

# 時を表す副詞類の前置詞省略 —Brown Family Corpora 比較分析—

## Zero-Prepositions in Time Adverbials in the Brown Family Corpora

### 序章 はじめに

本論文は、いくつかの目的を持って執筆されている。まず、言語学 (Linguistics) という研究分野においてコーパス言語学はどのような研究領域に属するのか、コーパス言語学の立ち位置、つまりコーパス言語学とは何を研究しようとする学問領域なのかを他の分野との関係において説くことを目的とする。比較的新しい研究分野であると言われる所以はやはりコンピュータの進化と大いに関係している。コンピュータ時代以前から、辞書編纂、文法書、英語教授法等で多くの学者・研究者が手で集めたデータを基にしていたのは周知の事であったが、言語学研究分野からの見地でコーパス (Corpus) という専門用語が使われ始めたのは、コンピュータの急激な進化と大いに関係が深いと思われる。コンピュータの急速な発展に伴いコーパス言語学 (Corpus Linguistics) という研究分野が確立され、そしてその興隆と学会活動が世界的、そして日本国内で急速に発展していった事を経験するのである。興味ある現象は、このコーパス言語学の始まりと活動の中心は、コンピュータの発展著しいアメリカがその中心であると思われがちであるが、イギリス、ドイツ、オランダ、北欧等のヨーロッパでがその中心であったことであろう。

この論文で特筆したいのは、世界的なコーパス言語学の興隆に同調する形で日本でもコーパス言語学が急速に発展してきて、学会活動、関係論文の執筆が急激に増えていった事である。そしてその発展に大東文化大学が深く関りをもっていて、人的交流を含めて大東文化大学が大きな役割と功績を担ってきたという点である。つまり、世界的なコーパス言語学発展の過程で、大東文化大学が多くの国際的に活躍している高名なコーパス言語学者 (Corpus Linguists) を招聘し、当時話題性の高いテーマについて直接講演、またはソフトを使った講習会を実施する事で広く日本のコーパス言語学界の発展、そしてコーパス学会会員、研究者、さらにはコーパス言語学を専攻した学生等に大きな影響を与えた事が関係学会で叫ばれた。それらの講演、講習会の様々なエピソード、さらに個人的な交流で知りえた情報等を可能な限り交えてここに述べている。

勿論、本論文はコーパス言語学研究報告であり、コーパスを利用した検索、解析による言語学的検証が主な目的である。タイトルにあるように言語学的研究において、「時を表す副詞類における前置詞省略」を Brown Family Corpora<sup>1</sup> を使って検索し、その結果を報告するの

---

<sup>1</sup> Brown Corpus (Standard Corpus of Present-Day Edited American English)のテキストカテゴリーを模して構成された 100 万語コーパスの総称。

が本論文の主な目的である。換言すれば、「時を表す副詞類における前置詞省略」とは、「時を表す副詞類」では、アメリカ英語、イギリス英語等どここの英語が、新聞記事、学術論文、小説等どのような英文で、それら英文の文頭、文末等どの位置で、どのような前置詞を省略する傾向にあるのかをコーパスを使って検索されたデータを統計的に処理して検証する事が主な目的である。

## 第1章 コーパス言語学とは何か

### 1.1 コーパス言語学とは何か

言語学 (Linguistics) は、言語を科学的に研究する学問分野であり、それに関する専門的な研究書がすでに多数出版されているが、特にその中で注目したい研究書二編について述べたい。David Crystal (1987) は、言語学について「言葉の意味、理想的な話し方、言語の起源などが考察され、観察は主観的であった。... 言語に対して関心を寄せる専門家の数はますます増加し、言語分析を中心とする科学的研究の新しい分野の登場」とその発展について述べており、それらの研究分野は文献学 (philology) と呼ばれていたが言語学 (linguistics) あるいは言語科学 (linguistic science) と呼ばれるようになったと説明している (風間喜代三/長谷川欣佑監訳 1992)。また城生伯太郎 (1990: 37) は「言語学は科学である」という説を述べる時に、言語学が「科学」であること条件として言語学の研究分野は、1) 音声、音韻、2) 文法、3) 意味の三つに分かれると説いている。それに言語学研究分野を付け加えて概説すると 1) 音声学 (この分野を前段言語学と位置付ける説もある)、音韻論、そして「語彙」を扱う辞書、形態論が加わる。2) 文、句などを扱う文法は統語論であり、3) 意味を扱うのは意味論である。さらに 4) それら言語学の応用として応用言語学が加わり、語用論、文体論、社会言語学、談話分析、言語教育、自然言語処理、コーパス言語学等が含まれる。その言語学の中で、英語という特定言語を研究する学問は英語学 (English Linguistics) である。

上記のようにコーパス言語学 (Corpus Linguistics) は応用言語学の枠組みに入る研究分野であるが、その一般的な定義は「言語分析のために集積された言語資料 (コーパス) をコンピュータを使って解析」する比較的新しい研究分野であると言われる。コーパス言語学は、理論と実践の言語調査・分析、人文学・自然科学的なアプローチ、質的・量的分析など異分野との融合の中で発展する学問であると言われる。

McEnery and Hardie (2012) はコーパス言語学について次のように解説している:

Corpus linguistics is not a monolithic, consensually agreed set of methods and procedures for the exploration of language... it is very important to realise that corpus linguistics is a heterogeneous field. (McEnery and Hardie, 2012: 1).

つまり、「言語学にはさまざまな分野がありますが、言語データの精緻な観察を基盤とし、他の研究分野との自由な連携を志向するコーパス言語学は、言語学の「混成的な場」 (heterogeneous field) であり、言語学を学ぶ入口として最適なものです」と説いている (石川慎一郎編、2012: 4)。

Biber et al. (2015) は、コーパス言語学は言語変異と用法についての経験的な調査を可能にする研究アプローチであり、普遍化と妥当性を可能にする研究結果を導き出していると説き、さまざまな研究目的と特徴を持っていると述べている。

Corpus linguistics is a research approach that facilitates empirical investigations of language

variation and use, resulting in research findings that have much greater generalizability and validity than would otherwise be feasible. (Biber et al., 2015: 1)

コーパス (corpus) という用語が言語学的意味で OED (Oxford English Dictionary) に初めて登場するのは 1956 年であると言われている。

The OED's first citation of the *corpus* in the linguistic literature is dated at 1956, in an article by W. S. Allen in the *Transactions of the Philological Society*, where it is used in the more familiar meaning of 'the body of written or spoken material upon which a linguistic analysis is based (OED: second edition, 2009). (O'Keeffe, A. and M. McCarthy (eds.) 2010: 5).

コーパス (corpus) の利用について、Crystal (1995) はその有用性を次のように述べている。

A well-constructed general corpus turns out to be useful in several ways. It enables investigators to make more objective and confident descriptions of usage than would be possible through introspection. It allows them to make statements about frequency of usage in the language as a whole, as well as comparative statements about usage in different varieties. (Crystal, 1995: 438).

しかし、コーパス言語学という用語は、Aarts and Meijs (eds.) *Corpus Linguistics* (1984) 以降一般的に広く使われはじめ、Leech (1992) によれば 1990 年前後に定着した用語であると言われている。

The term "corpus linguistics" made only occasional appearances until the publication of a book of that title edited by Aarts and Meijs (1984). (I am not aware that it made its appearance in the corpus linguistics heyday of the 1950s – simply because, for those who espoused this approach, corpus linguistics was simply "linguistics" – to them, no other linguistics deserved the name.) (Leech, 1992: 105).

換言すれば、コーパス (Corpus) とは、元来「身体」を意味するラテン語 (corpse) から由来することばであるが、所謂言語資料の集積という意味で使われたのは比較的新しいことが分かる。それら集積したコーパスのうち英語の電子テキストを機械可読処理し英語の分析・記述を行うのが英語コーパス言語学 (English Corpus Linguistics) である。その歴史を概観すると生成文法の台頭と重なり 1960 年代には閉塞状態であったことが、次のような Chomsky の 'sarcastic observation'<sup>2</sup> から推測される。

*I live in New York is more frequent than I live in Dayton Ohio.* (Halliday, 1991: 30, in K. Aijmer and B. Altenberg (eds.) *English Corpus Linguistics*).

しかし、コンピュータテクノロジーの目覚ましい発展とともに 1980 年代にはコーパス言語学が爆発的な発展を遂げ、英語研究論文の急激な発表がなされたことを、Johansson (1991: 312) は次の統計を出して明瞭に説明している。

---

<sup>2</sup> Chomsky によって、コーパス言語学に対して揶揄する表現が、Linguistic Society of America Summer Institute, July 1964 でなされた。

A study of the bibliography in Altenberg (1991), which focuses on publications based on or related to the English text corpora distributed through ICAME, reveals the following pattern:

Year	Number of publications
—1965	10
1966—1970	20
1971—1975	30
1976—1980	80
1981—1985	160
1986—1990	320

また、スウェーデンのコーパス言語学者 J. Svartvik は Nobel 財団 (Nobel Foundation) から財政援助を得て、1991 年にストックホルムで Nobel Symposium を開催し、世界的に活躍するコーパス言語学者を多く招待した。その Nobel Symposium には、国際的に高名な学者が招待され、その発表原稿はのちに著書にまとめられている (Svartvik, J. ed. *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82 Stockholm, 4-8 August 1991*. Mouton de Gruyter. 1992)。その中で、主催者の J. Svartvik は ‘Corpus linguistics comes of age’ (Svartvik 1992: 7) という表現を用いて、コーパス言語学の新しい発展を予見している<sup>3</sup>。

Towards the end of the 1980s some of us felt that corpus linguistics had come of age and should satisfy the criteria for Nobel Symposia: being a field of great scientific importance and great relevance to society. Hence it seemed fitting to celebrate the occasion by organizing this first Nobel Symposium on Corpus Linguistics and inviting a group of prominent colleagues both at home and abroad to take stock of past achievement, consider the present state-of-the-art and speculate about future development. (Svartvik 1992: 12).

## 1.2 コーパス言語学の特徴

1980 年代に急速に発展を遂げたコーパス言語学の特徴について、Leech (1992: 107) は次のように述べている (齋藤俊雄他、2005: 4)。

- 1) Focus on linguistic performance, rather than competence、  
言語能力 (linguistic competence) よりも言語運用 (linguistic performance) に中心をおく、
- 2) Focus on linguistic description, rather than linguistic universals、  
言語の普遍的特徴 (linguistic universals) の解明よりも、個別言語の言語記述 (linguistic description) に中心をおく、
- 3) Focus on quantitative, as well as qualitative models of language、  
質的な (qualitative) 言語モデルのみならず、数量的な (quantitative) 言語モデルにも中心をおく、

---

<sup>3</sup> この 1991 年に開催された Nobel Symposium を主催した J. Svartvik は、後に述べるように大東文化大学で開催した講演会に出席し、Nobel Symposium の後日談として「この大きなコーパス言語学会」が後のコーパス言語学の発展の礎になったという興味ある話をしていた。

4) Focus on a more empiricist, rather than rationalist view of scientific inquiry、

言語研究における合理主義的 (rationalistic) な立場より、より一層経験主義的 (empirical) な立場に中心をおく。

換言すると、コーパス言語学は言語運用、個別言語の記述にその中心をおき、質的・量的に経験主義的立場で言語解析をする学問であると説いている。

また、Biber et al. (1998: 4) は以下の4つを挙げてコーパス言語学の研究の目標、分析の特徴について説明している。

- they are empirical, analyzing the actual patterns of use in natural texts;
- they are based on analysis of a large and principled collection of natural texts, known as a 'corpus'; the corpus is evaluated for the extent to which it represents a target domain of language use;
- they make extensive use of computers for analysis, employing both automatic and interactive techniques;
- they depend on both quantitative and qualitative analytical techniques.

つまり、1) 自然言語のパターンを経験的に分析する、2) 大型で目的を持って集積されたコーパスの分析にその基礎をおく、3) 多角的にコンピュータを駆使する、4) 質的・量的な分析に基づく、といった説明をしている。Leech (1992) の説に近いことが伺える。

さらに、Kennedy (1998: 3) は、コーパス言語学の主要な問題と課題は次の4つに集約されると説いている。

- How can we best exploit the opportunities which arise from having texts stored in machine-retrievable form?
- What linguistic theories will best help structure corpus-based research?
- What linguistic phenomena should we look for?
- What applications can make use of the insights and improved descriptions of languages which come out of this research?

また、Kennedy (1998) は、コーパス言語学は言語学の一つの研究領域というより、言語データを正確に観察し、他の研究分野との自由な連携を志向する学問であると述べている。

It would be misleading, however, to suggest that corpus linguistics is a theory of language in competition with other theories of language such as transformational grammar, or even more that it is a new or separate branch of linguistics. (Kennedy, 1998: 7).

