の意味)
 - この条件が満足された場合に TE とし, 満足されない場合は $non TE$ とする

神澤 (2011) の手法

- 三つのパターン
 - A がNをX(し)くVする (source 1)
 - A がX(し)くNをVする (source 2)
 - A がNをVしてX(し)くなる (source 3)
- をそれぞれ
 - A が X(し)い N を V する (target = paraphrase of source i)
- に変形し, その結果が解釈可能か評価し. 解釈可能な場合に
 - AがXし{い, そうだ}
- の読みが得られるかを評価

手法の特徴

- 神澤 (2011) の方法は, Shibuya (2008) の手法と反対で次の利点がある
 - A. 探索範囲を有意義に限定できる
 - TEの発生率は低く, AN構文の事例数は膨大なので, Aの事例を絞り込む方法では正例が効率的に探せない
 - B. 作例を手順に含んでいるので, 正例だけでなく負例も扱うことが可能
 - C. 正例と負例の境界例の位置づけが容易
 - TEは修辭的な効果のある表現で, TE候補の大半は元々容認度が低い表現のはずなのだが, 先行研究では正例と負例の中間例の扱いに関する考慮が欠落している

手法の特徴

- Bに関しては，コーパス言語学のドグマの一つ“正例主義”を克服する期待がもてる
 - 従来のコーパス言語学では負例が扱えない
- 言語学での正例と負例の利用価値の違いに関しては，黒田 (in press) を参照

Source 1

- Sketch Engine の JpWac で、CQL: [tag = “Ai.free” & word = “. * く ”] [] { 1, 3 } [lemma = “を”] を入力すると 56,768 例が得られた。これを S とする。
- 後から考えると [word = “. * を ”] [tag = “A. * ” & word = “. * く ”] [tag = “V. * ”] の方が効率的
- S から 3,000 例をランダム抽出した後、手作業でノイズとなる 11 例を取り除くことで、1,727 例が得られた。これらはすべて “(A が) N を X(し) く V” の例。この例の集合を S* とする。
- S* のそれぞれを “(A が) X(し) い N を V する” にパラフレーズし、真理条件的な意味が保持されるものを目視で選出。この結果、170 例が得られた (S* のうちの 9.8%)。
- パラフレーズで真理条件を保持しない用例は転移修飾の条件を満たしていない

Source 1

- “(Aが) Nを X(し)く V” と “(Aが) X(し)い Nを V” とが交替を起こす形容詞 X の一覧
- 異なり数は46

X 位置の形容詞	延べ語数
つよい	40
かずおおい	19
たのしい	7
あつい(熱い)	6
おおきい	6
ひろい	6
きもちいい	5
たかい	5
なつかしい	5
あたらしい	4
いろこい	4
うまい	4
くわしい	4
ただしい	4
いちはやい	3
うらやましい	3
こまかい	3
すばやい	3
ふるい	3
ありがたい	2
いとおいしい	2
きそくただしい	2
さみしい	2
ちからづよい	2
てばやい	2

はばひろい	2
あかるい	1
あまい	1
うつくしい	1
うとましい	1
おそろしい	1
かんがいぶかい	1
きよい	1
きょうみぶかい	1
こいしい	1
すえながい	1
ちゅういぶかい	1
ながい	1
なまなましい	1
ねづよい	1
はがゆい	1
はなばなしい	1
はらだたしい	1
ひとあしはやい	1
ほこらしい	1
よぎない	1

Source 2

- Sketch Engine の JpWacで、CQL: [lemma = “を”] [tag = “Ai.free” & word = “.*く”] を入力すると41,840 例が得られた。これをSとする。
- 後から考えると [tag = “A.*” & word = “.*く”] [word=“.*を”] [tag = “V.*”]の方が効率的
- Sから3,000例をランダム抽出した後、手作業でノイズとなる例を取り除いて、1,290 例が得られた。これらはすべて “X(し)く (A が) N を V”の例。この例の集合を S* とする。
- S*のそれぞれを “(A が) X(し)い N を V する” にパラフレーズし、真理条件的な意味が保持されるものを目視で選出。この結果、454 例が得られた (S* のうちの 35.2%)。
- パラフレーズで真理条件が保持されない用例は転移修飾の条件を満たしていない

Source 2

- “(Aが) X(し)く Nを V” と “(Aが) X(し)い Nを V” とが交替を起こす形容詞 X の一覧
- 異なり数は74

X 位置の形容詞	延べ語数
おおきい	55
つよい	41
たのしい	26
ふかい	25
あたらしい	20
かるい	19
あつい(厚い)	18
くわしい	17
はやい	16
すごい	14
こまかい	13
きびしい	11
きもちいい	11
ものすごい	11
はばひろい	10
なにげない	8
こうりつよい	7
ちいさい	7
いちじるしい	6

ながい	6
ねばりつよい	5
さみしい	4
すばやい	4
ただしい	4
てばやい	4
あかるい	3
いちはやい	3
しつこい	3
ちゆういぶかい	3
ありがたい	2
いそがしい	2
うすい	2
うつくしい	2
うまい	2
えんりよない	2
おとこらしい	2
かたい	2
きそくただしい	2
きょうみぶかい	2

したしい	2
ちからづよい	2
なつかしい	2
めまぐるしい	2
やさしい	2
れいぎただしい	2
あたたかい	1
あわただしい	1
あつい(熱い)	1
おおい	1
おとなしい	1
かいがいしい	1
かっこいい	1
がまんづよい	1
かわいい	1
ぎこちない	1
けだるい	1
こころよい	1

しぶとい	1
ずうずうしい	1
すっぱい	1
ずぶとい	1
するどい	1
たえまない	1
たやすい	1
てびろい	1
とてつもない	1
なごりおしい	1
なまなましい	1
ねづよい	1
ひどい	1
ひとしい	1
ひんよい	1
まぶしい	1
やむない	1

Source 3

- Sketch Engine の JpWac で、CQL: [lemma = “する”] [lemma = “て”] [tag = “Ai.*” & word = “.*く”] を入力すると 3,019 例が得られた。これを S とする。
- 後から考える [lemma = “を”] [tag! = “V.*”]* [tag = “V.*”] [lemma = “て”] [tag = “Ad.*”]* [word = “.*く”] [lemma = “なる”] の方が効率的だった。この検索式に該当するのは 506 例でほぼノイズなし。
- S の前例をノイズとなる例を手作業で取り除いて、270 例が得られた。これらはすべて “(A が) N を V して X(し)くな(る|った)” の例。この例の集合を S* とする。
- S* のそれぞれを “(A が) X(し)い N を V する” にパラフレーズし、真理条件的な意味が保持されるものを目視で選出。この結果、27 例が得られた (S* のうちの 19.0%)。
- パラフレーズで真理条件が保持されない用例は転移修飾の条件を満たしていない

