

神戸大学 国際コミュニケーションセンター
石川慎一郎研究室
Dr. Shin Ishikawa, Kobe University



KUBEE1850

小学校英語教育のための基本語彙表1850 (V1.0.3)

"KUBEE 1850"

Kobe University, Basic English Words List for Elementary School Students

Updates

2006/12/10

- ・ V1.0.3 公開
- ・ 誤混入語 (bon) を削除し, 1847語とする。
- ・ 上記は, 株式会社アイアールエルの馬岡伸行氏のご指摘に基づく修正です。ご教示に感謝申し上げます。
- ・ JACET8000レベル表示, および, 簡易意味領域タグを付記 (※※タグはバグフィックス作業がまだ進行中です。)

2006/12/6

・ 大学英語教育学会英語語彙研究会において, 神田外国語大学の長谷川先生をリーダーとする研究グループが, 「早期英語教育のための語彙リスト開発」というご発表において, 小学校英語語彙の基礎資料の一つとしてKUBEEをお使いいただきました。

2006/8/05

- ・ V1.0.1公開
- ・ 暫定版公開

2006/8/10

V1.0.2 公開

- ・ 日本語から誤って混入していた2語 (miso, ukiyoe) をカットし, 1848語とする。
- ・ 上記に伴う順位修正。

(注意)

本語彙表は, 科学研究費 (若手B) による研究成果物です。

著作権は作成者にありますが, 商用以外であれば, ご自由にお使いいただけます。

なお, ご利用いただく場合は, 事前に[石川まで](#)おしらせください。

ご意見, ご感想もあわせてお待ちしております。

(概要)

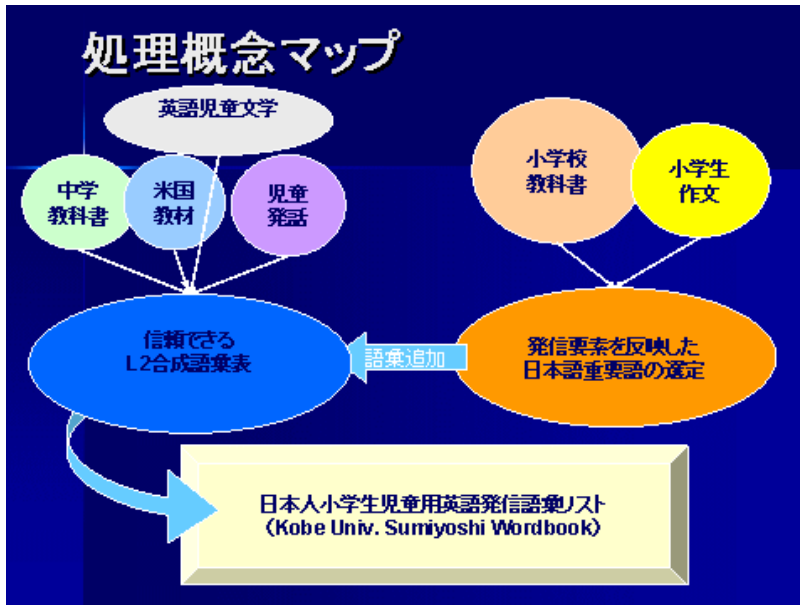
石川 (2005b), 石川 (in press) の枠組みに基づき, corpus-basedの手法で開発した語彙表です。
神戸大学発達科学部附属住吉小学校との連携プロジェクトをベースにしています。

本語彙表は, 日本の小学校英語教育という特殊な目的・方向性をふまえた<多層的L1/L2コーパスコンポーネント>を構築した上で, 目標言語としてのL2コーパスから抽出した語彙に, 日本人小学生にとっての母語であるL1コーパスの高頻度語を英語に変換して追加するというユニークな処理手法に基づいて開発されました。

KUBEE1850は、次のような特徴を持っています。

- 多様な小学校英語教育の実情に対応できる大型の基礎語彙資料となっていること
- 完全に1次言語資料 (attested language data) に準拠していること
- 客観的頻度データに準拠していること
- 日本人児童というユーザー環境をふまえていること
- 発信重視の語彙表であること
- すべての語に順位情報を付与しており、教育現場の実情に応じて、提示語彙量をコントロールできること (Top 500のみ使用, など)
- すべての語にPOSタグと意味タグを付与しており、現場のニーズに合わせたad hocな語選択ができること<現在, 最終調整中。近日公開>