つまり、コーパス言語学は生成文法のように母語話者の言語直観 (linguistic intuition) に依存するよりは、言語研究を目的で集積されたコーパスを経験的に検索して得られた言語結果を基礎として言語記述を行う言語学であると言える。換言するとコーパスを検索することで言語学のどんな分野の検証にも応用できる言語学であると言える。しかし、重要な事は、コーパス言語学は単なる方法論であるという領域をはるかに超えて、さまざまな言語研究に応用可能であるという事が実証されている (Tognini-Bonelli 2001)。

In this context, we take the view that although corpus linguistics belongs to the sphere of applied

linguistics, it differs from other partner disciplines under the same umbrella in that it can be seen as a *pre-application methodology*. While a methodology can be defined as the use of a given set of rules or pieces of knowledge in a certain situation, by “pre-application” we mean that, unlike other applications that start by accepting certain facts as *given*, corpus linguistics is in a position to define its own sets of rules and pieces of knowledge *before* they are applied; this leads the linguists to make use of some new parameters to account for the data, and this entails a change in which can be referred to the *unit of currency* for linguistic description. Corpus linguistics has, therefore, a theoretical status and because of this it is in a position to contribute specifically to other applications. Among the areas which have benefited from the input of corpus linguistics are lexicography, language teaching, translation, stylistics, grammar, gender studies, forensic linguistics, computational linguistics, to quote but a few. (Tognini-Bonelli, 2001: 1).

### 1.3 コーパス言語学の発展

コーパス言語学は、コンピュータ使用に限定したものではなくコンピュータ時代以前に辞書編纂、文法書、言語教育等の分野でさまざまなコーパスが集積され利用されていた<sup>4</sup>。

#### 1.3.1 コーパスを利用した辞書編纂

コーパスを利用して辞書編纂をする手法は、コンピュータの出現以前から採用されていたことがすでに明らかになっている。

コンピュータ利用以前の辞書編纂でコーパスが利用されていた区分を大きく二つに分けて考える。最初は‘Hard Word Dictionaries’（難解語辞書）でラテン語・英語の辞書編纂で、*Promptorium Parvulorum, Sive Clericorum* (ca. 1440), *Medulla Grammaticae* (ca. 1460), *The Dictionary of Syr Thomas Eliot* (1538) などが挙げられる。さらに英英辞書として *A Table Alphabeticall* (Cawdrey 1604), そして *The English Schoole-Master* (Coote 1596) がある。次の辞書編纂の区分として ‘Universal Dictionaries’（一般辞書）があるが、まず挙げられるのが *The New World of Words* (Phillips 1706) であり、その次に *Dictionarium Britannicum* (Bailey 1730) が編纂され、*A New Universal Etymological Dictionary* (Bailey 1755) も出された。

しかし本格的な辞書編纂として最も重要なものは、Samuel Johnson の辞書 *A Dictionary of the English Language* (1755) であり、‘Johnson collected a large corpus of illustrative sentences (ca. 40,000 entries)’(DeMaria, Jr 1993)と多くのコーパスを利用して辞書編纂をしたことが述べてある。また Johnson のコーパス編纂の方法について Hawkins (1961) は次のように解説している。

Johnson ... had, for the purpose of carrying on this arduous work, and being near the printers employed in it, taken a handsome house of Gough Square, and fitted up a room in it with desks and other accommodations for amanuenses, who, to the number of five or six, he kept constantly under his eye. An interleaved copy of Bailey’s dictionary in folio he made the repository of the several articles, and these he collected by incessant reading the best authors in our language, in the practice whereof, his method was to score with a black-lead pencil the words by him selected, and give them over to his assistants to insert in their places. The books he used for his purpose were what he had in his own collection, a copious but a miserably ragged one, and all such as he could borrow (Hawkins 1961: 77).

---

<sup>4</sup> コンピュータ時代以前のコーパス利用について、Yamazaki (1996) に詳細な記述がある。ここでは、言語教育における Thorndike (1921)、Palmer (1933)、Thorndike and Lorge (1944) は扱わない。

この Samuel Johnson の最初の辞書編纂で重要な事は‘slips of paper’であり、つまり the paper corpus を利用していたのである。この‘the corpus on slips of paper’の作業について他の研究者も言及している (O’Keeffe, A. and M. McCarthy, 2010)。

the most famous example of the ‘corpus on slips of paper’ is the more than three million slips attesting word usage that the Oxford English Dictionary (OED) project had amassed by the 1880s, stored in what nowadays might serve as a garden shed. (O’Keeffe, A. and M. McCarthy, 2010: 4).

また、James Murray が英語辞書編集を担当するようになって、‘thousands of slips were accumulated’ という作業に入り、1928 年に *The Oxford English Dictionary* (OED) が完成した。Murray はその編集作業の状況について詳しく述べている。

For the purpose of this historic illustration more than five millions of extracts have been made, by two thousand volunteer Readers, from innumerable books, representing the English literature of all ages, and from numerous documentary records. (Murray 1900: 47).

その後、James Murray の孫娘 (K. M. Elisabeth Murray) が貴重な著書を出版し、James Murray の辞書編集の様子を詳細に述べている。その中で、余りにも有名な James Murray の辞書編纂の様子が写った写真 (James Murray at work in the Scriptorium<sup>5</sup>) を表紙に載せている。つまり ‘Murray’s children had to help alphabetize and sort the slips’ と述べ、さらに詳細に slips of papers を使った辞書編纂について解説している。

Dictionary slips and their sorting became a major element in the lives of the Murray family... Every afternoon the children went to the Scriptorium to collect some packets of recently arrived slips. As each child reached an age when he or she could read, they were pressed into service. (K. M. Elisabeth Murray 1977: 178).

そして、このコーパスは ‘this corpus finally grew to over eleven million slips’ (Francis, 1992) と大きく成長したことが分かる。

同時期にアメリカでも大きな辞書編纂のプロジェクトが進行し、もともなったのは大きなコーパスであったと Francis (1992:22) は述べている (*Webster’s New International Dictionary, Second Edition* 1934 was based on a corpus of 1,665,000 citations)。さらに重要な事は、この Webster の辞書編纂は OED の編纂の反省を踏まえて次のような‘systematic’な体制を採用したと書いている。

A key word here is *systematic*. Unlike the hundreds of volunteers who built the OED corpus, the readers here were full-time professional lexicographers working together to make sure that all significant sources were searched. Other collections, some from the OED, increased the corpus for the new *Webster* to over ten million slips. (Francis 1992: 22).

---

<sup>5</sup> 有名な‘James Murray at work in the Scriptorium’の表紙の写真は本論文の補遺に挙げてある。



その後、コンピュータの急速な発展によりコンピュータコーパスの時代に入っていくのである。そして集積されたコーパスをコンピュータを駆使して語彙・用例の説明をする辞書編纂方法が本格的に始まり、Sinclair を編集主幹とする *Collins COBUILD English Language Dictionary* (1987) がその走りであり、辞書編纂の画期的な方法が示されていった。この辞書の特徴は OED の ‘historical principle’ の編集方針とは違って、語彙の意味検索の際に、現代英語ではその語彙の意味がどのように使われているかを説明するためにコーパス検索による ‘頻度順’ でその意味を説明する方針も採用したユニークな辞書である。それら編集方針についてこの辞書は次のように丁寧に説明している。

... the large group of texts, called the corpus, gives us reasonable grounds for omitting many uses and word-forms that do not occur in it. It is difficult for a conventional dictionary, in the absence of evidence, to decide what to leave out, and a lot of quite misleading information is thus preserved in the tradition of lexicography.

This dictionary makes a break with such traditions. We have gone back to basics and collected many millions of words, and put them into a very large computer. The dictionary team has had daily access to about 20 million words, with many more in specialized stores. (1987: xv).

Some dictionaries put the earliest recorded meanings first, so that the present-day meanings are to be found much later in the entry. It should be said at once that this is not a historical dictionary; it is a description of what the language is like at the present time. (1987: xix).

また、その後出版された多くの辞書も大型コーパスを利用して編纂されている。*Oxford Advanced Learner's Dictionary* (1995) は ‘the 100 million-word British National Corpus, the 40 million-word Oxford American English Corpus, 40 million-word Oxford American English Corpus, the Oxford English Dictionary research programme’ を使用していることを明記し、*Longman Dictionary of Contemporary English* (1995, 3<sup>rd</sup> Edition) は ‘the 100 million word British National Corpus, the 30 million word Longman Lancaster Corpus’ を利用していると強調している。

### 1.3.2 コーパスを利用した文法書

文法書に関しても、同様にコンピュータ時代以前にもコーパス利用がなされていた。まず挙げられるのが、Zachrisson (1913) で多くの ‘original letters’ を利用したと書いているが、Jespersen (1914) は文法書、新聞、雑誌等から多くのコーパスを集めて文法書を書いた。

With regard to my quotations, which I have collected during many years of both systematic and desultory reading, I think that they will be found in many ways more satisfactory than even the best made-up examples, for instance those in Sweet's chapters on syntax. (Jespersen, 1914: vi).

Kruisinga (1925) の文法書は、多くのコーパスを新聞から引用し、同時に口語英語は英語文法書作成に賛同した学生が英語話者から聞き取った生の英語を利用したと述べている。

Poutsma (1928) はその多くを Shakespeare と欽定訳聖書から引用し、次のように述べている。

mainly descriptive and attempts to ‘give a methodical description of the English language as it presents itself in the printed documents of the last few generations. (Poutsma, 1928: viii).

アメリカ構造主義言語学の Fries (1940) は *American English Grammar* の執筆にあたり、印刷されたものは使わず手書きの手紙を主にコーパスとした。

took only those in the original handwriting of the persons selected as suitable subjects... used for the particular study some two thousand complete letters and excerpts from about one thousand more in order to base a preliminary sketch of the inflections and syntax of American English with special reference to social class differences. (Fries, 1940: vii).

そのコンピュータ利用以前のコーパス利用について、Svartvik (1992) は文法書、辞書の編纂にコーパスは有効に利用されたことを次のように説明している。

Jespersen was not the first linguist to grind general laws out of authentic language data recorded on slips of papers ... Poutsma and Kruisinga, the Dutch masters of English grammar, also used corpus illustrative examples. (Svartvik, 1992: 7).

また、Jespersen (1938: 213-15) は文法書を執筆するときに手書きのコーパスを集めてそれを有効利用していたことを次のように述べている。

I have found it practical to use small slips of paper ... it is impossible for me to put even a remotely accurate number of the quantity of slips I have had or still have. (Jespersen, 1938: 213-15).

その後、コーパス言語学はコンピュータテクノロジーの急激な発展に伴って世界的な広がりを見せ、その発展には言語学解析の目的で集積された多くのコーパス編纂とそれに伴う検索エンジンの開発がなされていった。そのコンピュータ時代にコーパスを利用して多くの文法書が出版された。その典型として文法書作成にコーパスが大いに活用されたのが、Quirk et al. (eds.) *A Grammar of Contemporary English* (1972), Quirk et al. (eds.) *A Comprehensive Grammar of the English Language* (CGE) (1985) であり文法書の模範となった。その CGE (1985: 33) では、主にどのコーパスが利用されているかを次のように述べている。

But we have also drawn on our research and that of others into the frequencies of language phenomena in several important corpora, preeminently:

- (a) The corpus of the Survey of English Usage (SEU), covering spoken as well as written texts of British English
- (b) The Brown University corpus, comprising samples of American printed English
- (c) The parallel Lancaster-Oslo/Bergen corpus (LOB), comprising samples of British printed English.

それら大型文法書に付随して発行された中型文法書が発行され、Quirk & Greenbaum (eds.) *A*

*University Grammar of English*<sup>6</sup>(1973)、Leech & Svartvik (eds.) *A Communicative Grammar of English*<sup>7</sup>(1975) 等がその代表であろう。

また Biber, et al. (eds.) *Longman Grammar of Spoken and Written English* (1999) は新しい視点で文法書を発行し、そのなかで「英語の話ことばと書きことばのコーパスを駆使して ‘Corpus Findings’ を統計的に記述しているユニークな文法書」と謳っている。

The LGSWE adopts a corpus-based approach, which means that the grammatical descriptions are based on the patterns of structure and use found in a large collection of spoken and written texts, stored electronically, and searchable by computer. (Biber, et al., 1994: 4).