Source 3

- “(Aが) Nを Vして X(し)くな(る | り | っ)” と “(Aが) X(し)い Nを V” とが交替を起こす形容詞 X の 一覧
- 異なり数は16

X 位置の形容詞	延べ語数
うれしい	4
かなしい	3
かゆい	2
きもちいい	2
ぎこちない	2
くるしい	2
さみしい	2
たのしい	2

あつい(熱い)	1
あわただしい	1
いたい	1
くらい	1
こころぼそい	1
つらい	1
はらだたしい	1
わびしい	1

分類 1/2

- 以上の手法で収集された形容詞を，T: “A (は|が) Xし(い|そ
うだ)” の容認度によって，以下の第1群，第2群，第3群に分
類
- 第1群: Tが問題なく容認可能 (=日本のTEの確実な正例)
- 第2群: Tの容認度の問題あり (=日本のTEの境界例)
- 第3群: Tが問題なく容認不可能 (=日本のTEの確実な負例)

分類 2/2

- “AしいNをV” と交替するパターン
- 1群
 - “NをAしくV” と交替
- 2群
 - “AしくNをV” と交替
- 3群
 - “NをVし(て)Aしくなる”

lemma	S1Rate:N をXしくV	S2Rate:X しくNをV	S3Rate: N をVしてXし くなる
たのしい	0.29	0.49	0.11
気持ちよい	0.21	0.21	0.00
さみしい	0.08	0.08	0.00
いとおしい	0.08	0.00	0.00
あかるい	0.04	0.06	0.00
はらだたい	0.04	0.00	0.06
うしろめたい	0.04	0.00	0.00
かんがいぶかい	0.04	0.00	0.00
こいしい	0.04	0.00	0.00
にがにがしい	0.04	0.00	0.00
はがゆい	0.04	0.00	0.00
ほこらしい	0.04	0.00	0.00
いそがしい	0.00	0.04	0.00
あたたかい	0.00	0.02	0.00
おとなしい	0.00	0.02	0.00
かいがいしい	0.00	0.02	0.00
がまんづよい	0.00	0.02	0.00
けだるい	0.00	0.02	0.00
なごりおしい	0.00	0.02	0.00
まぶしい	0.00	0.02	0.00
うれしい	0.00	0.00	0.22
かなしい	0.00	0.00	0.17
かゆい	0.00	0.00	0.11
くるしい	0.00	0.00	0.11
いたい	0.00	0.00	0.06
こころぼそい	0.00	0.00	0.06
つらい	0.00	0.00	0.06
わびしい	0.00	0.00	0.06
SUM	24	53	18

S1由来の容認度の高い事例

- 私は楽しい余暇を過ごす
 - 私は余暇を楽しく過ごす
- 私は気持ちよいオフロードを走り抜ける
 - 私はオフロードを気持ちよく走り抜ける
- 私はすばやい朝食をとる
 - 私は朝食をすばやくとる

S1由来の容認度の低い事例

- *私はすでに深い彼女を愛していた
- 私はすでに彼女を深く愛していた

S2由来の容認度の高い事例

- 私は明るいブログ生活を楽しむ
 - 私は明るくブログ生活を楽しむ
- 私は忙しい商いをさせていただいております
 - 私は忙しく商いをさせていただいております
- 私は彼らに対して厳しい批判をする
 - 私は彼らに対して厳しく批判をする
- 私はしつこいアプローチを重ねる
 - 私はしくこくアプローチを重ねる

S2由来の容認度の低い事例

- *私は気持ちよい眼球を動かした
 - 私は気持ちよく眼球を動かした
- *私はすばやい周囲を見回した
 - 私はすばやく周囲を見回した

S3由来の容認度の高い事例

- 私は慌ただしい早朝出勤をした
- 土の中の成分が赤い発色をする
- うれしい勘違いをしちゃうそうです
- 空気のうれしい満喫をした
- ?気持ちのいい罵倒をする
- 気持ちいい浸透をする
- ?二人とも可哀相だなあ、と悲しい妄想をする
- ?離れる瞬間について寂しい想像をする
- ?冬になると肌がかゆい乾燥をする
- ?一人で頑張ってきた母も心細い病気をし...
- ?電車の中で寒い居眠りをする
- ?自分はクーラーの下で寒い長居をした
- ?彼女が美しいあか抜けをする

S3由来の容認度の低い事例

- ??頭が気持ちいいスッキリをする
- ??心臓が苦しいチクチクをする
- ??自分の言動に恥ずかしい反省をする
- ??楽しい熱中をしてきました!
- ??おとなしい酩酊をする
- ??筋肉が痛い硬直をして眠れない
- ??木が暗い密集をした森林では...
- ?*神父様の声を聞いていると眠い安心をする
- ?*やわらかい加熱をする
- ?*水と赤い反応をしたのに間違いないようです
- ?*門人達の方が青い心配をした
- ?*ヘアメイクに美しい挑戦をする
 - 美しいヘアメイクに挑戦する

まとめ

- 神澤 (2011) は
 - S1: SがOをAしくVする
 - S2: SがAしくOをVする
 - S3: SがOをVしてAしくなる
- の三つのソースからT: “SがAしいOをV” を作例し,
 - (i) T が真理条件を保存するか (交替条件C2)
 - (ii) AがSの述語になるか (ヒト主語指向述語条件C1)
- の二規準でTを評価し, 結果をまとめた.

Part III
英語の転移修飾の調査

調査の内実 1/2

- 英語でヒト主語指向条件C1と交替条件C2を独立に調査するが、特に人手評価の段階を可能な限り自動化するのが目的
- 理由
 - TECは逸脱性の高い(詩的な)表現なので、非母語話者には評価が難しいし、母語話者の評価も代表性を有するものになるか怪しい

調査の内実 2/2

- ただし
 - C1は終わっているが、時間不足で C2は P_1 を基にした候補 Q の事例数の調査までしか終わっていない
- 理由
 - 調査に必要なクエリが多く、完遂までに時間がかかる
 - 即答性のある調査には、ローカルで Sketch Engine を走らせないとダメ

ヒト主語指向条件C1の調査

- 手順

1. $P_1 = \text{VO A-ly}$, $P_2 = \text{V A-ly O}$, $P_3 = \text{A-ly VO}$ の3パターンのそれぞれについて
2. 実例 p_1, p_2, \dots, p_n から $Q = \text{S be A}$ の実例 q_1, q_2, \dots, q_n を自動生成
3. Sketch Engine の API を使って q_i のコーパス中の事例数を獲得
4. Q のうち、実例数=0でないものを目視で評価

