琉球大学学術リポジトリ

「現代日本語書き言葉均衡コーパス」(BCCWJ)を利用した「くらい」・「ぐらい」の研究

メタデータ	言語: ja 出版者: 琉球大学法文学部国際言語文化学科欧米系 公開日: 2011-11-10 キーワード (Ja): キーワード (En): 作成者: 金城, 克哉, Kinjo, Katsuya メールアドレス: 所属:
URL	http://hdl.handle.net/20.500.12000/22129

言語文化研究紀要 *SCRIPSIMUS* No.20, 2011

「現代日本語書き言葉均衡コーパス」(BCCWJ) を利用した「くらい」・「ぐらい」の研究

金城　克哉

0. はじめに

2011年4月現在、大学共同利用機関法人人間文化研究機構国立国語研究所と文部科学省科学研究費特定領域研究「日本語コーパス」プロジェクトが共同で開発した「現代日本語書き言葉均衡コーパス」(BCCWJ: Balanced Corpus of Contemporary Written Japanese) という大規模コーパスの検索デモンストレーション版がインターネット上で利用できる。(注1) 本稿はこの検索機能を利用して副助詞／取り立て詞とされる「くらい」と「ぐらい」の用例を収集・分析し、従来看過されてきた両者の差異を明らかにすることを目的とする。

1. 先行研究と問題点

まず、『新明解国語辞典第四版』(2001)の記述を引き、意味・用法を概観する。この辞書では「くらい」には「副助詞的に」との注の下、次の4つの用法があるとされ、「『ぐらい』とも」との記述からうかがえるように、いずれの場合も「ぐらい」と交換可能であるとされている（下線は筆者）(注2)：

(1) [数量を表す語や基準を示す語に付いて]だいたいの見当を表す。
　① 茶さじ一杯くらいの塩を入れる
　② 駅まで十分くらいかかる
　③ 猫の額くらいの広さ
(2) 比較的理解しやすい基準を示しつつ問題になっている状況の程度を表す。
　① 棚の物が落ちたくらいの大きな地震だった
　② いくら忙しいからと言ったって、電話を掛けるくらいのひまはあっただろうに

③　そんな事を言ってくれるのは君くらいなものだ（＝君以外にはあんまり無い）

(3)　それに比べると、ほかの場合は物の数ではないという意を表す。

①　異国で病気をするくらい心細いことは無い

(4)　自分の主義・主張や性格・趣味を基準とすれば、前件よりもむしろ後件を選択することを表す。

①　途中でやめるくらいなら、やらない方がましだ

従来の研究は主に「くらい・ぐらい」と類似表現「ほど」の比較対照を通して「くらい・ぐらい」の意味・用法を浮かび上がらせるというものが主であり（日本語教育事典1982，奥津1986，長田・辻村1997，森貞2000・2001,田中2003)、またその扱いについては副助詞として一括して扱う立場（倉持1969，日本語教育事典1982，加藤1996）や形式名詞・形式副詞・とりたて詞という細分化をとる立場（沼田1986）などがある。どのようなデータをもとに議論を進めているかという観点からは、日本語テキストに現れた文をもとに考察を加えたもの（長田・辻村1997)、国定教科書の用例を探ったもの（加藤1996)、『新潮文庫の100冊』のような大規模コーパスからの用例を分析したもの（朱2010）などが見られる。これらの研究では多少の言及は見られるもののいずれも「くらい」と「ぐらい」の用法に何ら差異がないものとして記述・議論がなされている。

また、一般書の中には現在は「くらい」と「ぐらい」の区別はないとしながらも、以前には次のような使い分けが行われていたとして、用法上の区別に言及しているものもある（『NHKことばのハンドブック第2版』）(注3)：

(1)　体言には「ぐらい」が付く。

(2)　「この・その・あの・どの」には「くらい」が付く。

(3)　用言や助動詞には、普通は「ぐらい」が付くが、「くらい」が付くこともある。

日本語学習者用のテキストを分析し「くらい（ぐらい）」を論じたものでは長田・辻村（1997）が詳しいが、例えば『みんなの日本語』という初級用テキストでは第11課で「ぐらい」と「どのくらい」が別々の見出し語として導入され、「どの」は「ぐらい」とは連接しない、すなわち「どの」とのコロケーションでは「くらい」と「ぐらい」の用法に差がある、ということが暗黙の裡に前提とされている。しかしながら、実例に当たると、「どのぐらい」が全く用いられないわけではないということがわかる：