## 1.4 コーパス言語学会

このようにコーパス言語学はコンピュータの急激な進化と関連して大きな飛躍を見せるのである。それは、ヨーロッパを中心とする地域の発展に端を発し、その後アメリカ、さらに日本へと波及してくるのである。

### 1.4.1 世界コーパス学会

コーパス言語学の研究大会は、ヨーロッパが中心で研究大会は例年5月前後に開催される ICAME (The International Computer Archive of Modern and Medieval English) がその中心を担っている。コーパス言語学の研究大会である ICAME は、大型文法書 (*A Grammar of Contemporary English* (1972), *A Comprehensive Grammar of the English Language* (1985)) を編集した4人の文法学者、R. Quirk, S. Greenbaum, G. Leech, J. Svarkvik が主催した学会であり、その後、多くの Corpus Linguists が参加して大きな発展を遂げてきた。

この ICAME は、*ICAME Journal: Computers in English Linguistics* という雑誌を、Bergen 大学の The HIT-centre—Norwegian Computing Centre for the Humanities から定期刊行している。さらに、ICAME 研究大会の後、Rodopi 社から *Language and Computers: Studies in Practical Linguistics* というシリーズで研究書を発行している<sup>8</sup>。

直接 ICAME 発行ではないが、コーパス言語学関係の専門書も多く出版され、Peter Lang 社の *Linguistic Insights*<sup>9</sup> もコーパス言語学関係の書物を扱っているが、John Benjamins Publishing Company は *Studies in Corpus Linguistics* シリーズを出し、多くの著書がコーパス言語学の論文を発表している。また John Benjamins Publishing Company 発行の *International*

---

<sup>6</sup> 邦訳『現代英語文法 大学編』池上嘉彦、他訳 紀伊国屋書店 1995

<sup>7</sup> 邦訳『現代英語文法 コミュニケーション編』池上恵子訳 紀伊国屋書店 1998

<sup>8</sup> そのシリーズ30の Kirk, J. (ed.) 2000. *Corpora Galore: Analyses and Techniques in Describing English*, に Kennedy, G. and S. Yamazaki. ‘The Influence of Maori on the New Zealand English Lexicon’ (33-44) を論文発表し、それはニュージーランドにおける原住民の言語マオリ語がニュージーランド英語にどのように影響しているかをコーパスを検索して研究したものである。

<sup>9</sup> そのシリーズ167で、Yamazaki, S, and R. Sigley. (eds.) 2013. *Approaching Language Variation through Corpora* を出版した。

*Journal of Corpus Linguistics* はコーパス言語学関係の雑誌として高く評価されている。

次に ICAME 開催の歴史を概観してみると、どの大学で開催され、どの大学でコーパス言語学が盛んに行われているのかが分かる。それぞれの大会の場所、大学、そして一部学会の内容も見ることできる。

### **ICAME Conferences & Proceedings:**

#### **1st ICAME (1979, March 29–30, Bergen, Norway)**

#### **"Grammatical Tagging of English Text Corpora in Machine-readable Form" Report in *ICAME News* 3 (1979), 9–14**

(第一回大会はノルウェーのベルゲン大学で 1979 年に開催されている。この学会の発足時には、高名な Corpus Linguists が一堂に会してコーパス言語学のその後の発展に大いに寄与したことが分かる。この大学には K. Hofland という LOB Corpus (Lancaster-Oslo/Bergen Corpus) 編纂で活躍したコンピュータ専門家がいた大学である。)

#### **2nd ICAME (1981, June 1–3, Bergen, Norway)**

#### **"Computer Corpora in Research and Teaching" No real proceedings, but the following includes most papers and later additions: Johansson, Stig (ed.) (1982). *Computer Corpora in English Language Research*. Bergen: Norwegian Computing Centre for the Humanities. [ISBN: 82-7283-027-2]**

(第二回大会もノルウェーのベルゲン大学で開催されていて、LOB Corpus 編纂の技術者 K. Hofland、オスロ大学の S. Johansson、そしてランカスター大学の G. Leech が揃い踏みしている。)

#### **3rd ICAME (1982, Sept 12–14, Stockholm, Sweden)**

#### **"Seminar on the Use of Computers in English Language Research" Report in *ICAME News* 7 (1983), 1–12**

(第三回大会は、スウェーデンのストックホルム大学で LJUNG, Magnus の主催で開催されている。)

この ICAME はほぼ毎年 5 月頃に開催されているが、第 1 回大会 (1979) から第 37 回大会 (2016) までの開催地を列記するとコーパス言語学の盛んな地域や大学が概観できる。主にヨーロッパ中心の地域で開催されていて、英語母語の英国、米国よりも北欧の開催が多いのは、英語学研究が主目的に加えて英語教育という観点が働いているのが伺える。

1<sup>st</sup>. Bergen, Norway (1979)、2<sup>nd</sup>. Bergen, Norway、3<sup>rd</sup>. Stockholm, Sweden、4<sup>th</sup>. Nijmegen, The Netherlands、5<sup>th</sup>. Windermere, England、6<sup>th</sup>. Röstånga, Sweden、7<sup>th</sup>. Amsterdam, The Netherlands、8<sup>th</sup>. Helsinki, Finland、9<sup>th</sup>. Birmingham, England、10<sup>th</sup>. Bergen, Norway、11<sup>th</sup>. Berlin, Germany、12<sup>th</sup>. , Ilkley, England、13<sup>th</sup>. Nijmegen, The Netherlands、14<sup>th</sup>. Zürich, Switzerland、15<sup>th</sup>. Århus, Denmark、16<sup>th</sup>. Toronto, Canada、17<sup>th</sup>. Stockholm, Sweden、18<sup>th</sup>. Chester, UK、19<sup>th</sup>. Newcastle, Northern Ireland、20<sup>th</sup>. Freiburg im Breisgau, Germany、21<sup>st</sup>. Sydney, Australia、22<sup>nd</sup>. Louvain-la-Neuve, Belgium、23<sup>rd</sup>. Göteborg, Sweden、24<sup>th</sup>. St. Peter Port, Guernsey, British Isles、25<sup>th</sup>. University

of Verona, Italy、26<sup>th</sup>. University of Michigan, Ann Arbor, Michigan, USA、27<sup>th</sup>. Helsinki, Finland、28<sup>th</sup>. Stratford-upon-Avon, UK、29<sup>th</sup>. Ascona, Switzerland、30<sup>th</sup>. Lancaster, UK、31<sup>st</sup>. Giessen, Germany、32<sup>nd</sup>. Oslo, Norway、33<sup>rd</sup>. Leuven, Belgium、34<sup>th</sup>. Santiago de Compostella, Spain、35<sup>th</sup>. Nottingham, UK、36<sup>th</sup>. Trier, Germany、37<sup>th</sup>. Hong Kong (2016)

筆者はこの ICAME 大会に 17 回大会から参加して、多くの Corpus Linguists との交流を持てたことは大きな喜びであり、また学問的な刺激を受けたとも言える。さらに日本のコーパス言語学会、ひいては大東文化大学との関係においてこの大会に参加したのがターニングポイントになった気がする。

#### **17th ICAME (1996, May 15–19, Stockholm, Sweden)**

**Liung, Magnus (ed.) (1997). *Corpus-based Studies in English*. Amsterdam: Rodopi**

(第 17 回大会<sup>10</sup>は、スウェーデンのストックホルム大学にて M. Ljung が主催した大会であった。)

### **1.4.2 日本におけるコーパス言語学の動向**

日本のコーパス言語学会も世界的な発展に呼応して盛んになり、その中心になったのが英語コーパス学会である。

#### **1.4.2.1 英語コーパス学会**

日本におけるコーパス言語学会の発足は、大阪大学の齋藤俊雄氏を会長とする英語コーパス研究会という名称で 1993 年に開催された研究会がその始まりである。初めは、大阪大学言語文化部の教員が呼びかけ人となり、過去にイギリスの UCL (University College London)、バーミング大学、ランカスター大学等で研究した主に関西を中心とした研究者を中心に、コンピュータを使った言語処理をする研究会がそのスタートであった。その英語コーパス研究会は大阪大学言語文化部が中心となり、他の研究者に呼びかけをした研究会であったが、全国的な研究会にすべく広く呼びかけを始めた。その大きな原動力となったのは、イギリスの大学・研究所に研修に行った時には、コンピュータを使った言語処理が盛んになって来た頃で、この新しい科学的な言語研究がとても新鮮で今後大いに盛んになることを予見していた研究者が多くいて、その学問的な動機が研究会発足に導いていくのである。

そのように、コンピュータを利用した言語処理の学会を発足させようという試みは、益々の広がりを見せ、下記にあるようにニューズレター第 1 号に第 1 回研究会の報告書を載せ、その初回研究会がとても盛会であったと綴っている。

---

<sup>10</sup> 筆者が最初に参加した ICAME 大会であり、後述するように S. Greenbaum との不思議な縁ができた大会であった。ICAME 大会には 13 回 (17, 18, 19, 20, 22, 23, 24, 25, 28, 30, 32, 33, 36) 参加し多くの高名な研究者に会ってコーパス言語学について意見交換ができたことと合わせて、後述のようにその後多くの Corpus Linguists を大東文化大学に招聘した。

=====

=====

JAECs Newsletter No. 1 (May 1, 1993)

英語コーパス研究会 会長 齊藤俊雄

事務局 560 豊中市待兼山町大阪大学言語文化部

今井研究室 Tel: 06-844-1151 内線 5711 (今井) / 5567 (齊藤)

Fax: 06-845-5290 (FAX ご利用の場合は今井宛または齊藤宛と明記下さい)

=====

=====

## 1. 第1回例会報告

英語コーパス研究会第1回例会は、去る4月3日(土)大阪大学で予定どおり開催されました。研究発表・講演への出席者は、北海道、関東、四国、九州からお越し頂いた方々を含め、56名の多数にのぼり、極めて活発な質疑応答がありました。また懇親会も41名が参加され、楽しい雰囲気の中で情報交換ができて有益な時を過ごすことが出来ました。皆様のご協力の賜と心から感謝いたしております。

その後、その研究会は年2回開催され、学会誌『英語コーパス研究』を年一回発行するようになった。筆者も第5回(1995)から参加したが、会を追うごとに会員が急増し、学会参加者も多くなっていったのを憶えている。世界的なコーパス言語学の興隆に連れて日本でもコーパスを利用した教育・研究が盛んになったことの証である。その後、「学会化」について議論され、1997年研究会から学会に変更され、名称も「英語コーパス学会」<sup>11</sup>(Japan Association for English Corpus Studies)(JAECs)とし、例会も「大会」と改められた。

日本におけるコーパス言語学の専門雑誌・著書は、JAECsが『英語コーパス研究』という雑誌を定期的に発行しているが、その他にJAECsの有志で齋藤俊雄、中村純作、赤野一郎編『英語コーパス言語学』(研究社、1998、2005改訂新版)という日本語で最初のコーパス言語学の専門書が発行された。また日本におけるコーパス言語学の現状を紹介する目的でRodopi社から英語版のコーパス言語学専門書、Saito, T., J. Nakamura, and S. Yamazaki. (eds.) *English Corpus Linguistics in Japan* (Rodopi 2002)が発行され、日本もRodopi社のコーパス言語学の仲間入りができたのである。この書物は、JAECsの総会で、新しく創設された第1回「英語コーパス学会賞」を2002年に受賞している。さらに、ICAMEと英語コーパス学会との関係において、Biber et al. *Corpus Linguistics: Investigating Language Structure and Use* (Cambridge, 1998)の邦訳をJAECsの有志で行い『コーパス言語学:言語構造と用法の研究』

---

<sup>11</sup> JAECsでは、2001年から東支部長、2007年から事務局長として学会に参加した。

(南雲堂 2003) の題名で出版した。また、ひつじ書房から堀正広、赤野一郎編『英語コーパス研究シリーズ』全7巻が出版され、多くの日本コーパス学会の研究者が参加・執筆している<sup>12</sup>。

その後、ICAME は、英語コーパス学会は日本においてコーパス言語学の活動に大いに貢献しているという感謝の意を込めて「ICAME の名誉団体会員」(Honorary Institutional Membership) という名誉ある賞を贈るという話を伝えてきた。それを強く推奨したのが、日本を訪問して、大東文化大学に招聘された Leech はじめ Svartvik 等であったことは ICAME に参加したときに聞いた。筆者は前述のように JAECS で活動しつつ ICAME にも数多く参加したので、その名誉ある賞、「ICAME の名誉団体会員」は日本から参加していた筆者が会長代理で受賞した。その時の報告を、「英語コーパス学会 Newsletter No. 41」(May 20, 2003) に投稿したものが残っている。

---

### ICAME 2003

大東文化大学 山崎 俊次

例年 5 月に開催される ICAME が、今年はリバプール大学の主催で 2003 年 4 月 23 日から 4 月 27 日にチャンネル諸島のガーンジー島 (Guernsey) で開催された。詳細な報告は『英語コーパス研究』に掲載の予定であるが、今年の ICAME 2003 “The Changing Face of Corpus Linguistics” について、例年と違う特徴と簡単な内容の説明をしたいと思う。

世界的な政治と社会の不安定な状況で開催が危惧された今年の学会であったが、新型肺炎 SARS (重症急性呼吸器症候群) の影響で、ニュージーランドからの 2 人とトロント (カナダ) の 1 人が参加を中止しただけのいつもと何ら変わったところのない学会の雰囲気であった。会議が行われたガーンジー島はリバプール大学レノフ教授 (Prof. A. Renouf) の出身地で、独自の司法、税法、行政を持ち、さらに独自の紙幣を有するフランスの影響が残る英国領の小さな島である。1996 年のストックホルム大会から参加して特に今回の大会で違った特徴は何かといえば、まず参加者の多様性である。それは、2 年程前にベルギーで開催されたときに ICAME の抜本的な改革と発展に対する参加者の直接討論があり、参加者の拡大と他のコーパス学会関係との密接な連携を重要な課題とて、Board も定期的に入れ替えを実施することが決定されたことの変化を表わしていると実感した。特にヨーロッパを中心に初めて参加するという若い研究者が多数いたことが特徴である。そして、他の学会との連携という点からは、すでに周知のように英語コーパス学会 (JAECS) が世界で初めて ICAME の名誉団体会員に認定されたことがあげられよう。現会長の命を受けて代理で名誉会員証を受領して感動したが、以下のような表現をしてあった。

*The International Computer Archive of Modern and Medieval English takes great pleasure in*

---

<sup>12</sup> そのシリーズの第 1 巻『コーパスと英語研究』に拙著「ニュージーランド英語におけるマオリ語の影響」(ひつじ書房 2019, pp. 146-163) を上梓した。