- Step 4 は未完了

- 使用コーパスは BNC

交替条件C2の調査

- 手順

1. $P_1 = \text{V O A-ly}$, $P_2 = \text{V A-ly O}$, $P_3 = \text{A-ly V O}$ の3パターンのそれぞれについて
2. 実例 p_1, p_2, \dots, p_n から $Q = \text{V A O}$ の実例 q_1, q_2, \dots, q_n を自動生成
3. Sketch Engine の API を使って q_i のコーパス中の事例数を獲得
4. Q のうち、事例数=0でないものを目視で評価

- Step 4 は未完了

- 使用コーパスは BNC

詳細 1/2

- いずれの調査でも Step 3で Python API を使用し BNC 中の実例を検索
- API の使用に関する重要な注意
 - query の数が膨大 (e.g., 30,000+) になります
 - 10秒ごとにqueryを再発行するとかして“紳士的”にAPIを利用しないとアカウントを閉じられます

詳細 2/2

- P1, P2, P3の事例獲得のための CQL

- [lemma!="be"] [tag="V.*" & lemma!="be"] [tag="PR.*"]* [tag="AT.* | DT.* | CDR | AJ.* | N.*"]* [tag="N.*"] [word=".*ly" & tag="AV.*" & lemma!="only"]
- [tag="V.*" & lemma!="be"] [word=".*ly" & tag="AV.*" & lemma!="only"] [tag="AT.* | DT.* | CDR | AJ.* | N.*"]* [tag="N.*"]
- [word=".*ly" & tag="AV.*" & lemma!="only"] [tag="V.*" & lemma!="be" & tag!="V.*N" & tag!="V.*G"] [tag="AV.* | V.*" & tag!="V.*N" & tag!="V.*G"] * [tag!="N.*" & tag!="V.*" & tag!="PR.*" & tag!="PU.*"]* [tag="N.*"] [tag!="N.*" & word!="(|)|-"]

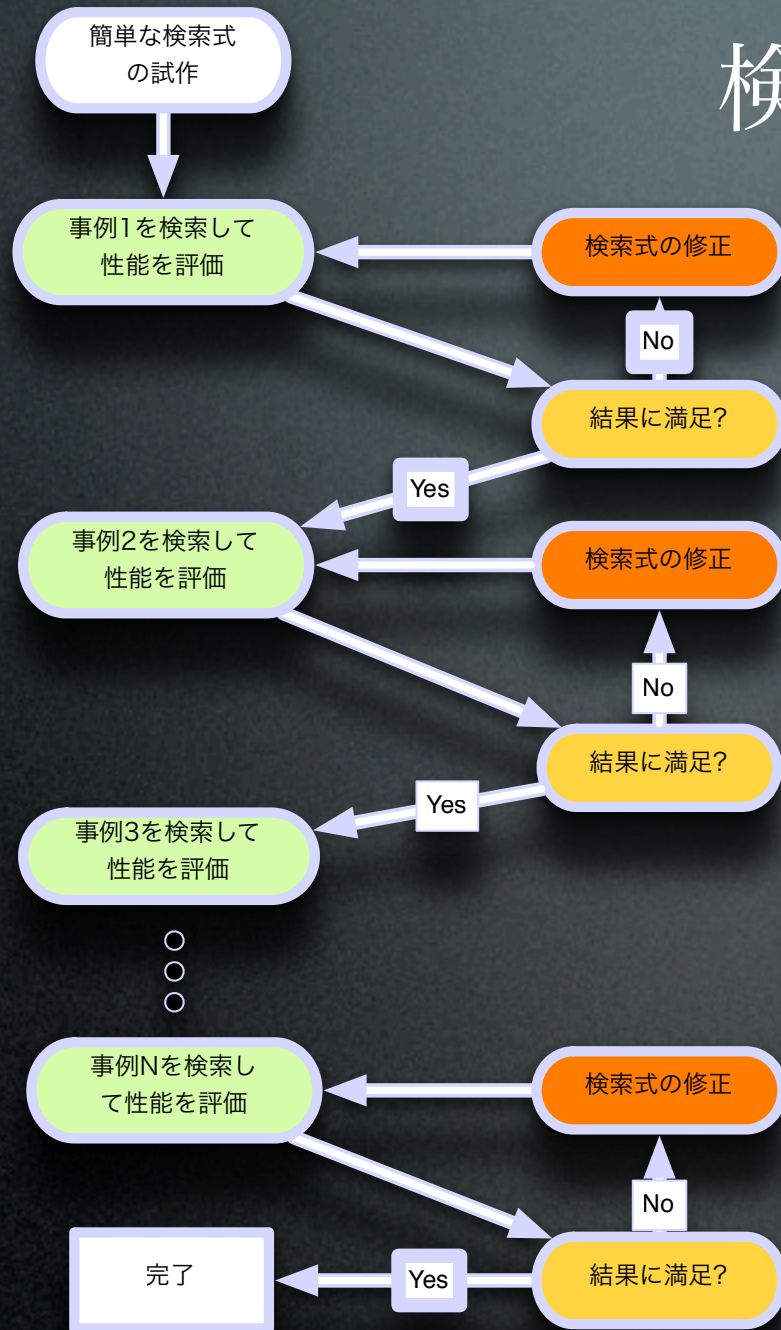
- P1; P2; P3 のBNC中の事例数

- 25,181; 10,411; 44,571 (合計 80,163)

CQL の選定に関する補足

- 検索性能のよい CQL を構築するには、それなりの試行錯誤が必要
 - 検索性能が高いとは,
 - なるべく true positives を取りこぼさず (i.e. coverage が大)
 - なるべく false positives を含まない (i.e. precision が高)
 - 例えば F値 = $\frac{\text{\#true positives}}{\text{\#true positives} + \text{\#false positives}}$ が高い
- 検索式を試行錯誤するのは大変ですが、やり甲斐のある作業
 - この“泥臭い”作業が楽しめない人はデータ駆動言語学に向いていないでしょう

検索式の開発サイクル



BNC上での調査結果

C1の調査結果

- S1, S2, S3 のいずれかに現われたA-lyを変形したAの異なり数は2322

- わずかに誤変換 (e.g., likely => like*, wholly=>whol*) も含む
- BNCのA-ly形が約9600個なので、全体の25%ほど

