

●第6章 120ページ web 資料

演習1

表 6.6

データ作成の手順

- ① 「中納言」で『現代日本語書き言葉均衡コーパス通常版』を選択する。
- ② 「長単位検索」で、下記の検索をおこなう。
(キー) 品詞の大分類が「名詞」 + (後方共起1) 語彙素「の」 +
(後方共起2) 語彙素「発生」
- ③ ダウンロードしたデータを Excel で開き、ピボットテーブルを用いて頻度表を作成する。

●第6章 121ページ web 資料

表 6.7

データ作成の手順

(まず「的發生」の直後に生じる1文字で行い、それが終わったら2文字、3文字でも同様の作業を行う。)

- ① 「北京大学中国语言学研究中心」の「在线资源」から「语料」に進み、『现代汉语语料库』を選択する。
- ② 検索画面が出てきたら、「的發生」を検索し、検索結果をダウンロード（下载）する（その際、「下载」の横の件数を多めに設定すること）
- ③ Excel を立ち上げ、「ファイル」→「開く」→「デスクトップ」（ダウンロードしたデータがある場所）→右下の「すべての Excel ファイル」を「すべてのファイル」に変更し、当該のデータを開く。
- ④ 「的發生」を含む列（A列）をコピーして、サクラエディタに貼り付ける（ここで1行目に「1文字」と書いておくと、後の作業の役に立つ）。
- ⑤ 「検索」→「置換」で「的發生」の前後の「[]」を消す。

置換前：[的發生]

置換後：的發生

- ⑥ 「検索」→「置換」で「正規表現」にチェックを入れ、「的發生」と直前の1文字のみにする（「的發生」の前の「.」を2つにすると、「的發生」と直前の2文字になる）。

置換前： $^.*(.的發生).*$

置換後： $\$1$

- ⑦ 「的發生」を削除する。

置換前：的發生

置換後：

- ⑧ ⑦で処理したデータに名前をつけてデスクトップに保存する（その際に「文字コードセット」を「UTF-8」にする）。
- ⑨ Excel を立ち上げ、「開く」→「デスクトップ」→右下の「すべての Excel ファイル」を「すべてのファイル」に変更して⑧のファイルを開く。

⑩ ピボットテーブルを用いて頻度表を作成する。

●第6章 122ページ web 資料

演習2

表 6.8

データ作成の手順

- ① 「中納言」で『現代日本語書き言葉均衡コーパス通常版』を選択する。
- ② 「長単位検索」で、下記の検索をおこなう。
(キー) 品詞の大分類が「動詞」+ (後方共起1) 語彙素「た」+
(後方共起2) 語彙素「ばかり」
- ③ ダウンロードしたデータをExcelで開き、ピボットテーブルを用いて頻度表を作成する。
- ④ 「～たばかりか」「～たばかりで（は）なく」「～たばかりに」といった直後を表すとは言えないものを除外する。

表 6.9

データ作成の手順

(まず「剛」の直後に生じる1文字で行い、それが終わったら2文字、3文字でも同様の作業を行う。)

- ① 「北京大学中国语言学研究中心」の「在线资源」から「语料」に進み、『现代汉语语料库』を選択する。
- ② 検索画面が出てきたら「剛」を検索し、検索結果をダウンロード（下载）する（その際、「下载」の横の件数を多めに設定すること）。
- ③ Excelを立ち上げ、「ファイル」→「開く」→「デスクトップ」（ダウンロードしたデータがある場所）→右下の「すべてのExcelファイル」を「すべてのファイル」に変更し、当該のデータを開く。
- ④ 「剛」を含む列（A列）をコピーして、サクラエディタに貼り付ける（ここで1行目に「1文字」と書いておくと、後の作業の役に立つ）。
- ⑤ 「検索」→「置換」で「剛」の前後の「[]」を消す。

置換前：[剛]

置換後：剛

- ⑥ 「検索」→「置換」で「正規表現」にチェックを入れ、「剛」と直後の1文字のみにする（「剛」の後の「.」を2つにすると、「剛」と直後の2文字になる）。

置換前： $^.*(\text{剛}).*$

置換後：\$1

- ⑦ 「剛」を削除する。

置換前：剛

置換後：

- ⑧ ⑦で処理したデータに名前をつけてデスクトップに保存する（その際に「文字コードセット」を「UTF-8」にする）。

- ⑨ Excel を立ち上げ、「開く」→「デスクトップ」→右下の「すべての Excel ファイル」を「すべてのファイル」に変更して⑧のファイルを開く。
- ⑩ ピボットテーブルを用いて頻度表を作成する。

表 6.10

データ作成の手順

(まず「刚刚」の直後に生じる 1 文字で行い、それが終わったら 2 文字、3 文字でも同様の作業を行う。)

- ① 「北京大学中国语言学研究中心」の「在线资源」から「语料」に進み、『现代汉语语料库』を選択する。
- ② 検索画面が出てきたら「刚刚」を検索し、検索結果をダウンロード（下载）する（その際、「下载」の横の件数を多めに設定すること）。
- ③ Excel を立ち上げ、「ファイル」→「開く」→「デスクトップ」（ダウンロードしたデータがある場所）→右下の「すべての Excel ファイル」を「すべてのファイル」に変更し、当該のデータを開く。
- ④ 「刚刚」を含む列（A 列）をコピーして、サクラエディタに貼り付ける（ここで 1 行目に「1 文字」と書いておくと、後の作業の役に立つ）。
- ⑤ 「検索」→「置換」で「刚刚」の前後の「[]」を消す。

置換前：[刚刚]

置換後：刚刚

- ⑥ 「検索」→「置換」で「正規表現」にチェックを入れ、「刚刚」と直後の 1 文字のみにする（「刚刚」の後の「.」を 2 つにすると、「刚刚」と直後の 2 文字になる）。

置換前： $^*(刚刚).*$

置換後：\$1

- ⑦ 「刚刚」を削除する。

置換前：刚刚

置換後：

- ⑧ ⑦で処理したデータに名前をつけてデスクトップに保存する（その際に「文字コードセット」を「UTF-8」にする）。

- ⑨ Excel を立ち上げ、「開く」→「デスクトップ」→右下の「すべての Excel ファイル」を「すべてのファイル」に変更して⑧のファイルを開く。

- ⑩ ピボットテーブルを用いて頻度表を作成する。