

言語学 A 第 1 章: 日常に見られる言語

1 言語学の教科書を見て答える

問 1 言語学とはどんな学問であるのかを話し合え。

問 2 教科書の中から言語学の定義らしい文を探せ。教科書をまだ入手していない人は、つぎのことばを参考にして考えよ。

- 人間に関する重要な問題の中には、言語と切っても切れない関係にあり、言語の機構が分かればその解決に実質的に役立つと思われるほど言語から深い影響を受けているものがある。さらに、人間活動ほどの面をとっても研究する価値がある。従って、実用的な理由からも、また人間の生来の好奇心を満足させるためにも、言語は慎重で知的な研究に値するものである (Gleason 1970: 3)。
- 言語学 (げんごがく、英: linguistics) は、人間の言語の特性、構造、機能、獲得、系統、変化などを研究する学問。下位分野として、音声学、音韻論、形態論、統語論 (統辞論)、意味論、語用論などの様々な分野がある。これらの下位分野は、(表出) 音 (手話言語の場合はジェスチャー)、音素、語と形態素、句と文、意味、言語使用に概ねそれぞれが対応している (Wikipedia)。

問 3 自分の考える言語学と教科書の説明との違いについて話し合え。

問 4 テキスト (田中・田中 2015) 中、pp. 2-5 の馴染みない単語の意味を確認せよ。

問 5 ことばの研究するのなら、どんなテーマがよいかを考えよ。もちろん、テーマは途中で変えても良い。

問 6 テーマが決まったら、リサーチクエスション (RQ) を作れ。RQ とは、学術領域において、自分の関心となる疑問について、調査・実験の結果、Yes か No かで答えられる疑問文である。

2 新しい言い方と古い言い方

問 7 ビデオを見て、新しい言い方あるいは古い言い方だと思うものを指摘し、意見を交換せよ。

問 8 なぜそれが古い言い方だと思う。また、その古い言い方は現代でも使われているか。

問 9 現代でも使われている古語を探し、その語を使うとどんな感じがするか、どんな時に使われているかを話し合え。

問 10 なぜ現代でも古い語が使われているのかを考えよ。

問 11 どれぐらいの時間が経過したら、古い言い方だと私たちは感じるのだろうか。

問 12 このように言語の経年変化を分析する言語学の領域を何というか、調べてみよ。

3 単語のカウント

問 13 あるテキストに含まれる単語の数を数えたい。単語の数え方について自分なりの方法を考えよ。

問 14 動詞、名詞、副詞、助詞など品詞別に単語のカウント方法を考えよ。

問 15 ある名詞、副詞、助詞など品詞別に単語のカウント方法を考えよ。

Listing 1 How to count words

```
1 % cat inoue.txt | mecab | grep "名詞" | wc
```

- MeCab (和布蕪) とは
- 形態素解析システム MeCab のページ

4 オンラインだけでできるテキストマイニング

4.1 テキストマイニングツール

User Local テキストマイニングツール

<https://textmining1.userlocal.jp/home/file>

このサイトでワードクラウドを作成することができる。

スコアが高い単語を複数選び出し、その値に応じた大きさを図示しています。単語の色は品詞の種類で異なっており、青色が名詞、赤色が動詞、緑色が形容詞・形容動詞、灰色が感動詞を表しています。

4.2 統計検定ツール χ^2 検定

問 16 2つのテキストが質的に違うことを数値で述べるにはどうすればよいか、考えよ。

問 17 次の2作家のことばの使い方を比較するにはどうすればよいかを考えよ。

- 井上ひさし (いのうえひさし、1934–2010) 日本の小説家、劇作家、放送作家、文化功労者、日本藝術院会員、自らを遅筆堂と名乗る。代表作品に「モッキンポット氏の後始末」「吉里吉里人」(日本 SF 大賞受賞作)「手鎖心中」などがある。
- 筒井康隆 (つついやすたか、1934–) 日本の小説家、劇作家、俳優。小松左京、星新一と並んで「SF 御三家」とも称される。代表作品に「時をかける少女」「日本以外全部沈没」「文学部唯野教授」などがある。