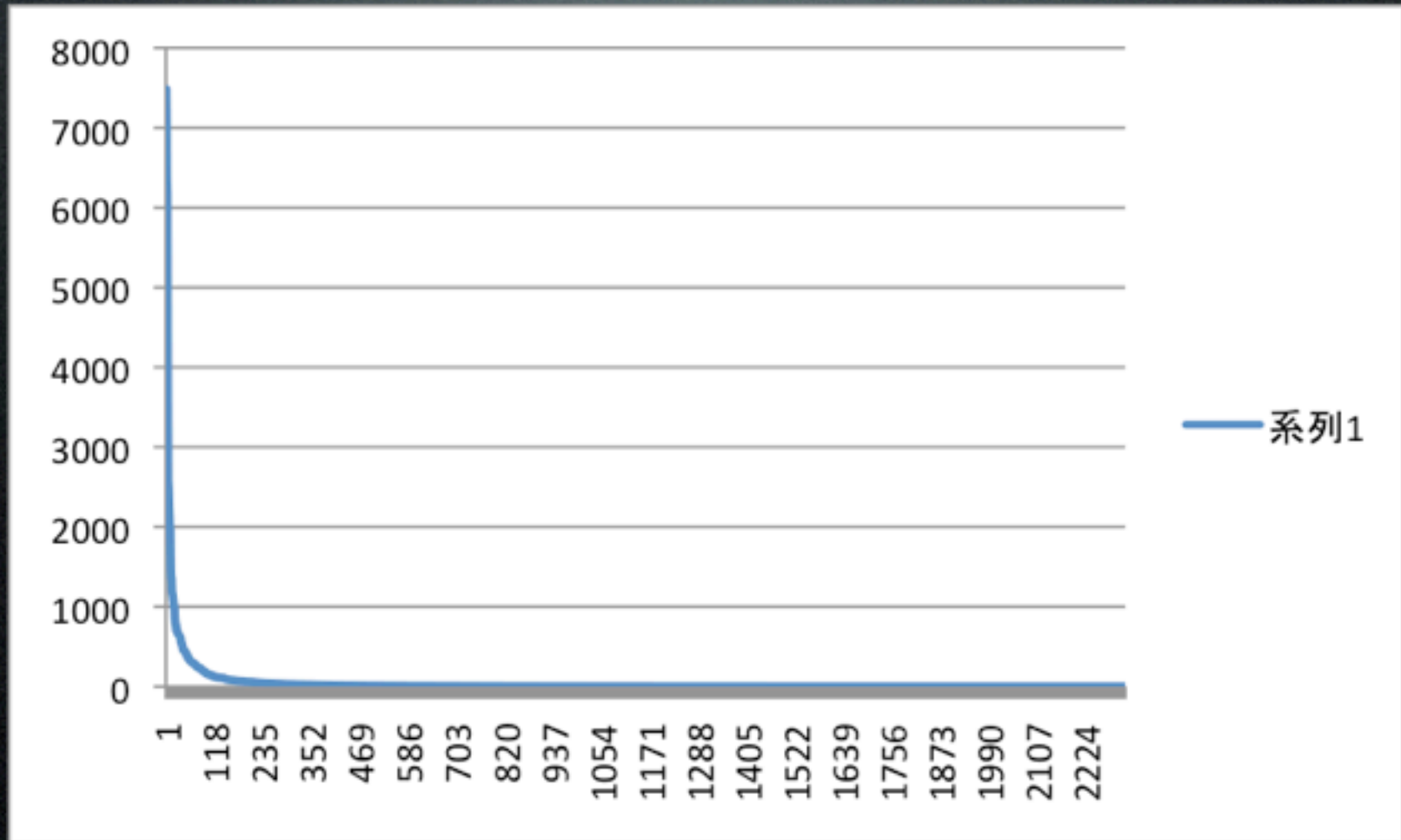
- C1のヒト主語指向条件を判定したCQL

- `[tag= "PNP" & lemma!="it" & lemma!="this" & lemma!="that" & lemma!="they" & lemma!="these" & lemma!="those"] [tag="V.*"]* [lemma="be"] [tag="AV.* | CJ.* | AJ.*"]* [word=ADJ]`

- 最高頻度はsurely=>sureの7482
- 64度以上の頻度であれば十分に強いヒト主語指向性をもつと言って良い結果

Freq	#Types	Rate
1以上	1454	62.8%
2以上	1180	50.8%
4以上	903	38.9%
8以上	651	28.0%
16以上	452	19.5%
32以上	298	12.8%
64以上	190	8.2%
128以上	118	5.1%
256以上	74	3.2%

Freq rank 分布はべき乗分布



C1の該当数でソートした結果

	A	B	C	D	E	F	G	H	I	J	K
1	ADV	SIRate	S2Rate	S3Rate	SI	S2	S3	Row_	ADJ	Person	note
2	surely	0.07	0.04	0.90	22	12	293	327	sure	7482	
3	verily	0.00	0.00	1.00	0	0	3	3	very	7126	誤変換
4	ably	0.22	0.22	0.56	2	2	5	9	able	6355	
5	justly	0.23	0.00	0.77	3	0	10	13	just	6176	
6	rightly	0.07	0.03	0.90	11	5	152	168	right	4059	
7	happily	0.30	0.06	0.64	50	10	108	168	happy	2585	
8	gladly	0.38	0.02	0.60	34	2	53	89	glad	2493	
9	onely	0.50	0.50	0.00	1	1	0	2	one	2268	
10	interestedly	1.00	0.00	0.00	1	0	0	1	interested	2143	
11	pleasedly	0.00	0.00	1.00	0	0	1	1	pleased	1981	
12	likely	0.01	0.14	0.85	1	10	61	72	like	1968	誤変換
13	gratefully	0.36	0.03	0.61	26	2	44	72	grateful	1485	
14	luckily	0.33	0.17	0.50	2	1	3	6	lucky	1365	
2312	wholly	0.38	0.26	0.35	38	26	35	99	whol	0	誤変換
2313	willy-nilly	1.00	0.00	0.00	6	0	0	6	willy-nil	0	誤変換
2314	willynilly	1.00	0.00	0.00	1	0	0	1	willynil	0	誤変換
2315	wistfully	0.94	0.00	0.06	16	0	1	17	wistful	0	
2316	wittingly	0.00	0.00	1.00	0	0	2	2	witting	0	
2317	woefully	0.50	0.50	0.00	1	1	0	2	woeful	0	
2318	wolly	0.00	0.00	1.00	0	0	2	2	wol	0	
2319	wondrously	1.00	0.00	0.00	1	0	0	1	wondrous	0	
2320	wordlessly	0.75	0.00	0.25	3	0	1	4	wordless	0	
2321	wrongfully	0.33	0.00	0.67	2	0	4	6	wrongful	0	
2322	wryly	0.71	0.05	0.24	15	1	5	21	wry	0	
2323	wurly	0.00	0.00	1.00	0	0	1	1	wur	0	
*											
2325					28905	12265	59406				

調査の制限

- 主語が Theme の場合と Experiencer の場合が未区別
 - He is amusing では He は Theme で他に Experiencer がいる
 - He is amusing to us.
 - He is happy (about it) では He は Experiencer (で Theme/Stimulus が it.)
- 一部の形容詞 (e.g., ready) では人が主語の場合に Theme 読みと Experiencer 読みが可能
 - He was ready for us. <= Theme
 - He was ready to go. <= Experiencer/Raised ?