処理手法の詳細は、石川 (in press) をご覧ください。



(C) 神戸大学石川慎一郎研究室

KUBEE1850

V1.0.3 (2006/12/15 release)

[Excel版ダウンロード](#) 最終確定版ではありません。バグフィックス中ですのでご注意ください。

- ・CLAWS品詞タグ, 意味タグは自動タガーの出力結果をそのまま記載したもので、現在検証作業中です。
- ・L2頻度, L1頻度, 確定頻度は、それぞれ、英語コーパスコンポーネントでの頻度, 日本語コーパスコンポーネントでの頻度 (対訳), 両者を調整した最終頻度を示します。なお, L1は補充用資料としての位置づけから, L2/L1の両方に出ている場合はL2頻度をそのまま利用し, L1のみに出ているものに限ってその頻度を最終頻度としています。
- ・JACET8000レベルは、大学英語教育学会 (JACET) 基本語リストにおけるレベルを示します。各レベルは1000語刻み (1レベルのみ1250語) で、1~8になるにつれて難しくなります。レベル9とは、JACET8000未収録語の意味です。
- ・第3列の「品詞」は、代表的英和辞典を確認して手作業で付与した品詞コードです。
- ・最終列の意味領域は、Tom McArthur(1981)のロングマンレキシコンの意味領域をふまえ、手作業で付与した意味コードです。
- ・手作業による品詞コード, 意味コードの付与に当たっては、神戸大大学院生の天谷詩美氏の助力を得ました。記して感謝申し上げます。

(A)の項目の一部)

ランク	語	品詞	CLAWS品詞タグ	L2頻度	L1頻度	確定頻度	JACET8000レベル	CLAWS意味タグ	Tom McArthur (1981)の区分に基づき手作業で付与した意味領域
5	A	Article	AT1	21269.5		21269.5	1	Z5	Z 機能語
739	ABLE	Adj	JJ	194.7		194.7	1	X9.1+	N 一般語
54	ABOUT	Prep	II	2967.1		2967.1	1	Z5	Z 機能語
791	ABOVE	Prep	II	180.7		180.7	1	Z5	Z 機能語
1277	ACCEPT	Verb	VV0	80.8		80.8	1	A9+	N 一般語
1592	ACCIDENT	Noun	NN1	54.7	200.1	54.7	1	Z3	N 一般語
591	ACCOMPANY	Verb	VV0		263.2	263.2	3	S3.1	M 運動・移動・位置
1170	ACCOUNT	Noun	NN1	96.4		96.4	2	I1	N 一般語
913	ACCUSTOMED	Adj	JJ		145.2	145.2	7	A6.2+	N 一般語
1560	ACHE	Verb	NN1	56.5	242.7	56.5	5	B2-	N 一般語
581	ACROSS	Prep	II	267.8		267.8	1	Z5	Z 機能語
1470	ACT	Verb	VV0	62.5		62.5	1	A1.1.1	N 一般語
382	ACTUALLY	Adv	RR	405.9		405.9	1	A5.4+	N 一般語
674	ADD	Verb	VV0	220.6	37	220.6	1	N5+	N 一般語
1264	ADDRESS	Noun	NN1	82.5		82.5	1	H4	N 一般語

