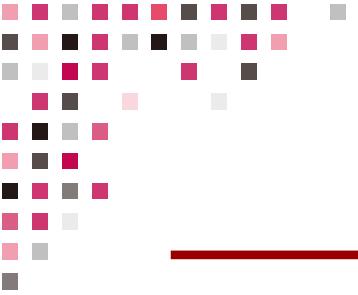


コーパスを活用した 外来語表記のゆれに関する調査

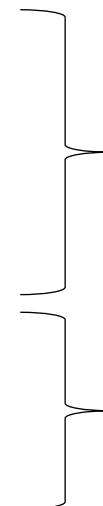
立命館大学文学部
小椋秀樹

Futurize.
きみの意志が、未来。



目次

1. はじめに
2. 外来語表記のゆれの調査(1)
 - 外来語表記のゆれの概要
 - 語末長音の表記
3. 外来語表記のゆれの調査(2)
 - [ei]の表記の経年変化
4. 外来語表記のゆれの背景
5. 国語施策におけるコーパスの活用



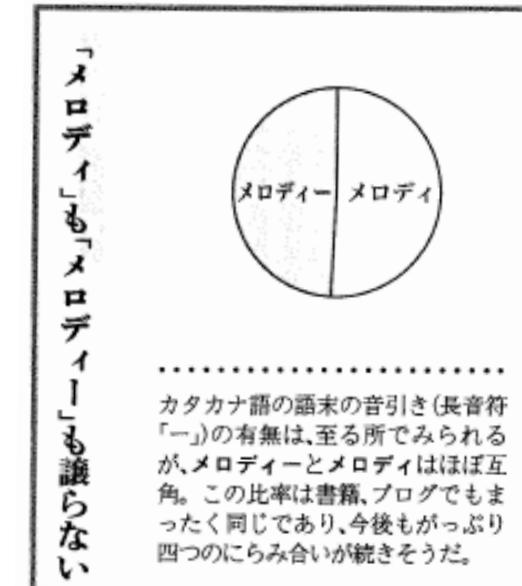
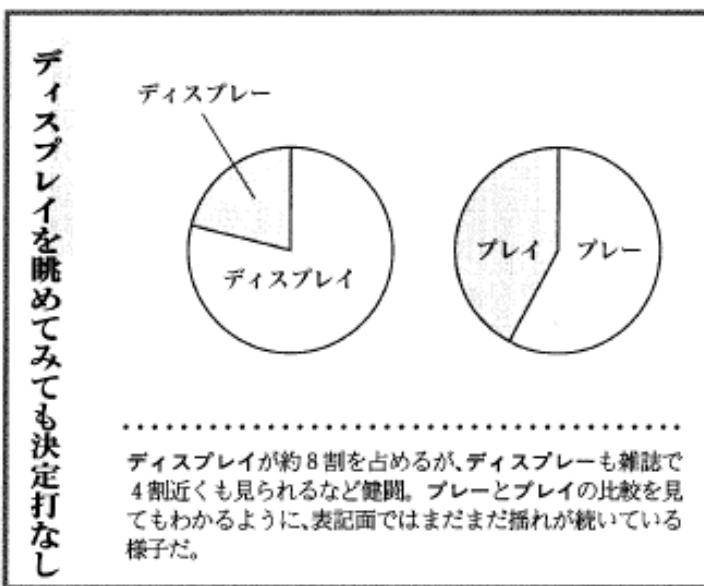
BCCWJ

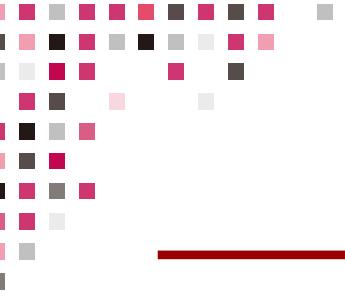
CHJ, SHC

1. はじめに

● 外来語表記のゆれ

林巨樹・松井栄一監修『現代国語例解辞典 第5版』（小学館）

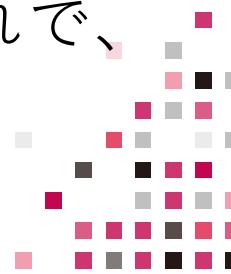


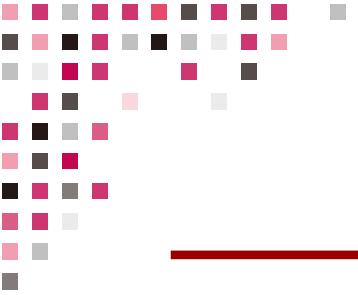


1. はじめに

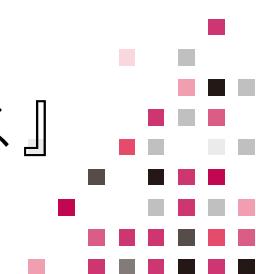
- 外来語表記のゆれ — 表記のゆれか発音のゆれか
- 宮島・高木 (1984:44)