*conferring upon The Japan Association for English Corpus Studies Honorary Institutional Membership In recognition of its contribution to the advancement of English corpus linguistics On behalf of the Board*

Matti Rissanen, Chairman of the Board

3 February, 2003

さらに参加者の多様性の特徴という点から、アメリカからの参加者が以前より多かったことである。北米コーパス学会を中心的に運営しているミシガン大学から、Rita Simpson、Anne Curzan が参加して、話し言葉のアカデミック・コーパスの MICASE の概略を述べていた。オンラインでアクセスできるコーパスの今後の課題として、ランカスター大学から David Leech が赴任して POS タグ付け、lemmatization を行い、さらに実際の音声とのリンク付けを実施したいと発表していた。今後さらに北米コーパス学会との連携が密になると確信している。

---

この ICAME 大会が開催された「ガーンジー島はリバプール大学レノフ教授 (Prof. A. Renouf) の出身地で、独自の司法、税法、行政を持ち、さらに独自の紙幣を有するフランスの影響が残る英国領の小さな島である」と報告書に書いてあるが、ガーンジー島は英語史上非常に重要な島であり、英国王室属領であるとはいえ元来ノルマンディーに属しフランス語が主要言語であったことは、滞在してどことなくフランス、フランス語の雰囲気がある島であることが肌で感じられてとても興味のある島であった。

後述のように、その後、大東文化大学に招聘した欧米のコーパス言語学研究者は殆ど同時に JAECS の学会、講習会にも参加して講演、講習会をすることになるのである。そして、それらの研究者 (G. Leech, S. Johansson, J. Svartvik, G. Kennedy 等多数) は後になって JAECS の「名誉会員」として推奨されることになる。

#### 1.4.2.2 日本における辞書編纂

日本も遅ればせながら、英和辞典の編纂でコーパス活用が一般的になってきて、数多くの英和辞典でコーパス利用の記述が見られるようになってきた。大型英和辞典の研究社『新英和大辞典』は「まえがき」で次のように説明している。

「旧版の刊行以降に英語の語彙に加わった新語・新語義を大幅に採り入れたのはもちろんであるが、最近急速に発展したコーパス言語学 (corpus linguistics) の資料を活用して、重要語を中心として語義などを原則として頻度順に配列し、検索の効率化をはかった。」  
(まえがき、2002)

また、同じ大型英和辞典の大修館『ジーニアス英和大辞典』も同様に「はしがき」でコーパス利用を次のように説明している。

「編集にコンピュータコーパスを利用した。主としてインターネット上の新聞・雑誌・TV 番組などをもとに 2000 万語 (話し言葉 1000 万語 + 書き言葉 1000 万語) のコーパ



スを構築し、さらにこれに加えて新聞・雑誌・百科事典の CD-ROM 版等を利用し、語や表現の頻度、用例、語法記述の参考資料とした。生きた言語資料に基づく新しい英和辞典の誕生である...コーパスに基づいて、A ランク語（最重要語約 3500）、B ランク語（重要語約 5400）を選定、大英和に初めて頻度表示を採用した。ランクという窓から現代の世相をうかがうことができる。」（はしがき、2001）

中型英和辞典で画期的な辞書編纂をしているのは、JAECS の会員を中心に編纂した井上永幸、赤野一郎編小学館『ウィズダム英和辞典』（2003）で、そのまえがきで丁寧にコーパス利用を書いている。

「初めて本格的にコーパスを活用した *Collins COBUILD English Language Dictionary* (Collins ELT, 1987) が発刊されて以来、英国を中心にコーパスに基づく ESL/EFL 辞典の出版が盛んである。一方、日本国内の現状はでは、コーパス利用は頻度表示を中心とする一部の内容にとどまっている。この度、世に問う『ウィズダム英和辞典』は企画段階からコーパス言語学の方法論を導入し、英米の辞書や参考書に多くを依存する従来の英和辞典編纂法にも再検討を加え、最新の英語辞書学の成果を取り入れて編纂されたものである。これにより、日本の英和辞典はもとより、英米の ESL/EFL 辞典にも漏れていた多くの情報を盛り込むことができたと自負している。」

つまり、「コーパスを全面活用し、「生きた英語」を精緻に分析した英和辞典」という触れ込みでコーパスを駆使して辞典編纂したことを謳っている。

### 1.4.3 大東文化大学とコーパス言語学の関係

ニュージーランド国ヴィクトリア大学の大学院博士後期課程（PhD course）で研究していた際に、当時すでにコーパス言語学の碩学であった指導教員 G. D. Kennedy 教授の推薦で、17th ICAME (1996, May 15–19, Stockholm, Sweden) に参加する事ができた。これが ICAME に参加するきっかけになり、その後何度も参加し発表して、世界で活躍するコーパス言語学者との交流が始まった。

最初に近い交流を持ったのは、その 17th ICAME に参加していた中国系研究者、Alex Fang であったが、後にとっても有名な学者であることが分かって来た。コーパス言語学の始まりであるといえる研究所 UCL (University College London) の SEU (Survey of English Usage) で所長 S. Greenbaum のアシスタントをしていることが判明した。その 17th ICAME 学会後にロンドンの SEU に回って、S. Greenbaum、A. Fang 両氏と SEU でコーパス言語学の最近の流れ、さらに日本におけるコーパス言語学の状況等種々話げできた時の感動を今でも忘れない。その時の話題の一つに日本訪問があって、所謂 ‘The Gang of Four’<sup>13</sup>の中で日本に行

---

<sup>13</sup> この表現は、G. Kennedy のことばを援用して、筆者が文法学者でもある Quirk, Greenbaum, Leech,

った事がないのは、私だけであると S. Greenbaum が力説したので、即座に大東文化大学を訪問し講演をしてもらえないか打診したら、笑顔で快諾してもらった。その時の S. Greenbaum の笑顔は今でも脳裏にはっきり焼き付いている。しかし、S. Greenbaum はその年の 8 月モスクワ大学で講演中に突然倒れて急死したことを A. Fang から聞いて大きなショックを受けた。そこで、研究者であると同時にコンピュータ解析のためのプログラムを作成するコンピュータ専門家でもあった A. Fang に大東文化大学で講演をしてもらえないかお願いして、それが後に実現した。そして“A Brief Guide to COMPAID 2.0”というタイトルで講演と実習をしてもらい、その後 S. Greenbaum の死を悼んで Fang, Alex, C. and S. Yamazaki. 1997 "The International Corpus of English and TEFL: In Memory of Professor Sidney Greenbaum." (『大東語学教育論文集』 第 5 号) という題名で共同論文を書いてコーパス言語学界の重鎮の早すぎる死を悼んだ。

大東文化大学とコーパス言語学の碩学との交流が参加に行われた理由はもう一つある。当時英語学科主任であった筆者は、1999 年 4 月に外国語学研究科英語学専攻を発足させるのに伴い修士課程主任として英語コーパス学会の会長を歴任した齋藤俊雄先生の招聘に尽力した。その英語学専攻の特徴として言語学特にコーパス言語学の充実発展を旗頭にし、そして予算積算も充実した。そして 1996 年から参加を始めた ICAME で知り合ったコーパス言語学の研究者を連続して招聘し、講演会等を開催することを計画した。その歴史の一部を年代順に挙げていきたい<sup>14</sup>。

### 1. 1996 年： Alex Fang (SEU, UCL)； ‘A Brief Guide to COMPAID 2.0’

自分で開発したソフト (COMPAID2.0) を使って言語解析の実習をした。

### 2. 1999 年 12 月： G. Leech (Lancaster University)； 「どんな英語を学ぶ？教える？米語？英語？それとも？ —コーパス言語学の立場から—」

G. Leech はコーパス言語学の発展に寄与した一人で、全ては SEU で R. Quirk のもとと言語学の研究をしたのが始まりである。そこで S. Greenbaum、J. Svartvik 等と繋がってくる。その後、Lancaster University で多くの若手研究者を育成し、コーパス言語学の着実な発展に寄与するのである。

### 3. 2000 年 10 月： D. Biber<sup>15</sup>, R. Reppen (Northern Arizona University)； ‘Using Corpora to Investigate the Lexical Associations of Related Words’

Biber 等との関係は、既出の *Corpus Linguistics: Investigating Language Structure and*

---

Svartvik の 4 人を親しみを込めて呼んだ言い方である。

<sup>14</sup> 巻末の Appendices に講演・講習会の flyers の一部を挙げてある。

<sup>15</sup> コーパス言語学は、言語学の知識を有する事は当然であるが、コンピュータを有効活用する事から、IT の知識も必要とされる。筆者が経験した研究者の中で、D. Biber は優れて両方に深い知識を持ち自分でプログラミングしたソフトで大型コーパスの検索をする理系・文型の知識に造詣の深い教養人であるというができれば。

Use (Cambridge, 1998) の邦訳を JAECS の有志で『コーパス言語学：言語構造と用法の研究』（南雲堂 2003）の題名で出版したことが大きい。

Biber, et al. は Biber, et al. (eds.) *Longman Grammar of Spoken and Written English* (1999) を出版している。

**4. 2000 年 11 月：G. Leech (Lancaster University)；「コンコーダンスラインから何が見えるか」**

G. Leech は大東文化大学への招聘の回数が最も多い学者であった。

**5. 2002 年 2 月：P. Rayson (Lancaster University)；「BNC 編集とタグ付けソフト」**

P. Rayson は Lancaster University の Director of the UCREL interdisciplinary research centre を長く務めたコンピュータの専門家で BNC 等多くのコーパス編纂を主導してきて、現在は Professor of Natural Language Processing である。

**6. 2002 年 4 月：J. Aarts (University of Nijmegen)；‘English Corpus Annotation and Grammatical Analysis: Past, Present, and Future’**

長い間、*Language and Computers: Studies in Practical Linguistics* (Rodopi) シリーズの編集 (W. Jeyts と共同編集) を手掛けてきたコーパス言語学の碩学である。

**7. 2002 年 10 月：S. Johansson (University of Oslo)；‘The Impact of Corpora on the Study of English Grammar’**

Stig Johansson は University of Bergen の K. Hofland と共同でコーパス編纂・検索を率先させてきて、LOB コーパスを編纂し、また ICAME の第 1 回会議を主催した第一世代コーパス言語学の学者である。

**8. 2002 年 11 月：S. Hofmann (University of Zurich)；「BNCWeb の利用」**

Sebastian Hofmann はチューリヒ大学で他の 2 人（筆者はこの 3 人を「3 人賢者」(three wise men) と名付けて称えたことを憶えている）と共に BNC の annotation, tag 付けなど、後の BNC 利用の作業を手掛けたコンピュータの知識もあるコーパス言語学研究者である。

**9. 2006 年 11 月：S. Hofmann (Lancaster University)；講演会 ‘Grammaticalization of Complex Prepositions in English’、講習会「BNCWeb の効率的な利用法」**

Sebastian Hofmann はその後、G. Leech のいる Lancaster University に移って研究をつづけた。その時の flyer が残っている。

**10. 2007 年 7 月：G. Leech and N. Smith (Lancaster University)；‘Corpus Linguistics and the Recent History of English Grammar’, ‘Exploring Recent Change in Standard Englishes: Methodology, Findings and Implications’**

G. Leech は若い研究者である N. Smith を伴ってこの講演をした後、Leech, G., M. Hundt, C. Mair, and N. Smith. *Change in Contemporary English: A Grammatical Study*

(Cambridge, 2009) を出版していて、その研究結果を出版前に大東文化大学で発表していると言われた時の感動を覚えている。

#### 11. 2014年6月：G. Leech (Lancaster University)：‘Corpus Linguistics and its Contribution to Descriptive Grammar’

大東文化大学の外国語研究科との共催で東松山校舎で初めてコーパス言語学の講演会を行った。講演後、G. Leech の滞在先まで電車で同行しているときに、G. Leech の住んでいる古いお城のような家は英国中部の地方では観光名所になっていて、筆者はまだ一度も訪問したことがないことを告げたら歓待するので訪問するようにと約束してくれた事をはっきり憶えている。しかし、英国に帰国後まもなく2014年8月に Lancaster University の研究室で倒れて帰らぬ人となった。筆者は大きなショックを受けたのと共に日本の大学において最後の講演を大東文化大学でしたのかと思いつつ大きな喪失感を憶えた。

その他のコーパス言語学の研究者で大東文化大学に招聘したのは、年月は明らかでない（恐らく1998年9月頃であろう）が、*An Introduction to Corpus Linguistics* (Longman, 1998)、*Structure and Meaning in English* (Longman, 2003)、*The New Zealand Oxford Dictionary* (Oxford, 2005) 等を出版した G. Kennedy がいる。また、‘The gang of four’ の中の一人、J. Svartvik も大東文化大学を訪れて講演を行っている。

大東文化大学に外国語学研究科英語学専攻が開設され、齋藤俊雄先生が赴任されたことで、学部・大学院のコーパス言語学研究が盛んになり、院生・学生の研究論文をまとめた論文集『*Studies in English Corpus Linguistics*』を2000年から発行し、No. IX まで年次発行したのは大きな功績であった。