(1) すずめの胃袋ってどのぐらい食べたら満腹か分かりますか？？？毎朝、パンをちぎってあげてるんだけど、すぐになくなります。（Yahoo!知恵袋）

果たして、実際に「くらい」と「ぐらい」は全く同じように用いられているのであろうか。

従来の研究が（表だって明言はしないものの）前提としてきたように「くらい」と「ぐらい」は「言葉のゆれ」の一種であり、どちらを使用しても文の意味に影響することはない（言い換えれば両者の区別は無い（に等しい））ということが正しいのであれば、大規模コーパスを用いた調査において「くらい」と「ぐらい」の出現頻度はほぼ同じとなるに違いない。逆に、両者の出現頻度に差異が認められるのであれば、「くらい」・「ぐらい」それぞれの用法がどのような特徴を持つものであるのかが問われることになる。(注4)

2. 分析方法とデータ

今回は大規模コーパスを用いて「くらい」と「ぐらい」の差異を検証するため、インターネット上で利用できるBCCWJの検索デモンストレーション版を用いて調査及び用例収集を行った。このデモンストレーション版コーパスは、2011年3月現在、無作為抽出された11種のデータ、合計約1億480万語から成る（KOTONOHA 「現代日本語書き言葉均衡コーパス」ウェブサイトより）。各メディア毎の資料収集期間、件数、総語数は表1の通り。

今回は「くらい」と「ぐらい」というひらがなの文字列のヒット数（検索結果）を調べ、検索結果として表示された用例を収集した（後述するように、検索結果の表示には制限があるため必ずしも検索結果として表示されるヒット数全てが収集したデータ数とはならない）。今回の解析は、立命館大学の樋口耕一氏の開発したKH Coder（Version:2.beta.24）を利用した（解析器は茶筌）。

表1　各メディアの対象期間・件数・総語数

メディア／ジャンル	期間	件数	総語数
書籍	1971～2005年	21,943件	約6,230万語
雑誌	2001～2005年	1,989件	約440万語
新聞	2001～2005年	1,479件	約140万語
白書	1976～2005年	1,500件	約490万語
教科書	2005～2007年	412件	約90万語
広報紙	2008年	354件	約400万語
Yahoo!知恵袋	2005年	91,445件	約1,030万語
Yahoo!ブログ	2008年	52,680件	約1,030万語
韻文	1980～2005年	253件	約20万語
法律	1976～2005年	346件	約100万語
国会会議録	1976～2005年	159件	約510万語

3. 結果

以下、まずジャンル毎の出現数を示し、次にジャンルを問わず全体として「くらい」・「ぐらい」に前接する語及び指示詞にはどのようなものがあるかという共起関係の結果を示す。

3-1. 検索結果（ヒット数）

各ジャンル毎の検索結果（ヒット数）を表2に示す。表2では単に「くらい」と「ぐらい」というひらがな3文字の文字列の出現数を示したものである。そのため、「酒をくらい遊んだ」などの形態素解析がうまくいかないものもわずかながら含まれる。しかしそれを考慮に入れたとしても、「くらい」と「ぐらい」の出現数には明らかに差が認められる。(注5) 表2のトップ項目「書籍」か

ら9番目の「韻文」にいたるまで、出現数で「くらい」が「ぐらい」のほぼ1.5倍から3倍となっている。その一方、話し言葉の性格を強く残している「国会会議録」ではそれが逆転し、出現数では「ぐらい」が「くらい」のおよそ3倍となっている（国会会議録を話し言葉のデータとして扱うことについては後述）。

表2　各ジャンルと検索結果

メディア／ジャンル	くらい	ぐらい
書籍	14,063	7,072
雑誌	1,273	597
新聞	88	69
白書	32	16
教科書	118	75
広報紙	145	47
Yahoo!知恵袋2005	10,518	3,317
Yahoo!ブログ2008	5,036	1,736
韻文	27	7
法律	0	0
国会会議録	652	1,800
計	31,952	14,736

3-2. 前接語全体の頻度

「くらい」と「ぐらい」の頻度差をさらに詳しく見るために、前接する語を調べた。表3と4にそれぞれの結果を示す（記号を除く）。(注6) 表3では、例えば「くらい」は「どのくらい」としての用例が2,465例、「どれくらい」が1,694例あることが示されている。4位と39位に「ない」が現れるが、前者は否定の助動詞、後者は形容詞である。また、7位の「た」は、過去を表わす助動詞、17位の「だ」も助動詞である。以下それぞれの例をあげる：