問 18 井上ひさし、筒井康隆、各作品で動詞、名詞比率が同じかどうかを調べよ。

問 19 Easy Chi-Square Calculator [link] を使って、 χ^2 検定を実施してみよ。

問 20 どんな比較をするとおもしろい結果が出るかを考えよ。



5 発展

問 21 $m \times n$ の分割表による χ^2 検定をしてみよ。

問 22 他の品詞や文の長さを計測・比較し、有意な差があるかどうかを確認せよ。

問 23 興味のあるのテーマを選び、リサーチ・クエスチョン (Research question: 研究上、追求すべき問題。Yes か No で応えられるような簡素な答が結論となるような問題文。「可能性があるかどうか」のような存在の有無でよいが、具体的に「.. であることが根拠で.. の可能性があるかどうかを検討する」のような文が望ましい) を作れ。

問 24 MVR という指数がある (樺島・寿岳 1965: 122–123)。MVR を計算し、2つの文章を比較してみよ。

$$MVR = \frac{\text{形} + \text{形動} + \text{副詞} + \text{連体詞}}{\text{動詞}} \times 100$$

問 25 MVR の数値が高いとき、低いとき、それぞれどんなことを意味するかを考えよ。

参考文献

- Gleason, Jr., Henry Allan (1970) 『記述言語学』, An Introduction to descriptive linguistics, 大修館書店, 東京, Japan.
- 樺島忠夫・寿岳章子 (1965) 『文体の科学』, 綜芸舎.
- 田中春美・田中幸子 (2015) 『よくわかる社会言語学』, やわらかアカデミズム・「わかる」シリーズ, ミネルヴァ書房.

Chi-Square Calculator

Note: You can find further information about this calculator, [here](#).

This is a chi-square calculator for a simple 2 x 2 contingency table (for alternative chi-square calculators, see the column to your right).

The first stage is to fill in the group and category information. So, for example, if you have collected data on smoking habits, and want to find out whether smoking is related to gender, you could have "Male" and "Female" as the two groups, and "Smoking" and "Non-Smoking" as the two categories.

	筒井康隆	井上ひさし	
動詞			
名詞			

Please enter group and category values.

A.

Chi-Square Calculator

The next stage is to fill in your values. Remember, the data is categorical - the number of subjects observed for each cell (for example, Male Smokers, Male Non-Smokers, Female Smokers, Female Non-Smokers). If you go wrong, you will get a chance to edit your data at the next stage.

	筒井康隆	井上ひさし	
動詞	21892	10515	
名詞	21624	14254	

Please enter data values for your categorical variables.

B.

Chi-Square Calculator

Okay, we've now set up the 2 x 2 contingency table, and we're almost ready to do the chi-square calculation. However, before you hit the "Calculate" button, you need to select a significance level. It defaults to .05, but you can choose .01 or .10 if you prefer. You should also take a moment to check your data, and make any changes you require by clicking "Edit".

	筒井康隆	井上ひさし	<i>Marginal Row Totals</i>
動詞	21892	10515	32407
名詞	21624	14254	35878
<i>Marginal Column Totals</i>	43516	24769	68285 (Grand Total)

Significance Level:

- .01
 .05
 .10

Remember, if you're ready to make the calculation, then you need to select a significance level.

C.

Chi-Square Calculator

Success! The contingency table below provides the following information: the observed cell totals, (the expected cell totals) and (the chi-square statistic for each cell).

The chi-square statistic, *p*-value and statement of significance appear beneath the table. Blue means you're dealing with dependent variables; red, independent.

	筒井康隆	井上ひさし	<i>Marginal Row Totals</i>
動詞	21892 (20652.02) [74.45]	10515 (11754.98) [130.8]	32407
名詞	21624 (22863.98) [67.25]	14254 (13014.02) [118.15]	35878
<i>Marginal Column Totals</i>	43516	24769	68285 (Grand Total)



If you've found this tool useful, please consider sharing it!

The chi-square statistic is 390.6453. The *p*-value is . This result is significant at $p < .05$.

D.

図1 井上作品、筒井作品の品詞による χ^2 検定: 両作品の品詞を数え、その比率を χ^2 検定してみると有意であった。