「ピアノ～ピヤノ」は、発音がちがうはずだが、その差は微妙である。 「ア」または「ヤ」がかいてあるからといって、そのとおり発音しているみるのは、あぶない。つまり、これは発音の差のはずだが、じつは表記の差にすぎないかもしれない。 （「タイプライタ」とかいている人も、「～ター」とのばして発音している可能性がないことはない。）それで、発音のゆれとおもわれるものも、純粋に表記上のゆれと区別せずに、とりあげることにした。





2. 外来語表記のゆれの調査(1)

- 資料：『現代日本語書き言葉均衡コーパス』（BCCWJ）
 - (1) 外来語表記のゆれの概要：コアデータ（人手修正済み、精度99%以上）
雑誌、書籍、新聞、白書、Yahoo!知恵袋、
Yahoo!ブログ
 - (2) 語末長音：雑誌、書籍、新聞、Yahoo!知恵袋、Yahoo!ブログ
 - (3) 語末長音 — 表記のゆれと発音のゆれ：書籍
『日本語話し言葉コーパス』
- 

2. 外来語表記のゆれの調査(1)

(1) 表記のゆれの概要

表1：外来語表記のゆれ（異なり語数）

| | 第1表 | 第1表- 第2表 | 語末 長音 | 語中 長音 | 長音- 連母音 | イア- イヤ | クス- キス | 促音 | 撥音 | 清濁 | その他 |
|-----|-----|-------------|----------|----------|------------|-----------|-----------|----|----|----|-----|
| 雑誌 | 7 | 10 | 10 | 3 | 14 | 0 | 0 | 0 | 1 | 2 | 6 |
| 書籍 | 2 | 4 | 3 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 |
| 新聞 | 1 | 1 | 5 | 2 | 5 | 1 | 0 | 1 | 1 | 1 | 3 |
| 白書 | 1 | 0 | 4 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| Web | 7 | 7 | 18 | 3 | 5 | 2 | 1 | 4 | 0 | 0 | 11 |

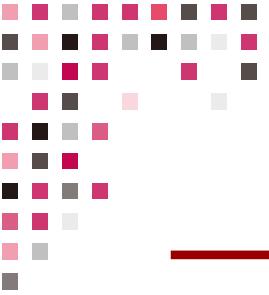
第1表：アイデア-アイディア

第1表-第2表：バイオリン-ヴァイオリン

語末長音：カテゴリー-カテゴリ

語中長音：マネージャ-マネジャー

長音-連母音：メール-メイル



2. 外来語表記のゆれの調査(1)

(2) 語末長音の表記のゆれ

表2：語末長音の表記（異なり語数、度数2以上）

| | 異なり | ゆれ | 割合 | ゆれなし | | |
|-----|------|-----|-------|------|-----|----|
| | | | | 計 | 符号 | 省略 |
| 雑誌 | 720 | 136 | 18.9% | 584 | 546 | 38 |
| 書籍 | 1042 | 259 | 24.9% | 783 | 695 | 88 |
| 新聞 | 255 | 14 | 5.5% | 241 | 238 | 3 |
| Web | 934 | 254 | 27.2% | 680 | 647 | 33 |

- ・長音符号による表記が優勢
- ・ゆれには、レジスターによる差異がある。

Web、書籍 > 雜誌 > 新聞

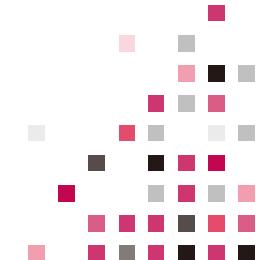
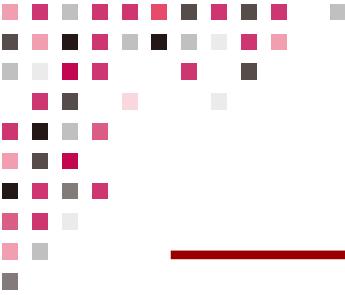


表3：長音符号を付ける表記のみ

| 雑誌 | | 書籍 | | 新聞 | | Web | |
|-------|-----|------|-----|-------|-----|------|------|
| センター | 699 | リーダー | 832 | センター | 430 | サッカー | 1169 |
| ユーザー | 417 | サッカー | 670 | サッカー | 224 | レーザー | 883 |
| サッカー | 306 | マスター | 506 | メンバー | 202 | ナンバー | 795 |
| メンバー | 288 | オーナー | 489 | エネルギー | 129 | アンサー | 516 |
| エネルギー | 267 | ナンバー | 387 | メーカー | 126 | オーナー | 505 |