表3「くらい」の前接語頻度順位表（頻度100以上）

順位	語	頻度	順位	語	頻度
1	どの	2,465	23	センチ	269
2	どれ	1,694	24	いくら	268
3	円	1,045	25	週間	260
4	ない	942	26	半分	246
5	年	847	27	時	235
6	同じ	831	28	こと	234
7	た	725	29	ヶ月	222
8	分	695	30	いい	192
9	歳	562	31	%	190
10	それ	551	32	ある	183
11	時間	469	32	いう	183
12	する	449	34	メートル	181
13	その	417	35	思う	180
14	人	410	36	いる	179
15	この	379	37	れる	147
15	回	379	38	ない	142
17	だ	361	39	キロ	137
18	これ	345	40	倍	130
19	日	324	41	前	119
20	度	312	42	個	110
21	cm	286	42	半年	110
22	なる	274	43	割	108

(2) 冠婚葬祭で使う喪服ってみなさんどのくらいでクリーニングに出しますか。 (Yahoo!知恵袋)

(3) 歌ってどれくらいの期間練習したらうまくなりますか？ (Yahoo!知恵袋)

(4) ラッシュ時は、携帯電話をいじることもできないくらい、周りの乗客と密着します。 (否定の助動詞,『読売新聞』2005年)

(5) 地面は燃え、足の踏み場もないくらいでした。 (形容詞「ない」,『広報まえばし』2008年)

(6) レールの両脇につまっている汚れは、雑巾でこすったくらいではびくともしません。　(助動詞「た」,『婦人之友』2003年)

(7) 特に労働の烈しい若い助修士など居眠りしなかったらその方が不思議なくらいだ 。　(助動詞「だ」,『フランス中世史夜話』)

表4　「ぐらい」の前接語頻度順表（頻度100以上）

順位	語	頻度	順位	語	頻度
1	年	680	17	同じ	201
2	円	593	18	これ	192
3	どれ	482	19	万	180
4	人	429	20	半分	161
5	分	417	21	する	160
6	歳	414	21	週間	160
7	どの	332	23	度	156
8	時間	283	24	いう	152
9	た	282	25	メートル	141
10	それ	271	26	センチ	140
11	%	268	27	割	136
12	いくら	264	28	キロ	125
13	回	259	29	その	112
14	ない	243	30	ｃｍ	108
15	日	235	31	少し	102
16	こと	224	32	この	101

表4では「くらい」の場合同様、「どれ」・「どの」が「ぐらい」と結びつく「どれぐらい」・「どのぐらい」が上位にあげられている。14位の「ない」」は否定の助動詞である。次に順位1位から5位までの例を挙げる：

(8) 商社マンだからといって、みなが英語がうまいとは限らない。特に入社三、四年ぐらいまではろくに喋れない者もいる。(『Domani』2001年)

(9) しみはしみ抜きに。しみの種類、大きさによって異なりますが三千円ぐらいからです。　(『ミマン』2001年)

(10) 食塩は、水にどれぐらいとけるのだろうか。実験１水にとける食塩の量を調べよう。（『新編新しい理科5下』2006年）

(11) 彼氏が友達７人ぐらいとよく飲みに行くんですがその中に女の人もいるみたいなんです。（Yahoo!知恵袋）

(12) 首里城を30分ぐらいの速さで見物。タクシーで首里駅。そこからモノレールで那覇空港へ。（Yahoo!ブログ）

表2の各ジャンルと検索結果は表3と表4における前接語の頻度にも影響を与えている。上述のように、ある日本語初級のテキストでは「どのくらい」が1つの連語として「ぐらい」とは別途導入されるということに触れたが、今回の調査では全体の中で「どのくらい」の頻度が突出して高いことが確認できた。しかしながら、「どのくらい」が出現するまさに同じコンテクストで「どれくらい」（「どのくらい」比68.72%）も現れうるということが示されている。その他興味深い点として、まず表2で示したように全体として「ぐらい」の頻度は「くらい」のおよそ3分の2から2分の1程度であるにもかかわらず、「年」という前接語では「～年ぐらい」の用例が「～年くらい」を上回っていること、また否定助動詞を用いて連体修飾とする「～ないくらい」がおよそ「～ないぐらい」の4倍の出現数となっている点があげられる。さらに、先行研究では「くらい」・「ぐらい」を形式名詞・取り立て詞・副助詞などとして扱っているが、形式名詞として連体修飾の「どの・その・この」とは高い頻度で共起するにもかかわらず、形容詞が前接語となっているケースが少ない点は注目してよい。

3-3. 指示詞との結びつき

次に上記頻度順位表から明らかになった指示詞と「くらい」・「ぐらい」の結びつきについて見てみよう。「くらい」では、前接要素の1位と2位に「どの」と「どれ」があげられていることから分かるように、この2つの疑問詞「どのくらい」・「どれくらい」の出現数は全体の約12.7%を占める。他の「こ・そ・あ」の指示詞とも共起するが、「あのくらい」・「あれくらい」の出現数

は「こ・そ」系に比べて著しく少ない。一方、「ぐらい」においても「どれ」・「どの」と共起する例が上位を占めているが、「どのぐらい」の出現数は「どのくらい」の約8分の1にとどまる。