S1の頻度でソートした結果

	A	B	C	D	E	F	G	H	I	J
1	ADV	S1Rate	S2Rate	S3Rate	S1	S2	S3	Row	ADJ	Person
2	directly	0.67	0.08	0.25	738	86	273	1097	direct	21
3	immediately	0.42	0.04	0.54	685	68	875	1628	immediate	4
4	really	0.16	0.06	0.78	646	262	3134	4042	real	134
5	early	0.75	0.18	0.07	615	144	57	816	early	166
6	actually	0.14	0.01	0.85	519	44	3153	3716	actual	1
7	quickly	0.38	0.02	0.61	496	21	792	1309	quick	201
8	carefully	0.46	0.11	0.43	411	98	380	889	careful	554
9	properly	0.65	0.08	0.28	382	45	164	591	proper	18
10	simply	0.17	0.04	0.79	381	95	1811	2287	simple	36
11	firmly	0.69	0.02	0.29	361	9	151	521	firm	44
12	recently	0.36	0.04	0.60	356	44	598	998	recent	2
13	seriously	0.41	0.16	0.44	326	126	351	803	serious	278
14	completely	0.38	0.14	0.49	321	116	416	853	complete	22
15	regularly	0.39	0.03	0.58	302	20	448	770	regular	7
16	shortly	0.77	0.05	0.18	264	16	62	342	short	254
17	generally	0.24	0.04	0.72	252	38	740	1030	general	20
2315	unimpeachably	0.00	0.00	1.00	0	0	1	1	unimpeachable	0
2316	unnecessarily	0.00	0.00	1.00	0	0	1	1	unnecessary	0
2317	unsuspectingly	0.00	0.00	1.00	0	0	1	1	unsuspecting	0
2318	untruthfully	0.00	0.00	1.00	0	0	1	1	untruthful	0
2319	vaginally	0.00	0.00	1.00	0	0	1	1	vaginal	0
2320	vengefully	0.00	0.00	1.00	0	0	1	1	vengeful	0
2321	virally	0.00	0.00	1.00	0	0	1	1	viral	0
2322	wheely	0.00	0.00	1.00	0	0	1	1	whee	0
2323	wurly	0.00	0.00	1.00	0	0	1	1	wur	0
*										
2325					28905	12265	59406			

S2の頻度でソートした結果

	A	B	C	D	E	F	G	H	I	J
1	ADV	SIRate	S2Rate	S3Rate	SI	S2	S3	Row_	ADJ	Person
2	exactly	0.20	0.64	0.16	232	738	179	1149	exact	4
3	nearly	0.14	0.46	0.39	138	445	375	958	near	299
4	relatively	0.16	0.81	0.04	72	366	16	454	relative	3
5	virtually	0.16	0.60	0.25	84	313	129	526	virtual	1
6	really	0.16	0.06	0.78	646	262	3134	4042	real	134
7	precisely	0.27	0.65	0.09	101	247	33	381	precise	24
8	absolutely	0.18	0.69	0.14	59	231	47	337	absolute	6
9	significantly	0.16	0.42	0.42	89	229	232	550	significant	10
10	highly	0.15	0.58	0.26	56	212	96	364	high	88
11	particularly	0.31	0.29	0.40	216	201	279	696	particular	32
12	clearly	0.14	0.12	0.74	218	195	1149	1562	clear	231
13	fairly	0.41	0.50	0.09	148	178	33	359	fair	93
14	extremely	0.28	0.69	0.03	70	172	8	250	extreme	0
15	slightly	0.46	0.36	0.18	213	166	81	460	slight	8
16	mainly	0.35	0.41	0.25	139	162	99	400	main	3
17	increasingly	0.14	0.36	0.50	58	147	208	413	increasing	33
2315	vasarely	1.00	0.00	0.00	1	0	0	1	vasare	0
2316	vegetatively	1.00	0.00	0.00	1	0	0	1	vegetative	0
2317	ventrally	1.00	0.00	0.00	1	0	0	1	ventral	0
2318	villainously	1.00	0.00	0.00	1	0	0	1	villainous	0
2319	virulently	1.00	0.00	0.00	1	0	0	1	virulent	0
2320	voluntairily	1.00	0.00	0.00	1	0	0	1	voluntairy	0
2321	whole-heartedly	1.00	0.00	0.00	1	0	0	1	whole-hearted	0
2322	willynilly	1.00	0.00	0.00	1	0	0	1	willynil	0
2323	wondrously	1.00	0.00	0.00	1	0	0	1	wondrous	0
*										
2325					28905	12265	59406			

S3の頻度でソートした結果

	A	B	C	D	E	F	G	H	I	J
1	ADV	SIRate	S2Rate	S3Rate	SI	S2	S3	Row_	ADJ	Person
2	actually	0.14	0.01	0.85	519	44	3153	3716	actual	1
3	really	0.16	0.06	0.78	646	262	3134	4042	real	134
4	probably	0.07	0.05	0.89	172	119	2294	2585	probable	0
5	usually	0.10	0.01	0.89	238	14	2107	2359	usual	19
6	simply	0.17	0.04	0.79	381	95	1811	2287	simple	36
7	certainly	0.04	0.02	0.94	73	29	1692	1794	certain	770
8	clearly	0.14	0.12	0.74	218	195	1149	1562	clear	231
9	eventually	0.09	0.01	0.90	117	11	1105	1233	eventual	0
10	normally	0.11	0.02	0.87	127	20	991	1138	normal	49
11	finally	0.13	0.01	0.86	145	10	988	1143	final	1
12	necessarily	0.04	0.01	0.95	37	7	909	953	necessary	38
13	immediately	0.42	0.04	0.54	685	68	875	1628	immediate	4
14	hardly	0.04	0.10	0.86	41	96	844	981	hard	230
15	merely	0.11	0.10	0.79	110	105	822	1037	mere	6
16	quickly	0.38	0.02	0.61	496	21	792	1309	quick	201
17	suddenly	0.21	0.02	0.77	207	19	767	993	sudden	13
2313	tortuously	0.00	1.00	0.00	0	1	0	1	tortuous	0
2314	unalterably	0.00	1.00	0.00	0	1	0	1	unalterable	0
2315	unarguably	0.00	1.00	0.00	0	1	0	1	unarguable	0
2316	uncommercially	0.00	1.00	0.00	0	1	0	1	uncommercial	0
2317	uncountably	0.00	1.00	0.00	0	2	0	2	uncountable	0
2318	unfeasibly	0.00	1.00	0.00	0	1	0	1	unfeasible	0
2319	unnoticeably	0.00	1.00	0.00	0	1	0	1	unnoticeable	0
2320	unprintably	0.00	1.00	0.00	0	1	0	1	unprintable	0
2321	unstimulatedly	0.00	1.00	0.00	0	1	0	1	unstimulated	0
2322	vanishingly	0.00	1.00	0.00	0	4	0	4	vanishing	0
2323	vectorially	0.00	1.00	0.00	0	1	0	1	vectorial	0
*										
2325					28905	12265	59406			