また、すでに述べた Yamazaki, S, and R. Sigley. (eds.) *Approaching Language Variation through Corpora* (Rodopi, 2013) (Linguistic Insights: Studies in Language and Communication, Vol. 167) は出版に際して、‘*A Festschrift in Hour of Toshio Saito*’として齋藤俊雄先生の退職記念号として発行し、Foreword で G. Leech が齋藤俊雄先生と日本コーパス学会の功績を丁寧述べている。この書籍には過去に大東文化大学に招聘した海外のコーパス言語学の研究者も多く論文掲載をしていてコーパス言語学集大成の一冊になった。投稿者を列記する： S. Johansson、R. Sigley、S. Hoffmann、G. Kennedy、T. Fukaya、S. Yamazaki、M. Rissanen、T. Saito、Y. Iyeiri、S. Tsukamoto、P. Peters、M. Hori、M. Stubbe。この書物を Linguistic Insights シリーズで出版したが、過去の例から出版には費用が課されると聞いていたが、全く無料で出版してくれたことに敬意を表したい。

このように、大東文化大学とコーパス言語学研究者との繋がりは強く深いものであったことが分かり、日本のコーパス言語学の発展に大きく寄与したと自負している。コーパス言語学研究者のみならず若手研究者にも、世界の研究の流れ、方向性を知ってもらうのみならずコーパス編集、検索ソフト開発といったコンピュータの有効活用に関する講習会も何度となく行ったことで大東文化大学が貢献したことは大きな意義があると思われる。

## 第2章 副詞類における前置詞省略の言語学的検証

### 2.1 副詞類とは何か

副詞 (adverb) は品詞の中で本来動詞、形容詞、副詞を修飾する語であると言われるが、しかし前置詞、名詞句なども修飾することがある。Quirk et al. (1972; 1985) は、副詞を 1) 「修飾要素」 (modifier) として一般的に句の中の特定の語を修飾する副詞 (例: *very young, quite strong*) と、2) 「副詞類」 (adverbial) と呼ぶ副詞の 2 つに大きく分類している (安井稔編, 1996: 30)。副詞は、その形態から単一形副詞 (例: *very, immediately*) と「前置詞+名詞句」といったような前置詞句の形が主にみられる複合形副詞に分けることができ、「時を表す」意味の用法 (例: *on Sunday, in February*) も多く見られる。これら副詞類は節の中では独立した構成素であり、一般的には移動可能である (‘An adverbial is generally mobile, *ie* is capable occurring in more than one position in the clause’ (Quirk et al. 1972: 349))。

本論文は、副詞類の中で、「時を表す前置詞句の副詞類」を扱い、普通は前置詞を伴う前置詞句でどのような場合にその前置詞が省略されるのか「時を表す副詞類における前置詞の省略」を、Leech (1992: 107) が述べたコーパス言語学の特徴に沿って述べている。分析に利用するコーパスは ‘Brown Family Corpora’ の中の次の 5 種類である。アメリカ英語 2 種類 (Brown Corpus: Standard Corpus of Present-Day Edited American English、Frown Corpus: Freiburg-Brown Corpus of American English)、イギリス英語 2 種類 (LOB Corpus: Lancaster-Oslo/Bergen Corpus of British English、FLOB Corpus: Freiburg-LOB Corpus of British English)、ニュージーランド英語 (WWC Corpus: Wellington Corpus of Written New Zealand English) のそれぞれ 100 万語コーパスを使って検証していく。

### 2.2 副詞類における前置詞の省略の諸理論

本論文は「時を表す前置詞句の副詞類における前置詞の省略」において、どのような前置詞を、どのような場合に、どの位置で省略するのかをアメリカ、イギリス、ニュージーランド英語 100 万語コーパスを使って検証する。その前置詞省略について、過去に様々な研究がなされたが、その論点を整理しておきたい。

まず前置詞省略は、absence of preposition, preposition omission, ellipsis, abbreviation 等様々な表現をしているが、Quirk et al. (1985: 692) は次のような前置詞省略の可能性を示している。

1. 副詞的目的格 (距離・時間) : We walked (for) three miles. Won't you stay (for) the night.
2. Of+名詞→形容詞句 (年齢・大小) : The children are (of) the same age. This suitcase is (of) the same size as that.
3. That 節、疑問詞の節・句→前置詞の目的語 : I'm sure (of) that he is honest. I'm worried (about) how the project was carried out.
4. 動名詞の前 : I had a lot of difficulties (in) finding the works on tax. He spent most of his space (in) reading.
5. 不定詞の中→副詞用法の不定詞に近づく : The moon is not a good place to live (on). He didn't have the money to buy the tent (with).

しかし、本論文はどのような条件で副詞類における前置詞の省略が起こるかを検証するためにコンピュータコーパスを検索して検証するものである。

Quirk et al (1985: 692) は、時を表す前置詞句の副詞類では使われるべき前置詞が省略されることがあり、前置詞句というより名詞句の形態を取り、直示的 (deictic) 語彙 (例: last, next, this) の直前では前置詞は省略すると説明している。

例: last Sunday, \*on last Sunday  
next Monday. \*on next Monday

そして、「前置詞が省略された時を表す前置詞句」は

On the whole, the sentence without the preposition tends to be more informal and more usual (Quirk et al. (1985: 693).

と述べている。つまり、副詞類における前置詞の省略は形式ばらない表現 (informal) で多く見られる現象であると述べている。例文: I'll see you (on) Monday で前置詞省略は、'variable overt vs. zero-marking of the adverbial by a preposition'の変異であると言えるのだろう。

また、Leech and Svartvik (1994)は

The preposition is also sometimes omitted directly before days of the week: I'll see you (on) Wednesday, then ... This omission is especially common in <informal AmE>. Leech and Svartvik (1994: 84).

と述べ、informal なアメリカ英語で省略することが多いと述べている。

地域変種としての英語において、Algeo (1988)は

British [English] usually requires a preposition (*on*) with days of the week, whereas American [English] can have the preposition or omit it. (Algeo, 1988: 14).

と述べ、イギリス英語では前置詞 *on* は必要である (例: on Monday) が、アメリカ英語では前置詞を付けるかどうかは任意となっていると解説している。

文体的な変種として、Sonoda (2002) は

*on* and *for* are omitted most frequently in informal styles. (Sonoda, 2002: 19).

と述べ、informal な形式では *on*, *for* はしばしば省略されると説いている。

### 2.3 Formality Index とは何か

副詞類における前置詞の省略について、概して省略は形式ばらない表現 (informal) で多く見られる現象である (Quirk et al. 1985: 693) と述べている事を勘案する時に、形式的 (formal)、形式ばらない (informal) とはどのような基準で図れるのか、またどの英語、どのテキストカテゴリーの英語がそのどちらに該当するのかをコーパスを使って統計的に

述べた論文がある。Sigley (2013) はアメリカ英語、イギリス英語、ニュージーランド英語を使って、英語変異における Formality を勘案して Formality Index を作り、どちらがより formal であるか、あるいは formal でないかについて詳細に研究している。その論文で Formality Index の分析の出し方を説明し、その後、結果を図に表してどの英語が more formal か、あるいは less formal かを説いた論文である。その論文は Sigley (1997: Text categories and where you can stick them: a crude formality index. *International Journal of Corpus Linguistics* 2, 199-237) で扱った Formality Index を簡略化して述べた論文である。少し長い引用になるが以下のように述べている。

First, a test corpus is chosen for the purpose of defining the measure... Second, a range of different linguistic feature counts are defined (in practice, mostly in terms of a list of functionally-related wordforms), each associated with some linguistic function(s) expected to be correlated either with a “formal” or an “informal” style of language... These counts are normalised for text length and then converted to standard normal variates (with an average of 0 and a variance of 1). Third, the variation shared between the normalised counts is extracted as a small number of summary dimensions using principal component analysis... The pattern of the count loadings on these dimensions forms one way of interpreting the dimensions. Finally, index scores are calculated on each principal component retained, as a weighted sum of the standard normal variates for each count in each text... At this stage, the interpretation of the principal components used can be checked by investigating the distribution of text categories... or alternatively, the linguistic characteristics of the texts occupying the most extreme positions on each dimension... The index defined on the test corpus can then also be applied to the texts in each of the corpora to be compared. (Sigley 2013: 76).

それによって出された結果が以下の図に示されている。それによると 5 つのコーパスの中で、Brown Corpus (アメリカ英語) が最も formal で、WWC (ニュージーランド英語) が最も informal であることが図から読み取れる。

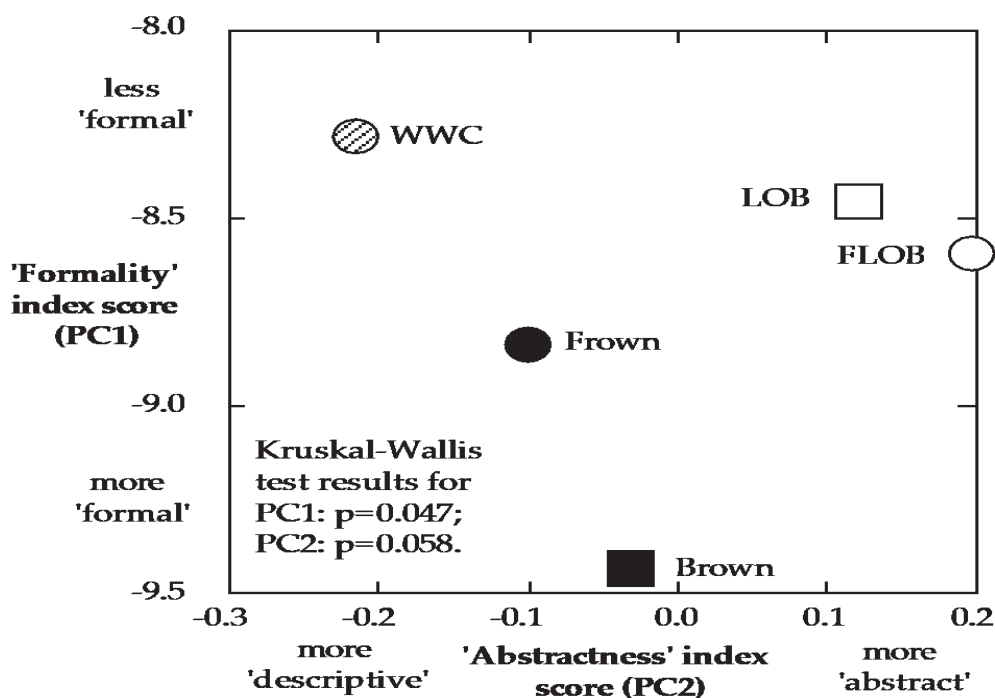


Figure 1 Mean factor scores for written corpora (Sigley 2013: 92)

また、WWC はより ‘descriptive’であり、イギリス英語（LOB Corpus, FLOB Corpus）は他の英語よりも ‘abstract’であること言う結果が出ている。

さらに、同じ英語でもイギリス英語（LOB Corpus, FLOB Corpus）の二つのコーパスにおいては大きな差は見られないが、しかし、アメリカ英語では古い英語（Brown Corpus）は新しい英語（Frown Corpus）より formal であるという結果は興味ある結果である。

there is some evidence of change in institutional norms, away from a noticeably more “formal” style in Brown to a more “informal” style in Frown. (Sigley 2013: 111).

さらに、後に詳しく述べる 15 のテキストカテゴリー（K-R の創作散文は一つにまとめている）の中で、どのテキストカテゴリーがより formal であるか否かの数値を出した Formality Index を詳細に分析してみると、大きな違いが見られる。下の分析図からテキストカテゴリーの K-R（創作散文）が極端に informal であることが明らかである。そして、カテゴリーH（政府文書等）、カテゴリーJ（学術）がより formal であることは容易に推測がつくが、カテゴリーA（新聞記事）とカテゴリーD（宗教）の一部のコーパスも formal であるという事は興味ある結果である。



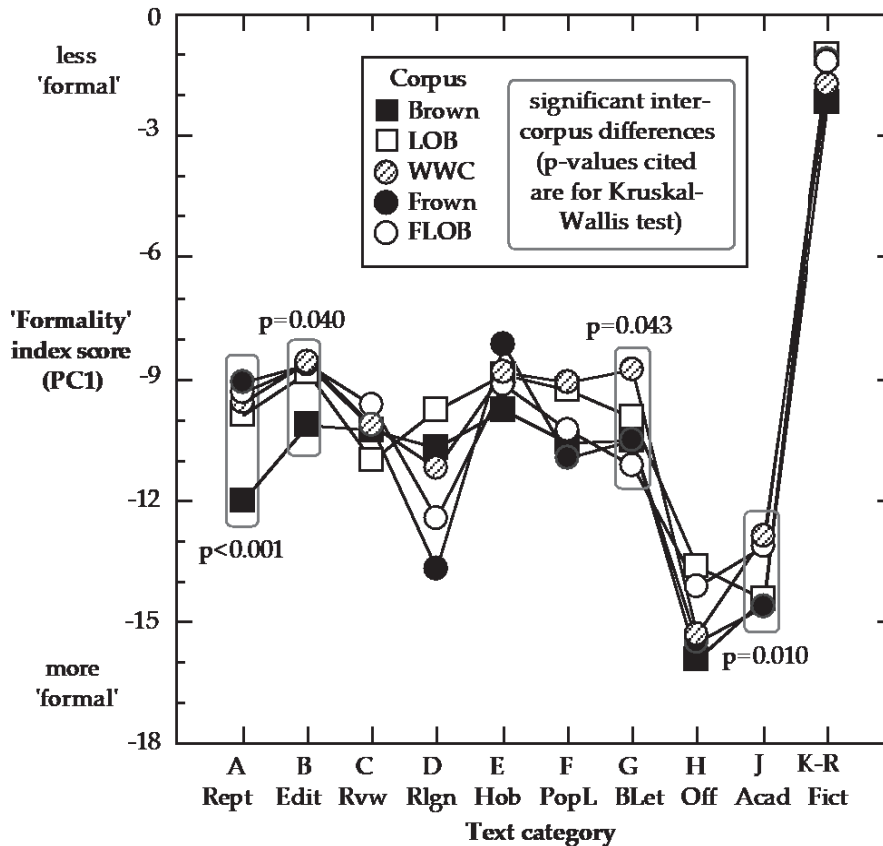


Figure 2 Mean text category scores for PC1, by corpus (Sigley 2013: 98)

この統計処理を行う際に使ったプログラムは、後に詳しく述べるように言語変異理論において使われたプログラムで、まだ日本では一般的ではなかった。何故一般的に使われていなかったかは後に述べるが、ここで使ったのは VARBRUL、特にその中で GoldVarb を使ったと説明がある<sup>16</sup>。

#### Variable-rule (= logistic regression) models :

(cf. Paulillo 2002; Tagliamonte 2006; Hoffmann & Sigley 2013:118)

Modelling allows us to correct for associations between factors. A logistic regression model relates the log odds of variant occurrence (in this case, preposition omission) to an overall tendency, plus a sum of independent effects from each factor (from factor groups 1 ... k) present in each cell.