表4：長音符号を省略する表記のみ

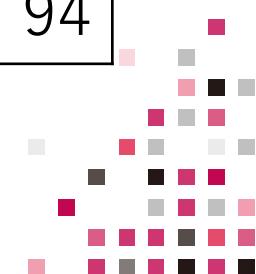
| 雑誌 | | 書籍 | | Web | |
|---------|----|---------|-----|---------|----|
| セレブリティ | 14 | モニタ | 132 | モニタ | 87 |
| ユーザビリティ | 12 | セクシャリティ | 94 | リーダ | 22 |
| スタビリティ | 8 | リーダ | 64 | バリスタ | 13 |
| クラリティ | 7 | バイナリ | 44 | コンパイラ | 8 |
| バイナリ | 7 | カウンティ | 41 | セクシャリティ | 8 |



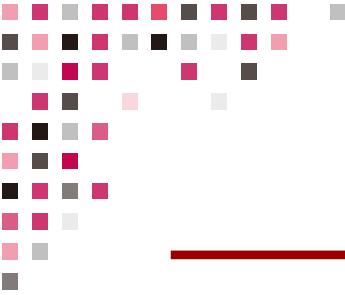
2. 外来語表記のゆれの調査(1)

表5：語末長音の表記にゆれの見られる語

| 雑誌 | | | 書籍 | | | Web | | |
|--------|-----|-----|--------|-----|-----|--------|-----|-----|
| | 符号 | 省略 | | 符号 | 省略 | | 符号 | 省略 |
| メモリー | 139 | 140 | パーティー | 362 | 469 | カテゴリー | 248 | 251 |
| パーティー | 121 | 156 | フィルター | 294 | 212 | プリンター | 209 | 240 |
| サーバー | 106 | 87 | メロディー | 110 | 118 | メロディー | 113 | 126 |
| バラエティー | 51 | 60 | ミステリー | 80 | 65 | バラエティー | 126 | 111 |
| フォルダー | 46 | 39 | コンテンサー | 50 | 53 | アダプター | 86 | 94 |



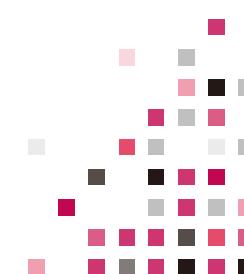
※ 長音符号を付ける表記の割合：40～60%

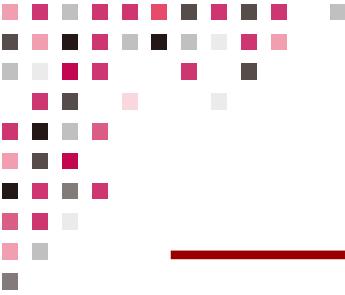


2. 外来語表記のゆれの調査(1)

表6：語末長音の表記のゆれ（語末別、度数2以上）

| | 異なり | ゆれ | 割合 | ゆれなし | | |
|-----|-----|-----|-------|------|-----|----|
| | | | | 計 | 符号 | 省略 |
| -ar | 45 | 7 | 15.6% | 38 | 38 | 0 |
| -er | 800 | 174 | 21.8% | 626 | 607 | 19 |
| -or | 151 | 57 | 37.7% | 94 | 82 | 12 |
| -gy | 42 | 6 | 14.3% | 36 | 36 | 0 |
| -dy | 26 | 21 | 80.8% | 5 | 5 | 1 |
| -ry | 115 | 36 | 31.3% | 79 | 79 | 1 |
| -ty | 122 | 76 | 62.3% | 46 | 12 | 34 |





2. 外来語表記のゆれの調査(1)

(3) 語末長音 — 表記のゆれと発音のゆれ

- 石野 (1991:47-48)

〔レモンティー〕は、だれもがこのように発音していると思われるが、これを〔レモンティ〕と書く人がいる。同様に〔シティー〕と発音している人がほとんどだと思われるのに、広告などでは「シティ」を目にすることがはなはが多い。昔、暗殺された「ケネディー」大統領の名も、大抵の人が語尾を伸ばして発音していたはずだが、マスコミは統一的に「ケネディ」と書いていた。なぜか世間には、「ティ」や「ディ」の音価を〔ティー〕〔ディー〕と誤解している人が多いようである。

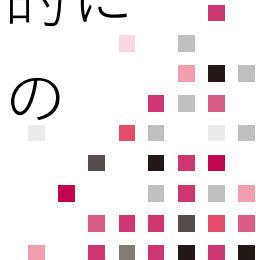


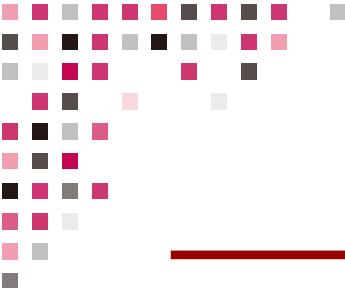
表7：語末長音の符号省略（BCCWJ・書籍、語末母音別、延べ）

| | 全体 | ア段 | イ段 | ウ段 | エ段 | オ段 |
|----|-------|-------|-------|------|------|------|
| 省略 | 16597 | 9145 | 7373 | 53 | 23 | 3 |
| | 17.0% | 16.1% | 25.1% | 0.8% | 1.2% | 0.1% |
| 語数 | 97583 | 56654 | 29341 | 6467 | 1976 | 3145 |