C2のS1の調査

- SV [o D ... N] A-ly の事例から AN の対を獲得し、これから自動生成した25,181個のCQL
 - `[tag="V.*" & lemma!="be"] [tag!="V.*"]* [word=A] [tag="CJ.* | CRD | AV.* | AJ.* | N.*"]* [lemma=N | word=N]`
- を Sketch Engine の Python API で処理し、BNC中の実例数を入手
- 実例数が0でない場合 (9,715例) のみ、より制限の厳しいCQL
 - `[tag="V.*" & lemma!="be"] [tag="PR.*"]* [tag!="PR.*" & tag!="V.*" & tag!="AJ.*" & tag!="PU.*"] [word=A] [tag="CJ.* | CRD | AV.* | AJ.* | N.*"]* [lemma=N | word=N]`
- で再調査

C2のS1の調査

- 示唆されること
 - S1から自動的に生成されたANの圧倒的大部分
は、SV_の環境では容
認不可能

Freq	#Types	Rate
1以上	4109	16.32%
2以上	2333	9.26%
4以上	1242	4.93%
8以上	601	2.39%
16以上	263	1.04%
32以上	93	0.37%
64以上	32	0.13%
128以上	11	0.04%

Nの異なりの多いA(上位64語)

1. real (166)	18. specific (56)	35. individual (28)	51. precise (19)
2. early (128)	19. previous (51)	36. original (27)	52. safe (19)
3. direct (124)	20. separate (51)	37. considerable (25)	53. sudden (19)
4. actual (105)	21. correct (49)	38. personal (25)	54. initial (18)
5. simple (97)	22. effective (48)	39. independent(24)	55. light (18)
6. recent (96)	23. quick (48)	40. slow (24)	56. high (17)
7. particular (93)	24. large (47)	41. firm (23)	57. most (17)
8. immediate (90)	25. regular (45)	42. fair(22)	58. traditional (17)
9. proper (87)	26. main (41)	43. national (21)	59. formal (16)
10. general (84)	27. usual (39)	44. public (21)	60. possible (16)
11. entire (68)	28. special (38)	45. significant (21)	61. private (16)
12. local (68)	29. short (37)	46. bad (20)	62. rapid (16)
13. current (66)	30. successful (36)	47. even (20)	63. substantial (16)
14. complete (64)	31. normal (34)	48. final (20)	64. automatic (15)
15. serious (63)	32. strong (34)	49. total (20)	
16. clear (59)	33. exact (33)	50. efficient (19)	
17. different (56)	34. certain (28)		

Aの異なりの多いN (上位64語)

1. work (46)	18. process (20)	35. changes (16)	52. level (12)
2. people (45)	19. question (20)	36. market (16)	53. paper (12)
3. place (39)	20. business (19)	37. part (16)	54. programme (12)
4. things (31)	21. life (19)	38. school (16)	55. task (12)
5. system (29)	22. government (18)	39. word (16)	56. attention (11)
6. time (29)	23. issue (18)	40. course (14)	57. ball (11)
7. information (26)	24. line (18)	41. material (14)	58. food (11)
8. problem (24)	25. matter (18)	42. experience (13)	59. force (11)
9. hand (23)	26. money (18)	43. house (13)	60. group (11)
10. job (23)	27. policy (18)	44. party (13)	61. project (11)
11. man (23)	28. situation (18)	45. questions (13)	62. results (11)
12. point (23)	29. world (18)	46. scheme (13)	63. room (11)
13. way (22)	30. case (17)	47. services (13)	64. statement (11)
14. area (21)	31. data (17)	48. words (13)	
15. power (21)	32. position (17)	49. control (12)	
16. problems (21)	33. thing (17)	50. decision (12)	
17. service (21)	34. water (17)	51. effect (12)	

Nの異なりは1365

C1とC2の統合

◇	A	B	C	D	E
1	ADJ	Freq	N-types-in- type1-	TE- candidac	Check
5	just	6176	1	1	No
6	right	4059	3	1	No
7	happy	2585	7	1	Yes
12	like	1968	1	1	No
13	grateful	1485	1	1	Yes
15	ready	1363	1	1	Yes?
17	worried	1164	1	1	Yes
18	supposed	1135	3	1	No
20	wrong	1050	5	1	Yes?
24	proud	811	2	1	No
25	determined	775	1	1	Yes?
26	certain	770	28	1	No
28	most	708	17	1	No
30	free	670	14	1	No
31	over	664	1	1	No
32	pretty	655	1	1	Yes?
33	angry	652	3	1	Yes
1524	sequential	0	1	1	
1525	spatial	0	1	1	
1526	speculative	0	1	1	
1527	syntagmatic syntagmatical	0	1	1	
1528	three-dimensional	0	1	1	
1529	transverse	0	1	1	
2324					
2325					

C1とC2によるTE候補 1/2

- | | | | |
|----------------|-----------------|-----------------------|------------------|
| 1. happy | 15. serious | 29. wonderful | 43. enthusiastic |
| 2. grateful | 16. excited | 30. polite | 44. sympathetic |
| 3. worried | 17. quiet | 31. easy | 45. gentle |
| 4. angry | 18. sad | 32. brilliant | 46. impatient |
| 5. careful | 19. clever | 33. intelligent | 47. charming |
| 6. close | 20. desperate | 34. calm | 48. smart |
| 7. safe | 21. silent | 35. brave | 49. bitter |
| 8. kind | 22. quick | 36. attractive | 50. unfair |
| 9. conscious | 23. reluctant | 37. positive | 51. nasty |
| 10. anxious | 24. open | 38. critic critical | 52. consistent |
| 11. nice | 25. comfortable | 39. fair | 53. excellent |
| 12. stupid | 26. funny | 40. bright | 54. powerful |
| 13. nervous | 27. suspicious | 41. patient | 55. modest |
| 14. successful | 28. warm | 42. soft | 56. unfortunate |

C1とC2によるTE候補 1/2

1. ready

2. wrong

3. determined

4. pretty

5. willing

6. beautiful

7. honest

8. bad

9. clear

10. early

11. slow

12. dangerous

13. perfect

14. distracted

15. firm

16. reasonable

発展的問い

- C2の実態は何か？

交替条件C2の解明

支援動詞構文との交替

支援動詞構文 (SVC) とは?