表5「くらい」と指示詞

どの（どん）	2,357(14)
どれ	1,701
それ	610
その	433
これ	390
この（此の）	367(1)
あれ	20
あの	20

表6「ぐらい」と指示詞

どれ	410
どの（どん）	288(8)
それ	255
これ	187
この	97
その	95
あれ	15
あの	5

「ぐらい」と指示詞との共起関係で特徴的なのは、「それ」（255例）や「これ」（187例）よりも「その」・「この」との結びつきが弱く出現数が少ないことである。「あ」系は、「くらい」の場合同様、「あれ」・「あの」とも他に比して出現数が少ないが、全体的に見ても「このぐらい」・「そのぐらい」・「あのぐらい」は対応する「くらい」の場合に比して現れにくい傾向にあると言っていいだろう。次に最も用例の少ない「あのぐらい」の例をあげる：

⒀ 元日本語使いのボボにしてもいまでは「火星人にはあのぐらいしなきゃダメよ」とボクシングのポーズをとり、怒りをあらわにしている。

（『バンコクの罠』）

4. 考察

ここでは従来の研究でそれほど取り上げられることのなかった点を中心に、「くらい」と「ぐらい」の差についてさらに考察を加える。まず、「くらい」・「ぐらい」と比較・対照されてきた「ほど」とのかかわりについて見てみよう。

4-1. くだけた表現（口語的）か固い表現（文語的）か

辞書を始め先行研究では「くらい」と「ぐらい」はこれまで相互に交換可能であり、どのようなコンテクストであれ、いずれを用いても構わないという立場がとられてきた。仮にこの前提が正しいとすれば、コーパス調査において出現数や用法にほとんど差は見られないという結果になると考えられる。しかしながら、今回のコーパス調査ではいわゆる「書き言葉」において「くらい」の用例数が「ぐらい」の用例数を大幅に上回るということが明らかとなった。

先に「くらい」は類似表現「ほど」と比較対照され論じられてきたことに触れたが、意味にかぎらず、「くらい（ぐらい）」と「ほど」が交換可能なコンテクストでは「くらい（ぐらい）」がくだけた表現、「ほど」は固い表現であり、後者はどちらかと言えばフォーマルな文章に用いられるとされてきた。その一つの例として、「こんな」・「そんな」といった話者（もしくは書き手）の心情を反映する言葉との結びつきがある（今回の調査では「どんなぐらい」・「あんなぐらい」の用例は見られなかった）：

⑭ 私もそうなのですがそこばっかり気になるので、そんなぐらいは気にならないというのであれば入札してもいいのではないでしょうか？
（Yahoo!知恵袋）

⑮ 自動車学校って儲かる商売なんですか？教官の給料とかはどんなくらいですか？
（Yahoo!知恵袋）

「そんなぐらい」や「どんなくらい」の用例はあるが、「そんな」などと「ほど」が共起することはない。連体詞と（形式）名詞の共起は一般的に認められるものであるため、この共起制限は文法的な制約ではなく語用論的・文体的なものであると考えられる。それぞれの語は用法上重なる点もあるため、一概に論じることはできないが、このような制限が指し示すように、くだけた表現か固い表現かという区分で考えるならば、図1に示すような従来の2分法的な区分では十分とは言えない。「ぐらい」が「くらい」よりもより口語的であり、用法上重なる点があるという側面を考慮すると、イメージとしては図2のよう

にグラデーションを成しながら、文語的な「ほど」の用法と連続性を持つと考えたほうがよいのではないかと思われる(注7)：

図1　従来の「くらい（ぐらい）」と「ほど」の区分

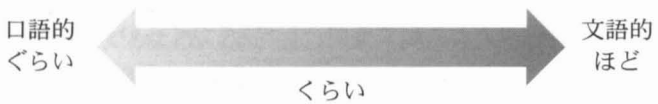

図2「ぐらい」「くらい」「ほど」の連続性

今回のコーパスに含まれる国会会議録以外に話し言葉のデータで分布に偏りが見られるのであれば、「ぐらい」のほうが「くらい」よりも口語的な表現であるという主張の傍証となる。そこで本稿では、『日本語話し言葉コーパス』サンプルデータを用いて検証を行った。(注7) その結果、「ぐらい」の総出現数10、「くらい」の総出現数2と、国会会議録同様に出現数に偏りが見られた(「0…L」の数字は『日本語話し言葉コーパス』固有のデータ番号)(注8)：

⑯ = 0397 00919.111-00922.337 L)

扉がないだけではなくて穴が掘ってあるだけぐらいのトイレにもチャレンジしたいなと

⑰ = 0015 00036.020-00038.874 L)