表8：語末長音の短呼（CSJ、語末母音別、延べ）

| | 全体 | ア段 | イ段 | ウ段 | エ段 | オ段 |
|----|-------|-------|------|------|------|------|
| 短呼 | 1630 | 1204 | 401 | 11 | 7 | 7 |
| | 7.7% | 9.4% | 6.9% | 0.7% | 1.1% | 1.5% |
| 語数 | 21208 | 12760 | 5806 | 1544 | 617 | 481 |

- 表記と発音とで、ずれが見られる。
- 語末がア段音・イ段音の語で、ずれが大きい。
→ 様々な語で生じているわけではない。



2. 外来語表記のゆれの調査(1)

表9：短呼率10%以上（語末ア段音、度数50以上）

| 語 | 度数 | 短呼 | 短呼率 |
|---------|------|-----|-------|
| ドアー | 160 | 141 | 88.1% |
| トリガー | 64 | 53 | 82.8% |
| クラスター | 616 | 206 | 33.4% |
| センサー | 73 | 13 | 17.8% |
| パラメーター | 1071 | 187 | 17.5% |
| サーバー | 69 | 10 | 14.5% |
| オペレーター | 77 | 9 | 11.7% |
| コンピューター | 540 | 56 | 10.4% |
| オーダー | 50 | 5 | 10.0% |

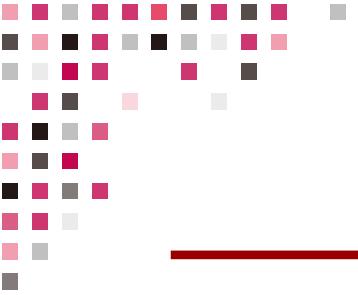
※1 異なり578語

度数50以上 50語、ゆれあり41語

※2 「ドアー」

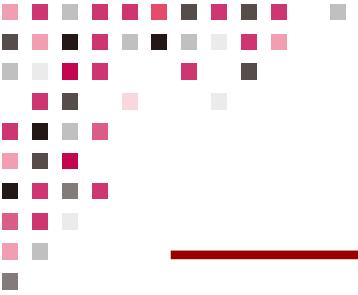
語彙素（見出し）の表記

短呼形「ドア」を語彙素とすると、
調査対象外。



3. 外来語表記のゆれの調査(2)

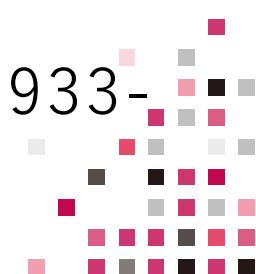
- 資料：『日本語歴史コーパス 明治・大正編Ⅰ雑誌』（CHJ）
 - 『国民之友』（1887・1888）『太陽』（1895-1925）
※『太陽』：収録期間 1901-1925年は8年間隔で4か年分
- 『昭和・平成書き言葉コーパス』（SHC）
 - 雑誌：『中央公論』（1933-1957）『文藝春秋』（1965-2013）
※SHC：雑誌、ベストセラー書籍、新聞
収録期間 1933-2013年、8年間隔で11か年分
- ◆ 発音との比較『日本語日常会話コーパス』
- [ei]の表記の経年変化 【例】メール-メイル、プレーヤー-プレイヤー

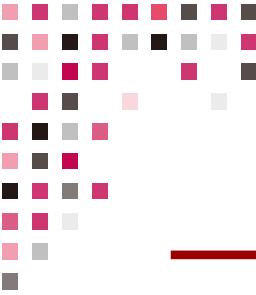


3. 外来語表記のゆれの調査(2)

表10：[ei]の表記の変化（延べ）

| | 長音 | | 連母音 | |
|------|-------|-------|------|-------|
| 明治 | 343 | 93.7% | 23 | 6.3% |
| 大正 | 318 | 90.1% | 35 | 9.9% |
| 昭和前期 | 1518 | 95.7% | 69 | 4.3% |
| 昭和後期 | 5357 | 86.7% | 822 | 13.3% |
| 平成 | 10244 | 84.3% | 1904 | 15.7% |

- 5期に区分：明治（1888-1909）大正（1917-1925）、昭和前期（1933-1941）、昭和後期（1949-1981）、平成（1989-2013）
- 



3. 外来語表記のゆれの調査(2)

- 長音表記が主流。平成期でも連母音は約15%。
→ 高頻度語の表記が影響

表11：各期の上位3語

【明治】

| | 度数 | 長音 |
|-----------|----|----|
| エーカー | 39 | 39 |
| メール | 36 | 36 |
| インスピレーション | 35 | 35 |

