

『多言語母語の日本語学習者横断コーパス』語数表 20180521 版 解説

1. データの概要

本資料は、『多言語母語の日本語学習者横断コーパス』（以下、I-JAS とする）の語数表について解説するものである。

語数表にはタスク別に調査協力者（K）、調査者（C）それぞれについて集計した語数を掲載している。ストーリーライティングなど、タスク内に調査者の発話が全くない場合は、語数表には掲載されていない。

語数は、基本的に UniDic の品詞体系に基づき、自動で形態素解析（短単位）を行った結果を集計しているが、一覧表には、「語数（全て）」と「語数（記号等除外・全て）」の 2 種類がある。「語数（全て）」には、記号等もすべて含め集計しているが、「語数（記号等除外・全て）」は記号等を除外して集計している。また、語数表には、語数に加え、協力者の属性（データセット、調査地、母語、年齢、性別、身分）も掲載している。

2. 集計方法

(1)基本的に UniDic の品詞体系に基づき、形態素解析（短単位）を行った。

(2)(1)の形態素解析で得られた書字形を集計した。この際、データは、以下の 2 種類について集計した。

- ・語数：記号類なども含め、すべての書字形を集計したもの。
- ・語数（記号等除外）：品詞が「補助記号」「記号」「空白」「非言語行動」となっているものを除外して集計したもの。

※一部の感動詞については、学習者特有のものがあるため、UniDic の品詞体系とは異なる場合がある。このようなものについては、I-JAS 独自の方針で品詞を付与した。詳細は別紙の通りである。

3. 語数表の見方

- ・ファイル名：I-JAS_語数表_20180521 版
- ・サンプル ID：学習者、調査協力者、それぞれの発話について、語数を集計している。
- ・協力者の属性：語数以外に、「母語」「年齢」「性別」「身分」「J-CAT の得点」「SPOT の得点」の情報を掲載している。調査の段階で、学習者に回答を拒否されたものは「未回答」となっている。また、「身分」は「学生」と「非学生」に区分されている。

4. 利用上の注意

- (1) 研究、教育目的であれば無償で自由に利用できる。『中納言』のユーザー登録と『I-JAS』

の利用申請が必要である。

- (2) I-JAS のデータの再配布は不可。商業使用（営利目的での利用）は要相談。
- (3) 論文等に語数表を引用する際は、出典とバージョンを明記すること。以下に、出典とバージョンの例を示す。
例) 『多言語母語の日本語学習者横断コーパス』「中納言」バージョン 2.4.2
短単位データ 20180502 版 I-JAS 語数表_20180521 版
- (4) 本データの著作権（編集著作権）は国立国語研究所が有する。
- (5) データの瑕疵による損害についてはいかなる場合でも補償しない。
- (6) 内容の改善のため、予告なくデータを更新することがある。

本データに関する問い合わせ先：lsaj[@]ninjal.ac.jp （[@]を@に変えること）

以上

【別紙】

表 1. I-JAS で付与した感動詞の書字形と語彙素・品詞の一覧

書字形	語彙素	品詞
あ	あっ	感動詞-一般
ああ	ああ	感動詞-一般
あー	あー	感動詞-フィラー
あああ	ああ	感動詞-一般
ああー	ああ	感動詞-一般
あーっと	null	感動詞-フィラー
あーっとー	null	感動詞-フィラー
ああと	null	感動詞-フィラー
あーと	あーと	感動詞-フィラー
ああとー	null	感動詞-フィラー
あーとー	あーと	感動詞-フィラー
あーむ	null	感動詞-フィラー
あーん	ああん	感動詞-一般
あい	あい	感動詞-一般
あいー	null	感動詞-フィラー
あう	null	感動詞-フィラー
あうー	null	感動詞-フィラー

あっ	あっ	感動詞-一般
あっと	あーと	感動詞-フィラー
あっとー	あーと	感動詞-フィラー
あと	null	感動詞-フィラー
あとー	null	感動詞-フィラー
あの	あの	感動詞-フィラー
あのー	あの	感動詞-フィラー
あは	あは	感動詞-一般
あむ	null	感動詞-フィラー
あん	あん	感動詞-一般
あんー	null	感動詞-フィラー
い	いー	感動詞-一般
いー	いー	感動詞-フィラー
いーん	null	感動詞-フィラー
いえ	いえ	感動詞-一般
いや	いや	感動詞-一般
いやー	いや	感動詞-一般
う	うう	感動詞-一般
うい	null	感動詞-フィラー
うー	うー	感動詞-フィラー
うーっと	null	感動詞-フィラー
うーと	null	感動詞-フィラー
うーとー	null	感動詞-フィラー
うーん	ううん	感動詞-一般
うーんと	うんと	感動詞-フィラー
うふ	うふふ	感動詞-一般
うわー	うわ	感動詞-一般
うん	うん	感動詞-一般
うんと	うんと	感動詞-フィラー
うんー	うん	感動詞-一般
え	えっ	感動詞-一般
えいー	null	感動詞-フィラー
えうー	null	感動詞-フィラー
ええ	ええ	感動詞-一般
えー	えー	感動詞-フィラー

ええーっと	null	感動詞-フィラー
ええーと	null	感動詞-フィラー
えーっと	えーと	感動詞-フィラー
えーっとー	えーと	感動詞-フィラー
ええと	えーと	感動詞-フィラー
えーと	えーと	感動詞-フィラー
ええとー	えーと	感動詞-フィラー
えーとー	えーと	感動詞-フィラー
えーむ	null	感動詞-フィラー
えーむー	null	感動詞-フィラー
えーん	null	感動詞-フィラー
えっ	えっ	感動詞-一般
えっと	えーと	感動詞-フィラー
えっとー	えーと	感動詞-フィラー
えと	えーと	感動詞-フィラー
えとー	えーと	感動詞-フィラー
えの	null	感動詞-フィラー
えむ	null	感動詞-フィラー
えむー	null	感動詞-フィラー
えん	null	感動詞-フィラー
えんー	null	感動詞-フィラー
お	おっ	感動詞-一般
おお	おお	感動詞-一般
おー	おー	感動詞-フィラー
おおん	null	感動詞-フィラー
おーん	null	感動詞-フィラー
おん	null	感動詞-フィラー
おんー	null	感動詞-フィラー
じゃあ	null	感動詞-フィラー
そう	null	感動詞-フィラー
と	と	感動詞-フィラー
とー	と	感動詞-フィラー
のー	null	感動詞-フィラー
はあ	はあ	感動詞-一般
はー	はあ	感動詞-一般

ふーん	ふん	感動詞-一般
ふん	ふん	感動詞-一般
ふんー	null	感動詞-フィラー
へえ	へえ	感動詞-一般
へー	へえ	感動詞-一般
ほー	ほう	感動詞-一般
ま	まあ	感動詞-一般
まあ	まあ	感動詞-一般
まー	まー	感動詞-フィラー
や	null	感動詞-一般
やー	null	感動詞-一般
わ	わあ	感動詞-一般
わー	わあ	感動詞-一般
ん	んっ	感動詞-一般
んっとー	うんと	感動詞-フィラー
んと	うんと	感動詞-フィラー
んとー	うんと	感動詞-フィラー
んふー	null	感動詞-フィラー
んん	んっ	感動詞-フィラー
んー	んー	感動詞-フィラー
んーと	うんと	感動詞-フィラー
んーとー	うんと	感動詞-フィラー
んーふ	null	感動詞-フィラー
んんー	null	感動詞-フィラー
んーん	うん	感動詞-一般