二十六歳か七歳かそれくらいなんですけど初めて中国に行くということになって

4-2. 疑問詞との結びつき

疑問詞「いくら」との共起（「いくらくらい」と「いくらぐらい」）では、指示詞で見られた傾向とは異なり、ほぼ同程度（「いくらくらい」275例と「い

くらぐらい」246例）用いられていることが分かった。しかし、「ぐらい」の出現数が全体で「くらい」の約2分の1であることに照らしてみると、「いくら」の場合は「ぐらい」と結びつく傾向が強いということが言える。

朱(2010)の調査では疑問詞に後接する場合は「いくら」・「どれ」・「何年」・「何回」に限られており、「いつ」・「だれ」との結びつきは無いとされる(朱2010:100)。しかしながら助数詞は「年」や「回」に限られるものではなく、今回のデータには「何泊」・「何歳」・「何秒」・「何人」・「何件」・「何機」・「何級」・「何台」など多くの「何＋助数詞」と「くらい」・「ぐらい」が共起する例が見られた。また、次に示すように、「いくつ」は「くらい」・「ぐらい」双方と共起するが、(「いくつぐらい」(44例)、「いくつくらい」(36例))、「いつ」に関しては「いつぐらい」(36例）があるのみで、「いつくらい」の例は見られなかった（後述の⒉も参照のこと）：

⒅ フリーのメールアドレスを作りたいのですが、いくつくらい作れるのでしょう？また、その方法を教えてください。 (Yahoo!知恵袋)

⒆ 私もあの系統のファッションが好きなのですが、彼女はおいくつぐらいの方かな？って思いました。 (Yahoo!知恵袋)

⒇ 新潟県の赤倉温泉や妙高のスキー場は、例年ではいつぐらいまで滑れますか？ (Yahoo!知恵袋)

4-3. 国会会議録と他のコーパスとの比較

4-3-1. 前接語の比較

BCCWJは国会会議録から159件（約510万語）の資料を収集しているが、この国会会議録は、松田（2008）が指摘するように国会でなされた発言がそのまま文字化されたものではなく、速記記号から漢字仮名交じり文にされ（これを「反訳」と言う），さらに校閲・編集部門にて誤りのチェックや字句の整理（「整文」と言う）・表記のチェック等のプロセスを経て公開されるが、この整文の段階でフィラーや語句・助詞等の脱落・省略の修正などが行われる(松田他2008)。これは自然談話の詳細な情報を含め文字起こしされた資料

(例えば会話分析などで扱うスクリプト等）とはおのずとその性格が異なるため、松田は国会会議録を「口語的特徴も多分に残しているが、書き言葉と話し言葉の中間的性格を持つものと位置づけるのが正しい」としている（松田2008：25)。本調査では限られた語彙のみに注目しているが、今回国会会議録における「くらい」・「ぐらい」の出現数が他の書き言葉と異なる分布を示したことから、国会会議録を話し言葉的要素が強いものと性格づけることも可能だと考える。

このような国会会議録の性格を踏まえたうえで、あらためて国会会議録と書籍や新聞その他に代表される書き言葉（Yahoo!知恵袋およびYahoo!ブログは除く）とを比較してみよう。(注9) 表7には「くらい」、表8では「ぐらい」について前接語を示す。

表7　国会会議録と書籍他の比較（くらい）

	国会		書籍その他	
1	どの	262	どの	1,086
2	どれ	77	どれ	733
3	この	38	ない	623
4	その	36	同じ	606
5	%	26	た	554
6	年	26	年	358
7	円	20	それ	352
8	いう	16	歳	327
9	これ	10	する	294
10	それ	10	分	292
11	ない	8	だ	276
12	人	7	円	275
13	いる	6	その	272
14	れる	6	人	251
15	思う	6	この	234
16	同じ	6	これ	214
17	割	5	時間	202
18	歳	5	こと	181

表7の「くらい」の比較からは、国会会議録では全651例中、指示詞、特に「どのくらい」での使用例が高い割合を占め（40.25％）、それに数量詞が続いていることがわかる。また連体修飾の「～ない」や「～た」を伴った「～ないくらい」や「～たくらい」、「する」動詞による連体修飾「～するくらい」などの用例は、書籍その他ではそれぞれ3位・5位・9位と高い頻度で出現しているが、国会会議録では逆に頻度が少ないことが読み取れる。

一方、表8では、国会会議録で15位までを助数詞、指示詞、動詞（「いう」のみ）が占めており、表7同様、「～ない」・「～する」による連体修飾は16例にとどまっている。書籍その他では「～た」・「～ない」・「～する」による連体修飾は多いものの、国会会議録に比べ「どのぐらい」（22位）や「その