【大正】

| | 度数 | 長音 |
|------|----|----|
| メーテー | 42 | 42 |
| テーブル | 39 | 37 |
| プレート | 38 | 23 |

【昭和前期】

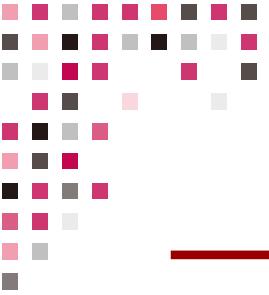
| | 度数 | 長音 |
|----------|-----|-----|
| インフレーション | 592 | 578 |
| メーテー | 148 | 148 |
| テーブル | 78 | 78 |

【昭和後期】

| | 度数 | 長音 |
|------|-----|-----|
| ケース | 385 | 385 |
| データ | 350 | 350 |
| テーブル | 266 | 241 |

【平成】

| | 度数 | 長音 |
|------|-----|-----|
| ケース | 801 | 801 |
| データ | 617 | 617 |
| メーカー | 610 | 610 |



3. 外来語表記のゆれの調査(2)

表12：[ei]の表記の変化（異なり）

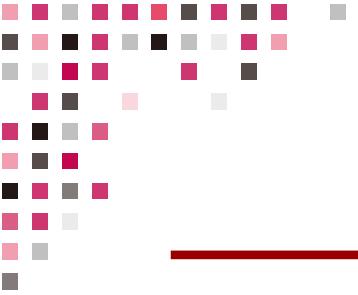
| | 長音 | | 連母音 | |
|------|-----|-------|-----|-------|
| 明治 | 70 | 90.9% | 7 | 9.1% |
| 大正 | 72 | 85.7% | 12 | 14.3% |
| 昭和前期 | 118 | 82.5% | 25 | 17.5% |
| 昭和後期 | 330 | 70.2% | 140 | 29.8% |
| 平成 | 391 | 66.6% | 196 | 33.4% |

- 集計方法

【例】メーン(5)、メイン(1)、ステーション(10)、テーブル(2)

→ 長音表記：異なり3語

連母音表記：異なり1語



3. 外来語表記のゆれの調査(2)

表13：III期の表記（異なり、度数2以上）

| | 語数 | % |
|-----|-----|-------|
| 長音 | 263 | 68.1% |
| 連母音 | 87 | 22.5% |
| ゆれ | 36 | 9.3% |

- 3期に区分：I期（明治・大正・昭和前期）、II期（昭和後期）、III期（平成）

- 集計方法：

「長音」 長音表記率70%以上

「連母音」 長音表記率30%以下

「ゆれ」 上記以外

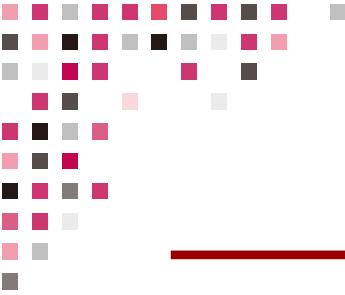
【例】《メール》20例

メール 15例

メイル 5例

長音表記率75%

→ 「長音」

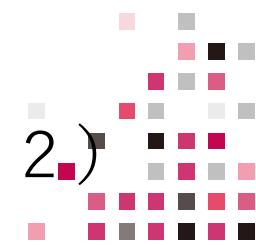


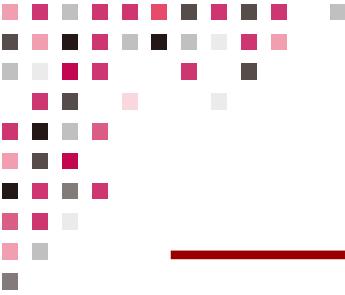
3. 外来語表記のゆれの調査(2)

- III期に出現した度数2以上の386語の初出時期

I期：95語 II期：116語 III期：109語

※ 初出とは今回の調査範囲内。初出時期の度数2以上の語

- ・ 勾配の坂路を往く際には、ブレーキをも掌つて、注意に注意を加へねばならぬといふ大變な役目（『太陽』、I期、1901年）
 - ・ 南洋やタイでは起らないというのは、やはり人間のクリエイティブな意思の有無によるわけでしょう。（『中央公論』、II期、1957年）
 - ・ 新しい作家登場という印象があった人で、『暖流』のリメイク（昭32）なんかとても印象的でした。（『文藝春秋』、III期、1989年）
- 



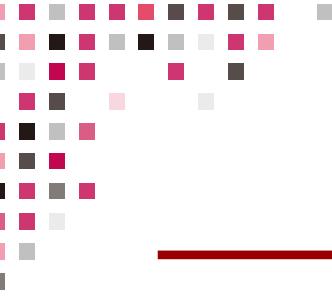
3. 外来語表記のゆれの調査(2)