475	ADJUST	Verb	VV0		323.1	323.1	3	A2.1+	N 一般語
1502	ADMIRE	Verb	VV0	59.9		59.9	3	S7.2+	F 感情と態度
1096	ADMIT	Verb	VV0	107.7		107.7	2	Q2.2	N 一般語
1559	ADULT	Noun	NN1		56.8	56.8	1	T3+	N 一般語
1516	ADVANCE	Verb	NN1	59.1		59.1	2	A9-	N 一般語
999	ADVENTURE	Noun	NN1	126.8		126.8	2	X8	N 一般語
1627	ADVICE	Noun	NN1	52.1		52.1	2	Q2.2	N 一般語
1545	AFFAIR	Noun	NN1	57.3	60.5	57.3	2	A1.1.1	N 一般語
425	AFRAID	Adj	JJ	363.9		363.9	1	E5-	F 感情と態度
160	AFTER	Prep	CS	965.9	42.2	965.9	1	Z5	Z 機能語
379	AFTERNOON	Noun	NNT1	406.8		406.8	1	T1.3	L 空間と時間
881	AFTERWARD	Adv	RL	153.7		153.7	3	N4	L 空間と時間
178	AGAIN	Adv	RT	861.5	111.6	861.5	1	N6+	N 一般語
478	AGAINST	Prep	II	320.5		320.5	1	Z5	Z 機能語
1054	AGE	Noun	NN1	115.5	335.5	115.5	1	T3	N 一般語

FAQ

株式会社アイアールエルの馬岡伸行氏より多くの貴重な指摘をいただきました。
 こうしたご意見が、語彙表の改訂、および新しい処理プロシージャの開発の原動力となります。記して感謝申し上げます。

下記に、FAQとしてご紹介させていただきます。
 なお、語の番号は古いバージョンに基づきますので、現行版では1～2番ほどずれています。

★1811 の BON ですが、これは「盆踊り」などの BON ではなく、BON JOURなどに用いられる BON でしょうか。
 ●bon danceからの混入です。最終処理で日本語起源は落としたのですが、bon (盆) については、ベースにした辞書で英単語としてエントリーされていたため、残ったものです。しかし、ご指摘のように、これだけが残ると、かえって不自然ですので削除いたします。(V1.0.3で対応済み)

★1010 の THY が小学生が知っておいてよい単語だとは思えないのですが、いかがでしょうか。
 ●ご指摘のとおりです。児童文学などでは非常に頻度が高くなり、結果としてここに残る程度の頻度になってしまいました。現代の英語コミュニケーションに反するという理由でフィルタを作って落とすか、あるいは、子供が身近にふれやすい物語の英語の実相として残すか、今後の検討課題といたします。

★1782 の PHOSPHORUS の使用頻度に意外感を持ちました。どのような使用例から採用されたものでしょうか。
 ●L2ではなく、L1からの補充語です。多くは小学校理科教科書からのエントリーです。

★604 の RECKON, 671 の REMARK, 1602 の BEHOLD は、もっとやさしい別の単語で代用できるといいますので、いらないのではないのでしょうか。
 ●これらは日本語コーパスから追加した語ではなく、英語コーパス（とくに児童文学）での高頻度語です。現場での語彙表の活用に当たっては、教育的判断でこういうものを削るという判断はありうると思いますが、今回は、頻度を基準に機械的な処理のアルゴリズムを作ることに研究の目的があり、主観的な調整は一切行っていません。ゆえに、ご指摘の語についても児童が触れる可能性のある英語に見られるの一つの実相として残しています。

★fixed がないのに 761 の FIXEDLY があるとか、fireplace がないのに1631 の HEARTH があるとか、convenience がないのに1736 の INCONVENIENCEがあるというのも違和感を感じました。(「コンビニで買った」という文を考えると、convenience という単語があってもよいのではと思った次第です。)

●頻度に基づく機械的な処理で作成した語彙表ですので、派生語の調整（派生語の基本形をエントリー語彙に入れる）や、同義語の調整（同義語でより平易なものが入っているかどうかを確認する）は行っていません。
 ご指摘のfixedlyについてですが、「しっかりと・きっちりと」は日本語コーパスからの対訳補充語です。この語は日本語では副詞の形で出てくるもので、そのまま副詞の形で補充されました。
 hearthは児童文学からの混入です。
 ご指摘のコンビニは、意外ですが、今回使用した小学生の作文コーパスでは基準値以上には出現しませんでした。

★1692 の A___ は、小学生は知ったら大喜びかもしれませんが、先生方は取り上げる気にはならないように思いました。
 ●確かにご指摘のとおりです。ただ、客観的な処理の結果として残った語ですので、今回はそのままにしています。ただし、f***のような語は、「卑語・俗語フィルター」というのを当方でつくり、この基準でカットしました。

[Back to top](#)