- 支援動詞構文 (Support Verb Construction: SVP) とは次の特徴をもつ
 - 他動詞構文 $S V^* [O D M N]$ で
 - V^* が *have, take, make, give* のような抽象的動詞 (別名 light verb)
 - V^* が閉じた集合かどうかは不明
 - $O = D M N$ の主要部名詞 N が *argument, promise, look, idea, picture* のような (主に動詞派生の) 未飽和名詞
 - ただし N と本動詞 V^* の対応は複雑
 - 意味上の特徴
 - 構文の実質的な意味が N で決まる

SVCの例と交替(相当)形

- He had a harsh **argument** with her.
 - He **argued** with her harshly.
- He made a promise to marry.
 - He **promised** to marry.
- He had a close **look** at it.
 - He **looked** at it closely.
- He took nice **pictures**.
 - He **photographed** nicely.
(cf. *He **pictured** nicely)
- He had the **idea** of trying it out.
 - He **thought** of trying it out.
- Se gave him a grateful **smile**.
 - Se **smiled** to him gratefully.

SVCとA-ly形との交替

- SVCはOを修飾するMによって交替条件C2を満足する
 - He had a **harsh** argument with her. \Leftrightarrow He quarreled with her **harshly**.
 - he had a **close** look at it. \Leftrightarrow he looked at it **closely**.
 - He took **nice** pictures. \Leftrightarrow He took pictures **nicely**.
 - **pictures nicely** で niceなのは pictures であり, かつ he?

A-ly以外の交替形

- $SV [{}_0 D A N]$ が交替しうるのは $SV [{}_0 D N]$ A-ly のみではない。例えば
 - He walked a long walk. \Leftrightarrow He walked for a long time.
 - He read a fun book. \Leftrightarrow He read a book with fun.
- これは SVC を更に利用価値の高い構文にする

同族目的語構文との関連

同族目的語構文

- 同族目的語構文 (Cognate Object Construction: COC) の例:
 - She **smiled** a subtle **smile**.
 - He **dreamed** a beautiful **dream**.
 - She **lived** a happy **life**.
 - He **laughed** a harsh **laugh**
 - he **walked** a long **walk**.
- COCの実例は自動詞が多いが、他動詞がダメだということではない
 - They are **giving** gifts.
- 木原 (2011)を参照

COC と A-ly 構文との交替

- A-ly 構文と交替する場合
 - She smiled a **subtle** smile. \Leftrightarrow She smiled **subtly**.
 - She lived a **happy** life. \Leftrightarrow She lived **happily**.
 - He laughed a **harsh** laugh. \Leftrightarrow He laughed **harshly**.
- A-ly 構文との交替で意味が変わる場合
 - He dreamed a **beautiful** dream. \Leftrightarrow ?He dreamed **beautifully**.
- A-ly 構文以外との交替
 - They walked a **long** walk \Leftrightarrow They walked **for a long time**.
- 交替形がない場合
 - They dreamed **American** dreams. \Leftrightarrow ???

同族目的語構文モドキとの 関連

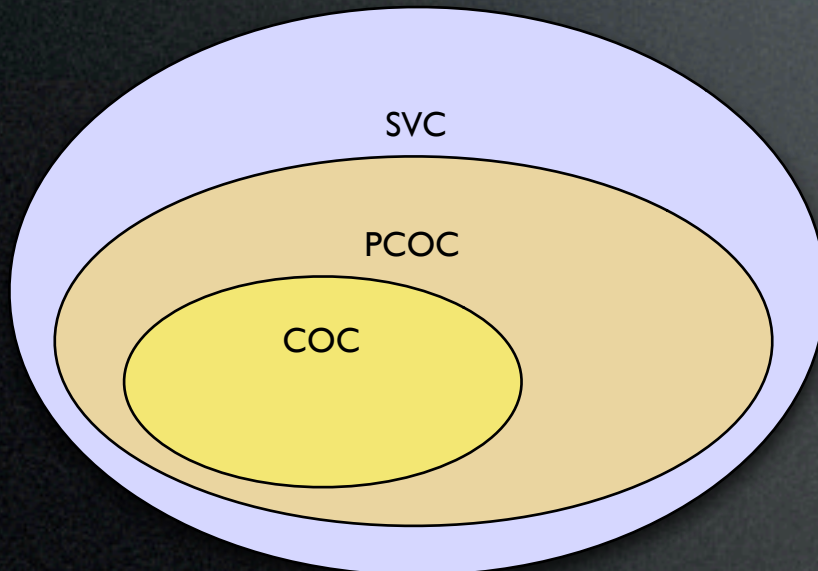
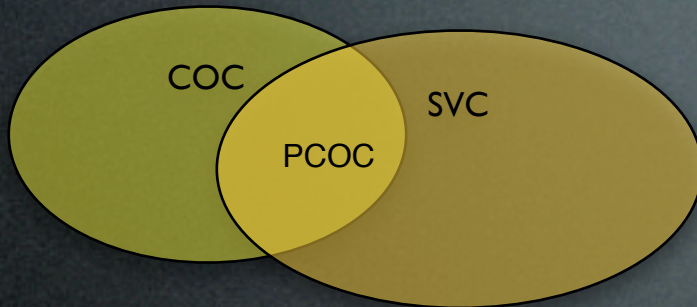
同族目的語構文モドキとは？

- 同族目的語構文モドキ (Pseudo Cognate Object Construction: PCOC) の例:
 - I feel like I found a “rare find” ... (Web)
 - He found a lucky find. (Web)
 - They found many interesting findings (Web)
 - you haven't saved a single save since the end of part 30. (Web)
 - There I discovered a shocking and unexpected discovery that fish oil lowers immunity. (Web)
 - he doctors explained a possible explanation being that the blood had pooled on the underside of the body after death ... (Web)
 - 10questions proved a valuable proof of principle, ... (Web)
- BNC級のコーパスでは実例は少ないが、Webにはかなり多くの実例が見つかる
 - Google で “found a * find” などの検索式を使えば見つけられる