- 初出時期における表記とIII期における表記

表14：初出時期の表記とIII期の表記（異なり、度数2以上）

| I期 | | | III期 | |
|-----|----|-------|------|----|
| 分類 | 語数 | % | 分類 | 語数 |
| 長音 | 87 | 91.6% | 長音 | 76 |
| | | | 連母音 | 8 |
| | | | ゆれ | 3 |
| 連母音 | 3 | 3.2% | 長音 | 1 |
| | | | 連母音 | 2 |
| | | | ゆれ | 0 |
| ゆれ | 5 | 5.3% | 長音 | 2 |
| | | | 連母音 | 2 |
| | | | ゆれ | 1 |

| II期 | | | III期 | |
|-----|----|-------|------|----|
| 分類 | 語数 | % | 分類 | 語数 |
| 長音 | 73 | 62.9% | 長音 | 64 |
| | | | 連母音 | 1 |
| | | | ゆれ | 8 |
| 連母音 | 29 | 25.0% | 長音 | 2 |
| | | | 連母音 | 24 |
| | | | ゆれ | 3 |
| ゆれ | 14 | 12.1% | 長音 | 4 |
| | | | 連母音 | 7 |
| | | | ゆれ | 3 |

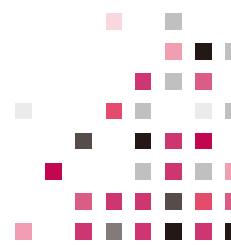


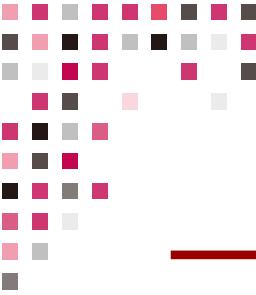
3. 外来語表記のゆれの調査(2)

- 萩野 (2014:100-103)

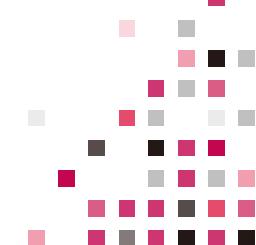
資料 : Web

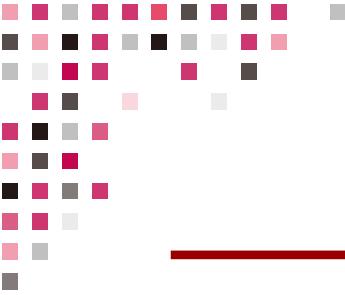
対象 : 20語 (「テークアウト」「ブレーキ」等)

- 長音表記が圧倒的に多い。
 - 長音表記か連母音表記かは語ごとに決まっており、同程度で表記がゆれている語は見られない。
 - 古い時代に日本語に入ってきた語は長音表記される傾向。
- 

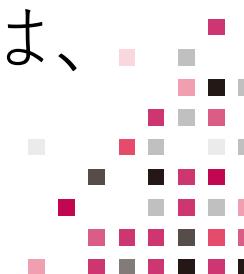


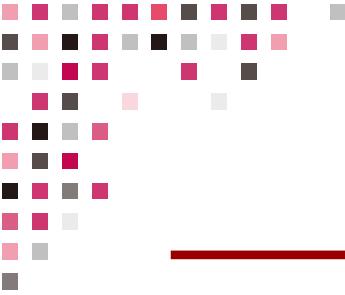
3. 外来語表記のゆれの調査(2)

- 変化の要因 — 発音との関係 —
 - 宮島・高木 (1984:55)
「原音が2重母音という意識があるとき」に連母音表記
 - 佐竹 (1986:416)
連母音表記は、「長音でないという意識が強いことの証明で」、
「そのような意識が強いというならば、長音符号と母音表記との
対立は、もはや長音表記のしかたのゆれではなく、発音の
ゆれの問題となる。」
- 



3. 外来語表記のゆれの調査(2)

- 変化の要因 — 発音との関係 —
 - 岡田 (2004) 資料：『日本語話し言葉コーパス』
対象：原音で[ei]がある英略語、外来語
 - 外来語では、長母音[e:]で実現される傾向
 外来語 1679例 [ei] 125例 (6.9%)
 - /ei/+/N/という音節構造（「メイン」「レーン」等）では、
連母音[ei]で実現される傾向
 外来語 56例 [ei] 54例 (96.4%)
- 



3. 外来語表記のゆれの調査(2)

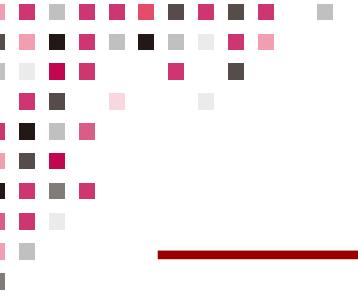
- 変化の要因 — 発音との関係 —
- ・『日本語日常会話コーパス』

表15：[ei]の発音（延べ）

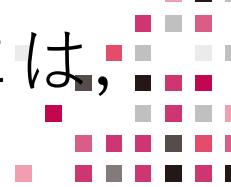
| | 用例数 | % |
|-----|------|-------|
| 長音 | 2994 | 93.9% |
| 連母音 | 195 | 6.1% |