Variable-rule model :

$$\ln\left(\frac{P_{cell}}{1-P_{cell}}\right) = \ln\left(\frac{P_o}{1-P_o}\right) + \sum_{i=1}^k \ln\left(\frac{P_i}{1-P_i}\right) X_i$$

<sup>16</sup> この VARBRUL については、使用したプログラム GoldVarb も併せて後のページで詳しく説明する。

Stepwise comparisons of model log-likelihood are used to establish the significance of each factor group's contribution to the model.

Variable-rule programs (such as VARBRUL 2000 or GoldVarb X) report these effects on a logit scale: the overall tendency is reported as an “input probability” ( $p_0$ ), and factor effects are reported as “factor weights” ranging from  $pi=0$  (categorically disfavouring the modelled variant) through  $pi=0.5$  (representing no effect) to  $pi=1$  (categorically favouring the modelled variant).

## 2.4 前置詞省略の仮説

前置詞を省略することで英語変異が起こるが、その省略が頻繁に起こるのはどのような条件で起こるのかということについて仮説を立ててみた。

### 1. 地域差によるヴァリエーション

イギリス英語より、アメリカ英語の方がより多く前置詞省略が起こる。

(Variation by regional variety: Prepositions are more often omitted in American English than British English.)

### 2. テキストカテゴリー (genre) によるヴァリエーション

フォーマルテキスト (formal text) とインフォーマルテキスト (informal text) における前置詞省略には明白な違いが生じる。フォーマルテキストは前置詞の zero-marking を避け、前置詞は「顕在範疇」 (overt marker) であるとする。

(Variation by genre: More formal genres should more often favour overt markers. More formal genres should disfavour zero-marking and favour overt prepositions.)

### 3. 副詞類の文中に位置によるヴァリエーション

文の位置によって副詞類における前置詞の省略に違いが見られる。文末にある副詞類は、文頭、あるいは文中にある副詞類よりも前置詞の省略がおこる。

(Variation with sentence position of the adverbial: Preposition omission may be expected with higher frequency in sentence-final adverbials than in sentence-initial or sentence-medial adverbials.)

大きく分類してこの3つの前置詞省略のヴァリエーションに沿って、Brown Family Corpora 中の5つのコーパスを使って、「時を表す副詞類における前置詞の省略」でどのような前置詞が、どのような位置で、どのような条件で省略されるかを検証していく。

## 第3章 副詞類における前置詞省略の分析・結果

### 3.1 Brown Family Corpora とは何か

本論文で使うコーパスは、Brown Family Corpora である。それは、1961年のアメリカ英語の書き言葉を15の詳細なテキストカテゴリーに分けて、500テキスト、各テキスト約2,000語のアメリカ英語を100万語コーパスとして編纂したBrown Corpusが範となっているサンプルコーパス<sup>17</sup>の総称である。そのBrown Family Corporaのうち、アメリカ英語、イギリス英語、ニュージーランド英語の5つのコーパスを使って分析する。基本的にアメリカ英語コーパス、イギリス英語コーパス、ニュージーランド英語コーパスと地域限定コーパスはそこで生まれ育った英語母語話者であることが条件であり、成人になってからその地域に移住した人の英語は入れないのが原則である。Brown Corpusのマニュアル(Manual)には、つぎのように明確にアメリカ英語を規定している。

This Standard Corpus of Present-Day American English consists of 1,014,312 words of running text of edited English prose printed in the United States during the calendar year 1961. So far as it has been possible to determine, the writers were native speakers of American English. (Francis and Kučera, 1964: 1).

また、LOB Corpusにも同様にイギリス英語としてのコーパスの範囲を明確にBritish Englishと述べて規定している。

The aim of our project has been to assemble a British English equivalent to the Brown University Corpus of American English...To facilitate a combined use of the two corpora, an attempt has been made to match the British English material as closely as possible with the American corpus. (Johansson, 1978: 1).

さらに、WWC Corpusも同じようにニュージーランド英語の書き言葉をコーパスとすると述べているが、ニュージーランドの総人口が3.3million(編集当時)と他の国と比較して少ないこともあり創作散文のテキストカテゴリーに変更が加えられていると述べている。

The basic aim of the Wellington Corpus of Written New Zealand English is to provide a computerized sample of written New Zealand English which will allow direct comparisons with the Brown University Corpus of American English, the Lancaster-Oslo/Bergen Corpus of British English...where the Wellington Corpus has Categories K and L for fiction, ... The fiction published in New Zealand, generally speaking, aims at a smaller market and is more consciously literary. (Bauer, 1993: 1-2).

---

<sup>17</sup> 均衡的収集法で特定の時点で切り取られた言語母集団を「サンプルコーパス」と呼び、このBrown Family Corpora、大型コーパスではBNC(British National Corpus)がその代表である。一方大規模収集法で拡大する言語の姿に注視して不断に集めたコーパスを「モニターコーパス」と呼びBank of Englishがその代表である。

下の表 1 に使用した Brown Family Corpora 100 万語コーパスの一覧を記す。

表 1 使用した 100 万語コーパスの一覧

Corpus	Variety
Brown Corpus	1961 American English
LOB Corpus	1961 British English
WWC Corpus	1986 New Zealand English
Frown Corpus	1992 American English
FLOB Corpus	1991 British English

それぞれのコーパスは普通その通称で呼ばれる事が多いが、正式名称とその編集に関わった主な研究者を以下に述べる。アメリカ英語の Brown Corpus, Frown Corpus、それにイギリス英語の LOB Corpus, FLOB Corpus は約 30 年の間隔をもって集積されたコーパスである。ニュージーランド英語コーパスは、オーストラリアの英語コーパス ACE (Australian Corpus of English 1986) と同時期に編纂されたコーパスである。

- **Brown Corpus (Standard Corpus of Present-Day Edited American English):** アメリカ、ブラウン大学 W.N.Francis, H.Kučera
- **LOB Corpus (Lancaster-Oslo/Bergen Corpus of British English):** イギリス、ランカスター大学 G.Leech、ノルウェー、オスロ大学 S.Johansson、ノルウェー、ベルゲン大学 K.Hofland
- **WWC Corpus (Wellington Corpus of Written New Zealand English):** ニュージーランド、ヴィクトリア大学 L.Bauer
- **Frown Corpus (Freiburg-Brown Corpus of American English):** ドイツ、フライバーグ大学 C.Mair, M. Hundt, A. Sand, P. Skandera
- **FLOB Corpus (Freiburg-LOB Corpus of British English):** ドイツ、フライバーグ大学 C.Mair, M. Hundt, A. Sand, R. Seimund

### 3.2 Brown Family Corpora のテキストカテゴリーとサンプル数

Brown Family Corpora のテキストカテゴリーは 15 テキストカテゴリー (WWC は一部統合して編集) に分けられていて、総サンプル数 500、一つのサンプル約 2,000 語を収集した合計 100 万語のコーパスである。Frown, FLOB コーパスは Brown コーパスを範としているのでサンプル数は同じであるが、LOB コーパスはテキストカテゴリー E, F, G が Brown コーパスと若干違っている (WWC コーパスは LOB コーパスを範としている)。

表2 Brown Family Corpora のテキストカテゴリーとサンプル数

Text category	Text samples (Brown)	Word count	LOB	Frown	FLOB	WWC
A. Press: Reportage	44	88,000	44	44	44	44
B. Press: Editorials	27	54,000	27	27	27	27
C. Press: Reviews	17	34,000	17	17	17	17
D. Religion	17	34,000	17	17	17	17
E. Skills and hobbies	36	72,000	38	36	36	38
F. Popular lore	48	96,000	44	48	48	44
G. Belles lettres and biography	75	150,000	77	75	75	77
H. Miscellaneous / official documents	30	60,000	30	30	30	30
J. Leaned	80	160,000	80	80	80	80
K. General Fiction	29	58,000	29	29	29	99
L. Mystery and detective fiction	24	48,000	24	24	24	29
M. Science fiction	6	12,000	6	6	6	
N. Adventure and western fiction	29	58,000	29	29	29	
P. Romance and love story	29	58,000	29	29	29	
R. Humour	9	18,000	9	9	9	
Total	500	1,000,000	500	500	500	500

その 15 テキストカテゴリーは大きく二つに分けられ、A-J: Informative prose (情報散文) と呼び、K-R: Imaginative prose (創作散文) と呼ぶ。更に A カテゴリーから R カテゴリーもそれぞれ細分化されていて、日本語でそれぞれ表記してみると以下のようなになる (サンプル数の率もあげてある) :

### A – J: Informative prose (情報散文) 約 75%

A:新聞記事(8.8%)、B:新聞論説(5.4%)、C:新聞書評等(3.4%)、D:宗教(3.4%)、E:技術  
娯楽(7.2%)、F:一般実用(9.6%)、G:書簡・伝記等(15%)、H:雑(政府文書等)(6%)、  
J:学術(16%)、

### K – R: Imaginative prose (創作散文) 約 25%

K:一般小説(5.8%)、L:推理小説(4.8%)、M:SF 小説(1.2%)、N:冒険小説等(5.8%)、P:恋  
愛小説(5.8%)、R:ユーモア小説(1.8%)

## 3.3 分析方法

この 5 つのコーパスを使って「時を表す前置詞句の副詞類における前置詞の省略」を分  
析するためには、いくつかの分析手順があり、その手順をそれぞれ、**Data extraction, coding  
and analysis** (データ抽出、コーディング(符号化)、分析)と呼ぶことができる。その三つ  
の手順について順を追って詳細に述べてみる。

### 1) データ抽出・コンコーダンス作成 (data extraction; concordancing)

この手順では、まず 5 つのコーパスについて下記のような方法でコンコーダンスを使っ  
て KWIC コンコーダンスを作成する (KWIC: Key word in Context) (コンコーダンス :  
KWIC Concordance for windows<sup>18</sup>を主に使用する) ことから始める。例えば Key word とし  
て :

- words containing *day* (*Friday, Sunday, etc*)
- abbreviations for weekday names (*Sun, Mon, etc*)
- month names and abbreviations (*April, Feb, etc*)
- season names (*Summer, Winter, etc*)
- four-digit numbers corresponding to years

等のコンコーダンスラインを作成する。

下に Brown Corpus を使って、‘April’ を key word として作成したコンコーダンスラインを  
示す。大文字、小文字は関係なく KWIC が作成できる。このコンコーダンスラインの左端  
の BROWN\_A.TXT は Brown Corpus のテキストカテゴリー A であることが分かる。その  
情報は右端にも、Category として A が書かれている。