表8　国会会議録と書籍他の前接語比較（ぐらい）

	国会		書籍その他	
1	年	173	年	341
2	％	163	歳	314
3	どの	153	人	250
4	円	138	分	207
5	どれ	113	円	200
6	人	96	た	199
7	割	70	こと	181
8	億	46	どれ	179
9	いう	40	それ	171
10	半分	36	ない	137
11	それ	34	時間	130
12	万	31	回	126
13	この	30	同じ	118
14	その	29	メートル	117
15	これ	26	日	117
16	倍	24	これ	101
17	年間	23	センチ	92
18	歳	21	度	92
19	一	19	いう	90
20	回	17	する	90

ぐらい」(29位)「このぐらい」(38位)の用例が少ない。特定の語に関して言えば、書籍その他では「同じ」という語を用いる場合、「同じくらい」の用例数(606例)が「同じぐらい」(118例)を上回っている(国会会議録では「同じぐらい」(16例)が「同じくらい」(6例)を出現数で上回っている)。

4-3-2. 後接続語の比較:助詞

次に国会会議録と書籍その他について、後接語の点から比較してみよう。表9と表10に後接語の頻度順位を示した。この2つの表からは助詞に差が見られることから、表12に格助詞のみに注目して「くらい」・「ぐらい」への後接語の頻度を調べた結果を示す:

表9 後接語の比較(くらい)

	国会		書籍その他	
1	の	202	の	3,334
2	に	89	だ	2,663
3	だ	77	に	1,221
4	ある	46	は	604
5	か	29	です	544
6	です	17	まで	346
7	は	15	しか	227
8	まで	11	ある	215
9	しか	7	か	197
10	かかる	6	前	162

表10 後接語の比較(ぐらい)

	国会		書籍その他	
1	の	534	の	1,800
2	に	271	だ	1,047
3	だ	168	に	674
4	は	133	は	517
5	です	72	で	341
6	が	68	です	276
7	で	62	まで	227
8	まで	57	か	224
9	と	49	しか	216
10	か	44	前	127

表11 「くらい」・「ぐらい」の後接助詞の比較

	国会		書籍その他	
後接助詞	くらい	ぐらい	くらい	ぐらい
の	206	534	3,567	1,817
に	89	271	1,239	660
が	13	68	208	124
を	4	34	80	39
と	3	49	55	46
へ	0	0	0	0

「くらい」・「ぐらい」双方とも「の」格を従え名詞句を形成する用法が最も多いことがわかる。頻度数ではそれに次いで「に」格が後接する例が続く。逆に、方向の「へ」はいずれの場合も用いられていない。さらに、主格「が」と目的格の比較では、主格との共起の方が目的格「を」の用法よりも多くなっている。また、国会会議録では「くらいと」及び「くらいを」の例が極端に少ない。

上述のように、後続に「の」が用いられる例が多いということは分かったが、それはさらにどのような言葉に連接していくのであろうか。朱（2010）でも「くらい」・「ぐらい」の後続要素について調査がなされているが、それが例えばどのような名詞句の一部を成すのかといった点については触れられていなかった。この点を示した表12を見てみよう。

大変興味深いことに、「～くらいのもの／こと／ところ」として不特定な物事や位置関係を示す用例が上位を占めるが、書籍その他には、「～くらいの大きさ」や「～ぐらいの大きさ」というように、特定の名詞「大きさ」を従えるケースがあり、これは国会の用例には見られない用法であった。次にその例を示す：

表12「～くらいの」・「～ぐらいの」の後接語

	国会				書籍その他			
	～くらいの		～ぐらいの		～くらいの		～ぐらいの	
1	もの	11	こと	28	こと	166	もの	109
2	こと	10	ところ	23	もの	158	こと	102
3	ところ	9	もの	18	大きさ	145	大きさ	62
4	数	8	間	18	時間	108	とき	41
5	規模	5	期間	13	ところ	60	ところ	41

(21) 軽い風邪をひいた子供では、これくらいの大きさのリンパ腺は腫れているのが、むしろ当たり前です。（『お母さんはホームドクター』）

4-4. 「くらい」・「ぐらい」の用法の変化

先行研究の概略および4－1でも触れたが、「くらい」や「ぐらい」の基本的

な用法は程度を表わすというものである。これらの語は基準となる具体的な数・量やそれに伴う助数詞とともに用いられたり（例：「茶さじ一杯くらいの塩」）、互いが了解できる具体的な事象や物事（人物を含む）を基準としてその程度であることを示したり（例：「電話をかけるくらいの暇はあっただろうに」）といった用法を柱とする。これらの用法は、具体的な高低・長短・軽重や程度が示せる場合であっても、またそういった具体性を欠いている場合であっても、「くらい」や「ぐらい」を用いる際には話者（書き手）には何らかの尺度が想定できている。