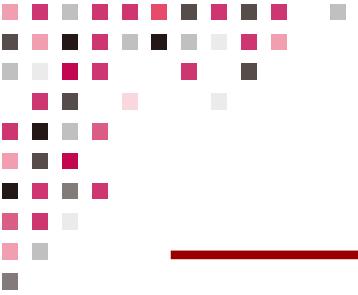
表16：[ei]の発音（異なり）

| | 用例数 | % |
|-----|-----|-------|
| 長音 | 278 | 95.9% |
| 連母音 | 12 | 4.1% |



4. 外来語表記のゆれの背景

- 『外来語の表記』（1991年、内閣告示第2号、内閣訓令第1号）
 - 留意事項その1（原則的な事項）
 - 3. 「ハンカチ」と「ハンケチ」、「グローブ」と「グラブ」のように、語形にゆれのあるものについて、その語形をどちらかに決めようとはしていない。
 - 4. 語形やその書き表し方については、慣用が定まっているものはそれによる。分野によって異なる慣用が定まっている場合には、それぞれの慣用によって差し支えない。
- 



4. 外来語表記のゆれの背景

- 留意事項その2（細則的な事項）

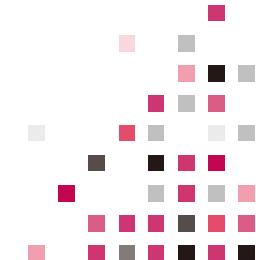
III 摭音，促音，長音その他に関するもの

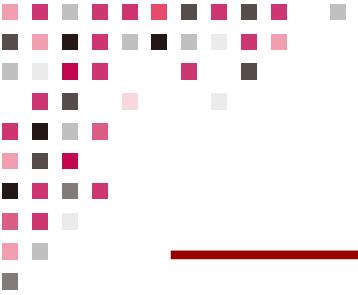
3 長音は，原則として長音符号「ー」を用いて書く。

〔例〕 エネルギー オーバーコート グループ ゲーム ショー¹
テーブル パーティー（以下略）

注1 長音符号の代わりに母音字を添えて書く慣用もある。

〔例〕 バレエ（舞踊） ミイラ





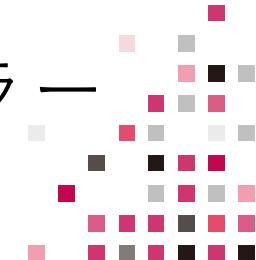
4. 外来語表記のゆれの背景

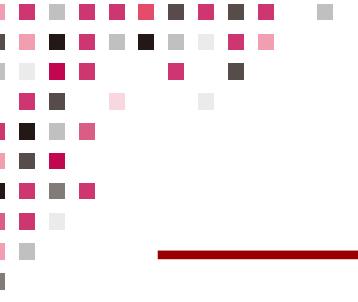
注2 「エー」「オー」と書かず、「エイ」「オウ」と書くような慣用のある場合は、それによる。

〔例〕 エイト ペイント レイアウト (以下略)

注3 英語の語末の-er, -or, -arなどに当たるものは、原則としてア列の長音とし長音符号「ー」を用いて書き表す。ただし、慣用に応じて「ー」を省くことができる。

〔例〕 エレベーター ギター コンピューター マフラー
エレベータ コンピュータ スリッパ





4. 外来語表記のゆれの背景

- JIS Z 8301 : 2008 規格票の様式及び作成方法

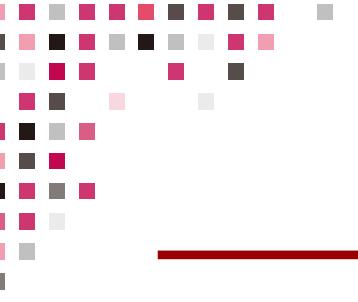
- 外来語語末長音の表記
 - a) 専門分野の用語の表記による。
 - b) 規格の用語及び学術用語にない用語の語尾に付ける長音符号は、表G・3による。