---

<sup>18</sup> 日本大学塚本聡氏が考案した日本人向けのコンコーダンスで無料で使える。

File	line	Left	Key	Right	Reference	Category
BROWN1_A.TXT	303	e a bill which would order the referendum on the	April	4 ballot, when Texas votes on a US&S senator.	A02 1030	A
BROWN1_A.TXT	310	majorities, before March 1 to be printed on the	April	4 ballot, Roberts said. Opponents generally ar	A02 1100	A
BROWN1_A.TXT	1151	ose four-year terms will expire a week after the	April	18 primary election, received carved wooden elp	A06 1830	A
BROWN1_A.TXT	1154	y be other 1981 state committee referendums come	April	18, but they will be leaving by choice of the Rep	A06 1860	A
BROWN1_A.TXT	3328	rt bridge and fashion show at 1 P& M& on Monday,	April	17, in the chapel assembly room, Wayne Mrs&	A10 0830	A
BROWN1_A.TXT	3358	es and sculpture at the Philmont Country Club on	April	8 and 9. A preview party for sponsors of the	A10 0930	A
BROWN1_A.TXT	3360	sors of the event and for the artists is set for	April	8. The event will be open to the public the follo	A10 0950	A
BROWN1_A.TXT	3848	rth into Wilmette in Prairie ST&S JOHN&S, MID&S,	APRIL	19. - A jury of seven men and five women found 2	A21 0010	A
BROWN1_A.TXT	4009	ent claims have been found exaggerated. _HAVANA,	APRIL	19. - Two Americans and seven Cubans were execute	A21 1610	A
BROWN1_A.TXT	4313	oodyard Bros& Grocery, 2025 E&S Burnside St&S, in	April	of 1959. Her husband, who was sentenced to 15	A23 0860	A
BROWN1_A.TXT	4315	ears in the federal prison at McNeil Island last	April	for robbery of the Hillsdale branch of Multnomah	A23 0700	A
BROWN1_A.TXT	4786	e had told the jury that he made an agreement in	April	with Stein to supply and supervise janitors in M	A25 1420	A
BROWN1_A.TXT	4994	the March pickup, dealers are optimistic that the	April	-June quarter will equal or top last year. The Mar	A26 1480	A
BROWN1_A.TXT	5405	s. _WASHINGTON - Consumer spending edged down in	April	after rising for two consecutive months, the Gov	A28 1740	A
BROWN1_A.TXT	5408	stores dropped to slightly under \$18 billion in	April	down 1% from the March level of more than \$18.2	A28 1770	A
BROWN1_A.TXT	5409	from the March level of more than \$18.2 billion.	April	sales also were 5% below those of April last yea	A28 1780	A
BROWN1_A.TXT	5410	billion. April sales also were 5% below those of	April	last year, when volume reached a record for any m	A28 1790	A
BROWN1_A.TXT	5412	takes into account such factors as Easter was on	April	2 this year, two weeks earlier than in 1980, and	A28 1810	A
BROWN1_A.TXT	5415	partment officials were inclined to explain the	April	sales decline as a reaction from a surge of consu	A28 1840	A
BROWN1_A.TXT	5480	ave come to Greer Garson are countless. Just this	April	she was nominated for the seventh time for an Aca	A29 0600	A
BROWN1_A.TXT	8455	ne of the main reasons for National Library Week,	April	16-22, and for its theme: "For a richer, fuller	A44 1830	A
BROWN1_B.TXT	554	uban invasion mounted by the administration last	April	, which, we trust, is not symptomatic of the meth	B03 1620	B
BROWN1_B.TXT	4893	PROPER BOSTONIAN? F&S L&S LUCAS'S article in "SR,	April	1 issue seemed to be a very fair and objective an	B27 0390	B
BROWN1_B.TXT	5017	ress, which is publishing the book this month of	April	, recognize and freely acknowledge that they invi	B27 1620	B
BROWN1_B.TXT	2741	daily reading", Pope Pius the Sixth, at Rome, in	april	1778, wrote the following: "The faithful should	D16 0060	D
BROWN1_E.TXT	2186	urr-filled Fishing Derby in Hot Springs, Arkansas,	April	19-23, and the Arkansas Band Festival in Hot Spri	E12 1780	E
BROWN1_E.TXT	2187	3, and the Arkansas Band Festival in Hot Springs,	April	20-22. #THE WEST& A WESTERN VACATION is pra.	E12 1790	E
BROWN1_E.TXT	3784	Prokofieff to honor his 70th Anniversary for the	April	issue of 'Sovietskaya Muzyka' is accepted with p	E22 0020	E
BROWN1_E.TXT	3876	Traditionalism for the creative art of music.	April	10 marked a memorable date in New York's musical	E22 1910	E
BROWN1_F.TXT	625	nter: the cabbage and squash keep until March or	April	. There is never enough corn, peas or strawberries	F04 0660	F
BROWN1_F.TXT	2752	eived imposition would be religion itself. ## On	April	17, 1610, the sturdy little three-masted bark, <D	F16 0010	F
BROWN1_F.TXT	3065	d that the ice on the Red was five feet thick. In	April	came a rapid thaw that produced high waters whi	F17 1830	F
BROWN1_F.TXT	3478	od will, and in exchange for armed support in the	April	election campaign, bestowed upon O'Banion a thir	F20 0020	F
BROWN1_F.TXT	3536	mity and in payment for the loan of gunmen in the	April	election, Torrio had given O'Banion a slice of C	F20 1800	F
BROWN1_F.TXT	3831	minutes after five on the morning of (Wednesday),	April	18, 1908, San Francisco was shaken by a severe e	F21 1710	F
BROWN1_F.TXT	5011	t until the adjournment of the legislature in mid-	April	Paltrey told his wife of his intentions for the	F28 0840	F

図3 Brown Corpus の'April'のコンコーダンスライン

このコンコーダンスラインの key word: April の左右を読むと、文章として読むことができる。上記の例の 6 番目のコンコーダンスラインの一部を取り出してみると、

and sculpture at the Philmont Country Club on **Friday** 8 and 9.

という文章を読む事ができ、on という前置詞が文末に来ている事が分かる。

次のコンコーダンスラインは、コンコーダンスで出たデータをテキストファイル(\*\*.txt) にコピーしたファイルであり、これから次の手順の coding に移る前のファイルである。このコンコーダンスラインは、Brown Corpus の 'May' を key word とするコンコーダンスラインで右側にはカテゴリーは記されていないが、左の 'Brown1\_A.TXT' で Brown コーパス

のテキストカテゴリーAであることが判明する。

BROWN1_A.TXT	2357	470 Monday night. A13 0480 ST& PAUL-MINNEAPOLIS,	MAY	1- ("AP")- {Billy Gardner's A13 0490 line double,
BROWN1_A.TXT	2381	in the last of the fifth. A13 0720 CODEN, UTAH,	MAY	1- ("AP")- Boston Red Sox Outfielder A13 0740 Jac
BROWN1_A.TXT	2434	1270 who reaches a decade in the big leagues come	May	25, joined six other A13 1280 sluggers who wallop
BROWN1_A.TXT	2740	.300.# The Pirates jumped off to an 11-3 start by	May	A15 0550 1 last year, when the Redbirds as well
BROWN1_A.TXT	2742	held them A15 0660 even over the season. On last	May	1, the Cardinals stood at 7-6, ending A15 0570 a
BROWN1_A.TXT	2744	estone. In 1958, the Birds were A15 0580 3-10 on	May	1. A year later they were 4-13. Since 1949, the
BROWN1_A.TXT	2745	he A15 0590 St& Louis club has been below .500 on	May	1 just four times. The A15 0600 '49 team was off
BROWN1_A.TXT	2747	beginning, then fell as low as 12-17 A15 0610 on	May	23 before finishing with 96 victories. The '52
BROWN1_A.TXT	2748	6 victories. The '52 Cards A15 0620 were 6-7 on	May	1 but ended with 88 triumphs, the club's top A15
BROWN1_A.TXT	2750	last season the Birds tumbled as low as 11-18 on	May	A15 0640 19 before recovering to make a race of
BROWN1_A.TXT	2755	wined a 21-2 mark and a nine-game A15 0680 lead by	May	8. The club that overcame the worst start in a co
BROWN1_A.TXT	2895	0 York Metropolitan Opera Company will be here in	May	, and the board A16 0240 will sponsor the Saturday
BROWN1_A.TXT	2896	e board A16 0240 will sponsor the Saturday night,	May	13, performance of "Turandot" A16 0250 as a benef
BROWN1_A.TXT	3096	ertain at a black tie dinner A17 0260 Wednesday,	May	3, in the Officers' Club at the Air Force A17 027
BROWN1_A.TXT	3099	ilio Bassi} at a A17 0290 cocktail party Tuesday,	May	2, from 6 to 8 p.m. in their home. The A17 0300 B
BROWN1_A.TXT	3123	first A17 0530 committee meeting will be held on	May	19. {Mr& Mrs& A17 0540 Andrew S& Kelsey} of
BROWN1_A.TXT	3142	ters A17 0710 there are to be written (and during	May	, which is National Salvation A17 0720 Army Week,
BROWN1_A.TXT	3266	members A18 0020 of the Thrift Shop Committee for	May	6 at Philmont Country Club. A18 0030 The women ha
BROWN1_A.TXT	3991	A 14-power conference on Laos should then meet on	May	5, A21 1440 he said. #PLEA FOR ARMS# The Laos gov
BROWN1_A.TXT	5997	31 1980 "The Marriage Altar" at 7:30 p.m. Sunday,	May	14. A31 1990 The resignation of the Rev& Warre
BROWN1_A.TXT	6519	deliberations A34 1310 that Premier Khrushchev on	May	4 made new inquiries through the A34 1320 U& S& E
BROWN1_A.TXT	6522	told Moscow he would give his answer A34 1340 by	May	20 after consultation with the Allies. The respon
BROWN1_A.TXT	6539	1500 and a 44-year-old President (his birthday is	May	29) with a A34 1510 demonstrated capacity for pol
BROWN1_A.TXT	7559	apped golf's confident Arnold Palmer (TIME cover,	May	A40 0600 2), who staged two cliffhanging rallies
BROWN1_A.TXT	8176	A43 0980 accompanied by a member of our staff, on	May	10 toured plants of A43 0990 two of Italy's bigge
BROWN1_B.TXT	1130	President Lyndon B& Johnson's B06 1810 visit last	May	. Gen& Taylor, the President's B06 1820 special
BROWN1_B.TXT	3047	hem to. The bill which passed the Assembly last	May	and B17 0540 is now pending in the Senate should
BROWN1_B.TXT	3091	led B17 0970 the parkway from East Orange to Cape	May	and I found the most B17 0980 courteous group of
BROWN1_B.TXT	3170	majority report of the Joint Economic Committee (	May	2) by B17 1770 Senators Proxmire and Butler alleg
BROWN1_B.TXT	4828	ng missions. He has been in B26 1580 prison since	May	, 1958, when his aircraft was shot down over Moluc
BROWN1_C.TXT	3064	0 Khrushchev at the Summit Conference in Paris in	May	1960, and that C16 1340 either China or Russia wa
BROWN1_E.TXT	267	re far more exciting in February E02 0780 than in	May	. Try that late winter pickup when you are so tire
BROWN1_E.TXT	2127	0-mile auto race in Charlotte, E12 1190 N& C&, on	May	27; the Florida Folk Festival, White Springs, E12
BROWN1_E.TXT	2128	he Florida Folk Festival, White Springs, E12 1200	May	5-7; Singing on the Mountain in Linville, North C
BROWN1_E.TXT	2178	re-enactment E12 1700 of the Battle of Lexington,	May	18th at Lexington, Missouri; E12 1710 the world-r
BROWN1_E.TXT	2180	ianapolis, Indiana, E12 1720 plus a festival from	May	27-30; "Song of Hiawatha", in E12 1730 Elgin, Ill
BROWN1_E.TXT	2182	ely E12 1740 Tulip Festival in Holland, Michigan,	May	12-14; the USGA E12 1750 Open in Birmingham from
BROWN1_E.TXT	2218	0170 it became capital of the Eastern section. On	May	11, 330 A&D&, E13 0180 its name was changed again
BROWN1_E.TXT	3093	is document began: "NO& E18 0110 1 NEWBURY PORT,	MAY	30TH, 1791 "Whereas, a Bridge over E18 0120 Me
BROWN1_E.TXT	3149	ished E18 0650 in the "Massachusetts Magazine" in	May	1793, which is reproduced E18 0660 herewith (Fig&
BROWN1_E.TXT	3198	st land battle of the Civil War at its home site,	May	E18 1150 28th to June 3rd. Colonel Frederick W&
BROWN1_E.TXT	3242	18 1580 while on the famous Jones-Imboden raid in	May	, 1863. General E18 1590 Jones was fresh from a lo
BROWN1_E.TXT	3262	8 1780 a week-long series of events, beginning on	May	28th and continuing E18 1790 through June 3rd, to
BROWN1_E.TXT	4602	t the prize E26 0740 was never awarded to him. In	May	, 1803, Ritter, in another flight E26 0750 of fanc

図4 Brown Corpus の 'May' のコンコーダンスライン

このコンコーダンスライン作成については、コンピュータでコンコーダンサーを使って自動的にコンコーダンスラインを出す作業であり、驚くほどの速さ、つまり数秒で出せる作業であってコンピュータコーパス言語学といわれる所以である。しかし、そのコンコーダンスラインの数は、key word の数だけラインができて、その総数は想像を絶するほどのラインになる。下記の表は、トライアルで出した 100 万語のアメリカ英語 2 つ、イギリス英語 2 つの「週 (Friday, Monday, etc) の出現数」であり、Brown Corpus だけで 457 個出現していて、総数で 1,592 のコンコーダンスラインを手動で検索しなければならない (トライアルには WWC は含まれていない)。つまり、「週 (Friday, Monday, etc) の出現数」だけで、1,592 コンコーダンスラインであるが、残りの「月 (month)」、「年 (year)」を加えると膨