今回の調査では、このような「尺度上の程度」という「くらい」・「ぐらい」の基本的な意味とは異なる用法が散見された。「くらい」・「ぐらい」が「ほど」と似通った意味用法を持つことは従来から指摘されているが、次の用例では、「くらい」・「ぐらい」が表わす意味内容は「ほど」ではなく、「など」に近い：

(22) 「家事は家族まかせ」は「一部の家事または自分の身の回りのことくらいをしている」、「ほとんど家族まかせである」と回答した人の割合の合計。（『国民生活白書』平成15年版）

また、次の例では時間に幅を持たせる表現「頃」の意味として用いられており、「程度」を表しているとは言えない：

(23) 毎年、GWくらいに、ホームセンターに並びはじめるニューギニア　インパチェンス。今ぐらいから、10月くらいまでの半日陰の夏の花として植えて楽しんでいるんです。（Yahoo!ブログ）

さらに、形式名詞や取り立て詞であった「くらい」が次の例では形容動詞として用いられている：

(24) それにしても、この光景を道端から見るとかなりの大迫力だ。　藤沢の

駅で「小さいな」と感じた車両がここでは「象さんの大名行列」くらいな存在感である。　（『PHPほんとうの時代』）

これらは用例数も少ないため、単に誤用として扱うこともできよう。単なる誤用か、用法の拡大かは例えば10年といったまとまった単位での用例の経年変化を見る必要がある。

5. おわりに

本稿では従来の研究ではほとんど取り上げられることのなかった副助詞／取り立て詞の「くらい」と「ぐらい」について、現代日本語書き言葉均衡コーパス（BCCWJ）で用例の出現数及び前接・後接の語を検証することによって意味・用法は確かに似通っているが、出現パターンに相違があることが明らかとなった。国会会議録では書籍や新聞などに代表される書き言葉とは異なり、「ぐらい」の出現数が多く、また頻出のパターンでは「どの」と「くらい」の共起例が最も多いこと、「くらい」は後続の助詞としては「を」と共起しにくいこと、最後に用法に変化が見られることが示された。

国会会議録を話し言葉として扱うことについては議論の余地があるが、他の書き言葉とは異なるパターンが見られるということは注意しておいてよいだろう。今後は書き言葉のみならず、話し言葉のコーパスを用いてさらに詳細に調査・検討する必要がある。それについては今後の課題としたい。

註

1．2011年8月現在、KOTONOHA「現代日本語書き言葉均衡コーパス」の検索はhttp://www.kotonoha.gr.jp/shonagonで利用できる。

2．大修館書店の『明鏡国語辞典』には、「『くらい/ぐらい』は、今は区別なく使うが、伝統的には、体言には『ぐらい』、コソアド系の連体詞には『くらい』、活用語には『ぐらい』『くらい』ともに付いたという。」との記述がある。

3．書籍などの出版物は編集作業の間に何度か校閲を経ることになるが、講談

社校閲局編『日本語の正しい表記と用語の辞典』（第二版）では「～くらい・ぐらい」という見出しのあとに「二十年―」という記述があり、両者は区別されていない。一方、三省堂の『現代国語表記辞典』では、「くらい[位]」として、用例に「この（その）くらい」、参考として「これ（それ）ぐらい」をあげ、「この・その」の場合と「これ・それ」の場合で「くらい・ぐらい」の使い分けがあることが示されている。

4．本稿では当然ながら「くらい」と「ぐらい」は異形態であるという考え方はとらない。「くらい」と「ぐらい」が同一形態素の異形態であるならば、互いの出現パターンは相補分布を成し、各々出現条件が定められるが、「くらい」と「ぐらい」のケースはこれらの条件に当てはまらない。

5．同音異義の形態素の抽出割合は書籍のジャンルで多く見られたが、「くらい」の収集数13,457例に対し同音異義の形態素抽出は42例(0.312%)、書籍の「ぐらい」では7,065例中15例（0.212%）であった。

6．ウェブ上で表示される検索結果（ヒット数）は表1の通りであるが、このデモンストレーション版では実際の用例の表示が1回につき500件に限ら

別表1　各ジャンルと収集用例数

メディア／ジャンル	くらい	ぐらい
書籍	13,415	7,050
雑誌	1,273	597
新聞	88	69
白書	32	15
教科書	106	75
広報紙	145	47
Yahoo!知恵袋2005	9,226	3,315
Yahoo!ブログ2008	3,011	1,594
韻文	12	6
法律	0	0
国会会議録	651	1,800
計	27,959	14,568