→ 長音符号を除いた部分の拍数 3拍以上…付けない（エレベータ）

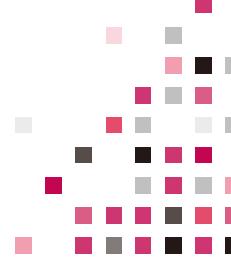
2拍以下…付ける（カバー）

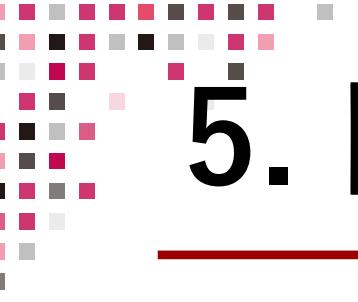
- JIS Z 8301 : 2019 規格票の様式及び作成方法

- 外来語の表記：主として『外来語の表記』による。



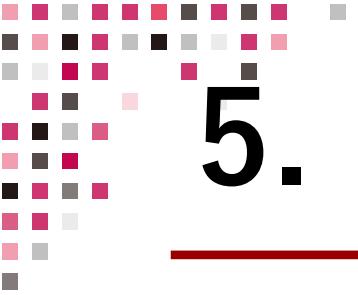
4. 外来語表記のゆれの背景

- NHK 放送用語の基準
 - 山下 (2015, 2016)
give-and-take, take out, nature, age, taste など
→ 連母音表記を採用 例：ギブアンドテーク→ギブアンドティク
 - 『読売新聞』 用字用語の基準
 - 読売新聞社 (2014)
main event メーンイベント→メインイベント
- 



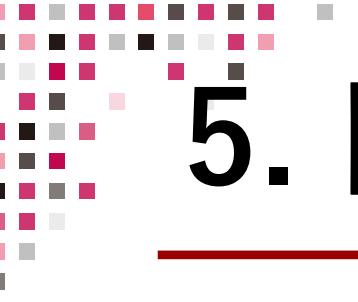
5. 国語施策におけるコーパスの活用

- 「国語施策は、空気のようなものでなければならない」
(氏原基余司 元主任国語調査官)
→ 日本語の今の姿を捉えることが必要。
- 実態調査
 - 「改定常用漢字表」 文化庁「漢字出現頻度数調査」(書籍、新聞、Web)
同 「出現文字列頻度数調査」
国立国語研究所 BCCWJ (出版・書籍、1730万語)
- 意識調査：国語に関する世論調査
 - 「改定常用漢字表」 追加候補漢字に対する印象
 - 「改定ローマ字のつづり方」 ローマ字表記に関する意識



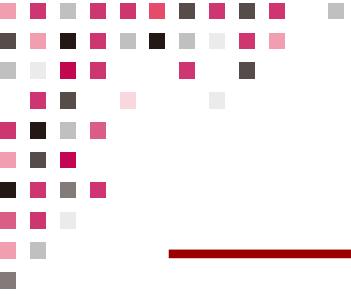
5. 国語施策におけるコーパスの活用

- 日本語の今の姿を捉える。
- 現代のような変化の激しい時代にあっては、「言葉に関する施策」についても、定期的な見直しが必要である。（中略）定期的・計画的な漢字使用の実態調査をしていくことが重要である。 （「改定常用漢字表」p.(16))
- BCCWJ：主要なデータは、2005年まで
→ 「日本語の今の姿」を捉えるためには、拡張が必要
- BCCWJ2：書籍（2006-2025）、教科書、SNS/CMC
→ 「日本語の今の姿」 + どのような変化を経て今の姿があるか。
- 繙続的な拡張による言語の使用実態の把握が必要



5. 国語施策におけるコーパスの活用

- 単語情報の活用
 - 表記に関する基準：語という観点からの実態調査が必要
 - 国立国語研究所のコーパス：形態論情報、特に短単位情報の活用
 - 漢字：どのような音訓で、どのような語（語の構成要素）の表記に、どのくらい使われているのか。
 - 外来語：音の表記と合わせて、どのような語に、どのような表記が、どのくらい用いられているのか。



参考文献

- 岡田祥平 (2004) 「『日本語話し言葉コーパス』に観察される母音連續/ei/のバリエーション－外来語の場合－」、『電子情報通信学会技術研究報告〔音声〕』104-148、pp.35-40
- 荻野綱男 (2014) 『ウェブ検索による日本語研究』、朝倉書店
- 小椋秀樹 (2013) 「現代日本語における外来語表記のゆれ」、相澤正夫編『現代日本語の動態研究』、おうふう、pp.151-171
- 小椋秀樹 (2014) 「外来語語末長音の表記のゆれについて」、『論究日本文学』102、pp.1-12
- 小椋秀樹 (2024) 「明治期から平成期にかけての外来語における[ei]の表記」、「通時コーパス」シンポジウム
2024
- 佐竹秀雄 (1986) 「外来語表記法の問題点」、宮地裕編『論集 日本語研究(1) 現代編』、明治書院、pp.407-422
- 宮島達夫・高木翠 (1984) 「雑誌九十種資料の外来語表記」、『国立国語研究所報告79 研究報告集』5、pp.43-76
- 山下洋子 (2015) 「外来語の発音・表記について～用語の決定～」、『放送研究と調査』65-5、pp.86-93
- 山下洋子 (2016) 「用語の決定」、『放送研究と調査』66-4、pp.86-95
- 読売新聞社 (2014) 『読売新聞用字用語の手引き 第4版』、中央公論新社