PCOCの特徴

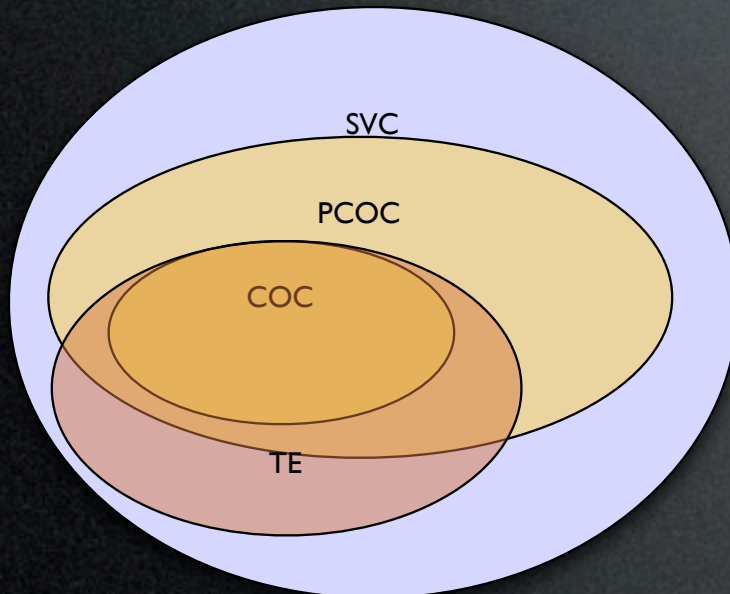
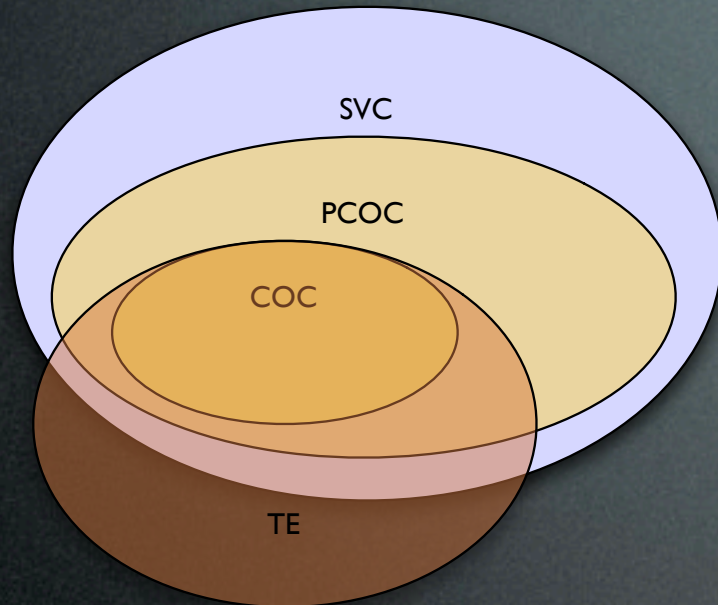
- $S V [{}_0 D M N]$ が同族目的語構文モドキ PCOC なら
 - 目的語名詞 N が文意の核となる項構造を決めているという意味で支援動詞構文 SVC の一種であり
 - 見かけの主動詞 V と N が派生的な関係にあるという意味で同族目的語構文 COC の拡張例である

SVC, COC, PCOCの関係



- SVC, PCOC, COCの関係を正しく表わしているのは, 上図と下図のどちらか?
- 私の見解では下図
 - PCOCは単なる SVC と COC の積集合ではない
- 含意
 - 従来のSVCの定義は $SVC - COC$ (か $SVC - PCOC$) の部分集合を排他的に指しているが, その定義では狭い

TECとSVC, (P)COCの関係 1/2



- TEとSVC, PCOC, COCの関係を正しく表わしているのは上図と下図のどちらか?
- TEはSVCの真部分集合か否か?
- COCがTEの部分集合であるのは確実
- この問題は未解決

TECとSVC, (P)COCの関係 2/2

- 動詞派生名詞 A, B であれ非動詞派生名詞 C であれ, 見かけの目的語名詞に固有の項構造を考えれば, 転移修飾構文 (TEC), 支援動詞構文 (SVC), 同族目的語構文(モドキ) ((P)COC) の三つを統一的に扱えるようになる
 - A. explanation, proof, gift, discovery, finding (名詞派生)
 - B. smile, laugh, dream, find (ゼロ派生: 動詞と名詞が同形)
 - C. idea (非動詞派生)
- しかし, そうしないと TEC, SVC, (P)COC の三つはバラバラのまま

発表後の補足 1/2

- 支援動詞構文 $S V [O D M N]$ の本質は文の中核となる意味構造=項構造が見かけの主動詞 V ではなく、見かけの目的語 O によって決まるどころ
 - これは N が動詞派生でなくても変わらない
 - SVC での本動詞 V の振る舞い実質的に助動詞と同じ
- 同族目的語構文 (モドキ) $(P) COC$ が支援動詞構文の一種であれば、同じことが同族目的語構文 (モドキ) についても言える
 - $(P)COC$ の特殊な点は支援動詞が軽い動詞 (e.g., have, take, make, give) でなく N と同じか派生の元になった動詞だということ
- 帰結
 - “ COC では自動詞が例外的に目的語を取る” という特徴づけは誤り

発表後の補足 2/3

- TECがSVCと関連づけられるのであれば， TECの認可は目的語にも依存する
- 実際， TECの例でも V と O の主要部には COCに似た “行為/道具 or 手段” の呼応関係がある
 - wave genial hands
 - plod a weary way
 - smoke a sad cigarette
- TEC の成立条件にC2が必須であるならば， TECの記述的に妥当な特徴づけは目的語名詞の特徴づけを必要とする

発表後の補足 2/3

- 今後の発展
 - S1, S2, S3 の全調査を終えた後で, TECの全実例で目的語名詞Nの粗い出しを行ない主動詞Vとの意味上の関連を見れば, NとVの間にどんなタイプ意味関係が必要か明確化できるだろう

まとめ

- コーパスの利用は思いがけない正例を見つけるのに非常に有益
 - 思いがけない正例の実例の一つが PCOCの実例
- 非典型的な表現はコーパス規模が大きくなると見つかる保証がない
 - sad cigarette の実例は Webの規模になってようやく見つかる
- どれぐらい大きなコーパスが必要なのか?に関して、見通しが得られているわけではない
- コーパスに負例の実例を見つけることは原理的に不可能
- しかし、神澤 (2011)と本研究が示したように、実例から負例の候補を生成し、それに人手検証を加えることでそれを実行するのは可能

謝辞

- 次の方からの助言やコメントが有益でした.
- 吉川 正人 (慶応大学大学院)

References

- Yoshikata Shibuya (2008). Transferred epithets: An MSFA approach. In 言葉と認知のメカニズム, pp. 129-144. ひつじ書房.
- Robert A. Hall (1973). The transferred epithet in P. G. Wodehouse." *Linguistic Inquiry* 4: 92-94.
- I. エアーズ (2007). その数学が戦略を決める (訳: 山形浩生), 文芸春秋社.
[原典: Ian Ayres (2007). *Super Crunchers: Why Thinking-by-Numbers Is the New Way to Be Smart*. Bantam.]
- 神澤克徳 (2011). 日本語の転移修飾への認知的アプローチ. Unpublished M.A. Thesis. 人間・環境学研究科, 京都大学.
- 木原恵美子 (2010). 構文現象への認知言語学的アプローチ: 同族目的語構文を中心に. Unpublished Ph.D. Thesis. 人間・環境学研究科, 京都大学.
- 黒田 航 (in press). 自作例を使った研究の基礎. In 認知言語学研究の方法 (中本敬子・李在鎬編). ひつじ書房.