れている。そのため、実際に収集できた用例は表1で示した数ではなく、次の別表1の通りである。今回はこのように収集数が検索結果と一致しないこと、また各ジャンルおよび総語数が2011年6月現在で概数でしかわからないために、カイ二乗検定のような仮説検定は行えなかった。表4と表5で示した前接する語の度数順位および本文中に示す用例数は「くらい」27,959例、「ぐらい」14,568例に基づく数値である。

7．ここでの指摘は「ほど」が必ずしも「固い表現」として用いられるということを意味するものではない。表1が示すように、「法律」の項目（これには憲法や法律の条文などが含まれる）には「くらい」も「ぐらい」も現れないが、かといって「ほど」であれば現れるかというと、そうではない。

8．ここで利用したものは『日本語話し言葉コーパス』（音声データで約660時間（短単位で約752万語））から2講演分の転記テキストと形態論情報を取り出したものである。当該サンプルデータの著作権は、大学共同利用機関法人人間文化研究機構国立国語研究所と独立行政法人情報通信研究機構が保持している。検索には『ひまわり』（ver.1.3β05、著作権は大学共同利用機関法人人間文化研究機構国立国語研究所）を用いた。

9．注6で触れたように、検索結果の表示数に制限があり、今回の調査ではヒット数がすなわち用例収集数とはならなかった。用例のとりこぼしは書籍の2001年～2005年のみで起こるため、このセクションでの国会会議録と書き言葉の比較では、書籍のデータに限り1971年から2000年までに限定した。

参考文献

NHK放送文化研究所編（2005）『NHKことばのハンドブック第2版』日本放送出版協会

奥津敬一郎・沼田善子・杉本武（1986）『いわゆる日本語助詞の研究』凡人社

奥津敬一郎（1986）「形式副詞」奥津敬一郎・沼田善子・杉本武編著『いわゆる日本語助詞の研究』凡人社，29-104.

加藤安彦（1996）「国定読本における副助詞『くらい』と『ぐらい』」『国立国

語研究所研究報告集』(17), 93-125.
倉持保男 (1969)「くらい (ぐらい) :副助詞<現代語>」松村明編『古典語現代語助詞助動詞詳説』學燈社, 516-518.
講談社校閲局編 (1992)『日本語の正しい表記と用語の辞典』(第二版) 講談社
朱武平 (2010)「『くらい (ぐらい)』の分布と意味の構文論的考察」『千葉大学人文社会科学研究』 (20), 99-110.
田中聡子 (2003)「『くらい』の意味的特徴:—『ほど』との比較を中心に—」名古屋大学日本言語文化専攻『言葉と文化』 (4), 277-292.
長田紀子・辻村俟子 (1997)「『ほど』と『くらい』の用法に関する考察」早稲田大学日本語研究教育センター『講座日本語教育』(32), 34-55.
中西久美子 (2005)「ナンカとクライ」日本語教育学会編『新版日本語教育事典』大修館書店, 158.
日本語教育学会編 (1982)『日本語教育事典』大修館書店
沼田善子 (1986)「とりたて詞」奥津敬一郎・沼田善子・杉本武編著『いわゆる日本語助詞の研究』凡人社, 105-225.
松田謙次郎 (2008)「国会会議録検索システム総論」松田謙次郎編『国会会議録を使った日本語研究』ひつじ書房, 1-32.
松田謙次郎・薄井良子・南部智史・岡田裕子 (2008)「国会会議録はどれほど発言に忠実か?—整文の実態を探る—」松田謙次郎編『国会会議録を使った日本語研究』ひつじ書房, 33-62.
森貞 (2000)「比較の基準を表す『より』『ほど』『くらい』について」『國語學』51(2), 173-174.
森貞 (2001)「『くらい』と『ほど』について」『國語學』52(1), 89-90.

Abstract

A comparative study of “kurai” and “gurai” based on BCCWJ

Katsuya Kinjo

This paper aims to reveal some aspects of the usages of "kurai" and "gurai" in Japanese based on analysis of Balanced Corpus of Contemporary Written Japanese (BCCWJ). The dictionary descriptions and previous studies on these words assume that "kurai" and "gurai" are interchangeable in any context. If this assumption is correct, their distributions in any corpus should be at the (almost) same ratio; however, the present study shows that this is not the case. The occurrence of "kurai" exceeds that of "gurai" in written materials, but the opposite is true in spoken data (the Diet Record). Moreover, the result of the analysis reveals that (i) "dono-kurai" was the co-occurring pairs of words with highest frequency, (ii) "kurai" is hardly followed by the objective case marker "o", and (iii) there are some usages which are close to "nado" rather than "hodo", which has been discussed as close to "kurai" and "gurai" in previous studies.