大なコンコーダンスラインを一つ一つ処理していくことを考えると、その処理にどのくらい時間と労力が掛かるかは容易に想像できると思われる。

表3 4つのコーパスに出現する週の数

	Brown	Frown	LOB	FLOB
Friday(s)	64	65	31	45
Monday(s)	72	91	46	49
Saturday(s)	73	59	72	65
Sunday(s)	117	104	97	119
Thursday(s)	35	53	31	28
Tuesday(s)	59	30	25	29
Wednesday(s)	37	37	26	33
Total	457	439	328	368

## 2) コーディング (coding; 符号化)

作成したコンコーダンスラインを具体的に処理するのは、その一つ一つにコードを付与する手作業である。この手作業のコーディング (符号化) とは、それぞれのコンコーダンスラインを一行毎に読んでその一つ一つにさまざまな情報を付与する作業であり、とてつもない時間と労力を必要とする根気のいる作業である (一般的にコーパス言語学を誤解する大きな要因は、コンピュータで出したデータをそのまま鵜呑みに即座に利用できるという間違った見解であり、このようなデータの読みを手作業で行う部分を度外視して理解することには大きな誤解が生じやすい)。本論文に出てきたデータの処理には想像を絶するほどの時間と労力を要したことは論を待たない。そのコーディングの手順を詳述する。

1. コンコーダンサーを使って出したコンコーダンスラインをコピーして、EditPad Lite 7 というアプリケーションソフトに張り付ける。その際、出てきたデータで分析の目的ではないもの (Bycatch tokens) も全て検証対象として考察していく。そうすることによって、全てのデータの読み込みが可能になる。
2. 下記の表4のあるように、1 から 10 までの factor group を一つ一つのコンコーダンスラインにそれが持つ情報として手作業で付与していく。例えば、この key word のコンコーダンスラインには、1. Preposition がある P、 ない 0 といった符号をつけ、最後の 10. Corpus は Brown Corpus であれば B、Frown Corpus であれば R とそれぞれ一つづつ符号を付与していく作業である。この factor group は model によっては順不同になることもある。この論文において、この手作業が最も時間と労力を必要とし、また同時に集中して付与する作業は必要とされていることは、想像に難くないと思われる。

表4 符号化に使われる情報

Factor group (and sample factors)	
1.	<b>Preposition presence</b> (P=present; 0=absent)
2.	<b>Preposition identity</b> (e.g. i= <i>in</i> ; o= <i>on</i> ; f= <i>from</i> ; \$= <i>by</i> …)
3.	<b>Type of head noun</b> (e.g. W=weekday; M=month; Y=year, T=date…)
4.	<b>NP definiteness</b> (e.g. d=definite; i=indefinite; t=deictic determiner)
5.	<b>Position of adverbial in clause</b> (I=initial; M=medial; F=final)
6.	<b>Head premodification</b> (e.g. 0=none; J=adjective…)
7.	<b>Head postmodification</b> (e.g. 0=none; P=prepositional phrase; R=finite relative clause…)
8.	<b>Total length of adverbial NP</b> (number of words; >9 coded as “+”)
9.	<b>Text category</b> (A-R, with fiction categories later combined in recoding)
10.	<b>Corpus</b> (B=Brown; R=Frown; F=FLOB; L=LOB; W=WWC)

上記の符号化のための情報について説明すると、大きく 10 の項目に分かれている。

1. 前置詞の有無：例 有 P, 無 0
2. 前置詞の区別：例 i=*in*, o=*on*
3. 名詞の種類：例 W=weekday, M=month
4. 名詞句の限定：例 d=definite, i=indefinite
5. 文中の位置：例 I=initial, f=final
6. 前置修飾：例 J=adjective, 0=none
7. 後置修飾：例 0=none, P=prepositional phrase
8. 副詞類の語数：例 3=three words
9. テキストカテゴリー：例 A=Text category A
10. コーパスの区別：例 B=Brown, F=FLOB

下に別の表現で略式表記したものをあげ、かっこの中にいくつかの例を挙げてみる。

### 1. Preposition presence

0: absent (*Monday*)

P: present (on his birthday)

### 2. Preposition identity

i: in, o: *on*, x: other

### 3. Type of head-noun

W: day of week (*on Tuesday*)

M: month (*April, January*)

D: other specified day (*on Christmas Day, that day*)

P: part of day (*morning, midday, afternoon, evening, night, dawn*)

T: date/specific day without day (*3rd March 2014, Christmas*)

### 4. Definiteness

d: *the* + head (*the day*)

0: no determiner (*on Wednesday*)

i: indefinite (*a(n)*) (*on a beautiful day*)

t: deictic words + head (*last, next, this, that, some, every, / today, yesterday*) (ex. *last Sunday*)

j: indefinite without *a(n)* (*one day*)

x: other (#, p, u, o)

### 5. Position of adverbial

I: sentence initial (*On Monday he sang a song.*)

M: medial (*He sang a song on Thursday when he visited his uncle.*)

F: final (*He sang a song on Tuesday.*)

X: heading (not part of sentence. ex. *Monday, 21 March, 2015*)

### 6. Premodification

J: adjective (*on the beautiful day*)

N: noun (*on his son's birthday*)

0: none (*on Saturday*)

### 7. Postmodification

∅: none

P: prepositional phrase

### 8. Total length of the time adverbial noun phrase (number of words):

1: 1 word (=default code) ... (*on Monday*)

2: 2 words (*on his birthday*)

3: 3 words (*on his last day*)

9 words (*the day when I visited my grandfather in Chicago*)

+: 10 words or over ( ??? )

## 9. Text categories

A, B, C, D, E, F, G, H, J, K-R

## 10. Corpus

B: Brown

L: LOB

R: Frown

F: FLOB

W: WWC

それら factor group をそれぞれ符号化してサンプル・コーディングをしてみた例を下記に挙げてみる。

### <Some examples with codes>

- *on Good Friday*; (PoD0F002AB)
- *on the morning of Friday, Nov., 9, 1927*; (PoPdF0P7FB)

サンプルのコーディングを説明すると、例の ‘on Good Friday’ のコーディングは、「前置詞があり、文末にあり、副詞類の数は2語で Brown Corpus のテキストカテゴリーA」にあるという情報の付与ができたわけである。さらに ‘on the morning of Friday, Nov., 9, 1927’ は、「前置詞があり、1日のある時間、文末にあり、前置句で、語数7、Brown Corpus のテキストカテゴリーF」であるという情報付与である。

コーディング作業は手作業であるので、これら膨大なデータに1行毎に code sets 10 個を付けていくのには想像を絶するくらいの時間と労力を必要とする。実際に上に挙げた code sets の情報を左端に1行ずつ書き入れたサンプルを挙げてみる<sup>19</sup>。

(PoWiM002DB 654	disobeyed, namely, on a	Friday, as a parallel to the death of Christ on Good Fr	D04 1010 D
(PoD0F002DB 655	as a parallel to the death of Christ on Good	Friday; he sees a parallel also to the Jewish day of	D04 1020 D
(0oD0F002KB 693	doorbell was pitched too high. It was still Good	Friday, after all. A nigger boy opened the door. Payn	K05 0130 K
(PoPdF0P7FB 751	and it was still raining on the morning of	Friday, November 4, 1927, when officials of the Post	F05 0160 F
(PoW0F001PB 902	having a stag dinner over at the Pagan Room on	Friday. Imagine a stag dinner with Toodle Williams". He	P06 0160 P
(PoW0F001PB 906	Reno for stag dinners. "Thanks", Owen said, "but	Friday is a long way off and anything can happen".	B P06 0200 P
(PoW0F001CB 1546	ompany appeared at the Henry Street Playhouse on	Friday and Saturday evenings and Sunday afternoons in t	C09 0020 C
(PoW0M001EB 1591	the classiest things on the grounds. Rain on	Friday prevented many workouts, but there were a few mil	E09 1310 E
(0oW0F001AB 1651	will be presented to the Montgomery County Court	Friday, it was disclosed Tuesday by Horace A& Davenport,	A09 1010 A
(PoW0I001AB 1823	Young Republicans Thursday night in Salem. On	Friday he will go to Portland for the swearing in of De	A10 0780 A
(0oW0F001AB 1861	fs and a firmer grasp on their future were taken	Friday by delegates to the 29th general council of the	A10 1160 A
(0oP0FN02AB 1883	the right hand of God. #SUPER AGAIN ELECTED#	Friday afternoon the Rev& T& F& Zimmerman was reelected	A10 1380 A
(0oP0IN02AB 1887	Mo&. Election came on the nominating ballot.	Friday night the delegates heard the need for their fort	A10 1420 A
(0oW0F001AB 1896	the church or religious life", he told delegates	Friday. #CHURCH LOSES PACE# Talking of the rapid	A10 1510 A
(0oW0F001AB 2235	said, "and probably will have a good scrimmage	Friday. We'll work out about an hour on Saturday, then w	A12 1070 A
(P@IjI009PB 2266	Easter holidays, with her at one of those awful	Friday Evening Dancing Class parties her mother had ma	P14 0210 P
(PoW0I001AB 2579	and the local National Leaguers yesterday. On	Friday, Roger Maris, the Yankee outfielder and winner o	A14 0830 A
(0oP0FN02AB 2706	to open with Larry Jackson against the Cubs here	Friday night. Harvey Haddix, set back by the flu this	A15 0200 A

図5 コーディング (Friday) の例

<sup>19</sup> このコーディングの作業は、一つ一つの事例を吟味し長い日数と労力を掛けてコーディングする作業であるのでその作業効率にも大きく影響してくる事は容易に想像できる。

また、同じようなコーディングを 'key word: Friday' として行ったコーディング結果が各コンコードンスラインの左端に挙げてある。

■ brownday20150414Day of week coding - X76帳

ファイル( F) 編集( E) 書式( O) 表示( V) ヘルプ( H)

File	Line	Left	Key	Right	Reference	Category
(OoWFO01AB	70	arl Williams Hartsfield, in Fulton Superior Court	Friday	. His petition charged mental cruelty. The coup	A01 0695	A
(PoWFO01LB	127	ed on it in his <Sun-Times> column. We opened on	Friday	and closed the following Monday. Out of the enti	L01 1260	L
(PoWFO01AB	142	nday in the House to rescind the body's action of	Friday	in voting itself a \$10 per day increase in expen	A01 1410	A
(PoWFO01AB	165	ution, which Barber tossed into the House hopper	Friday	, will be formally read Monday. It says that "in	A01 1640	A
(PoWIM002DB	654	on the same day that they disobeyed, namely, on a	Friday	, as a parallel to the death of Christ on Good Fr	D04 1010	D
(PoDF0002DB	655	ay, as a parallel to the death of Christ on Good	Friday	; he sees a parallel also to the Jewish day of or	D04 1020	D
OoDF0002KB	693	doorbell was pitched too high. It was still Good	Friday	, after all. A nigger boy opened the door. Payn	K05 0130	K
(PoPFO07FB	751	l days and it was still raining on the morning of	Friday	, November 4, 1927, when officials of the Post Of	F05 0160	F
(PoWFO01PB	902	e having a stag dinner over at the Pagan Room on	Friday	. Imagine a stag dinner with Toodle Williams". He	P06 0160	P
(PoWFO01PB	906	Reno for stag dinners. "Thanks", Owen said, "but	Friday	is a long way off and anything can happen". B	P06 0200	P
(PoWFO01CB	1546	ompany appeared at the Henry Street Playhouse on	Friday	and Saturday evenings and Sunday afternoons in t	C09 0020	C
(PoWMO01EB	1591	the classiest things on the grounds. Rain on	Friday	prevented many workouts, but there were a few mil	E09 1310	E
(OoWFO01AB	1651	will be presented to the Montgomery County Court	Friday	, it was disclosed Tuesday by Horace A& Davenport,	A09 1010	A
(PoWJ001AB	1823	Young Republicans Thursday night in Salem. On	Friday	he will go to Portland for the swearing in of De	A10 0780	A
(OoWFO01AB	1861	fs and a firmer grasp on their future were taken	Friday	by delegates to the 29th general council of the	A10 1160	A
(OoP0FN02AB	1883	to the right hand of God. #SUPER AGAIN ELECTED#	Friday	afternoon the Rev& T& F& Zimmerman was reelecte	A10 1360	A
(OoPJ0N02AB	1887	Mo&. Election came on the nominating ballot.	Friday	night the delegates heard the need for their fort	A10 1420	A
(OoWFO01AB	1896	the church or religious life", he told delegates	Friday	. #CHURCH LOSES PACE# Talking of the rapid popula	A10 1510	A
(OoWFO01AB	2235	ek said, "and probably will have a good scrimmage	Friday	. We'll work out about an hour on Saturday, then w	A12 1070	A
P01J1009PB	2266	e Easter holidays, with her at one of those awful	Friday	Evening Dancing Class parties her mother had ma	P14 0210	P
(PoWJ001AB	2579	es and the local National Leaguers yesterday. On	Friday	, Roger Maris, the Yankee outfielder and winner o	A14 0830	A
(OoP0FN02AB	2706	to open with Larry Jackson against the Cubs here	Friday	night. Harvey Haddix, set back by the flu this	A15 0200	A
(PoWJF002LB	2721	ay for a few days, and not likely to return on a	Friday	. Had Muller made the same deduction? Muller was	L16 1340	L
(OoWJFJ02AB	2812	tie. The tie was against Southeast Missouri last	Friday	. Scherer also had a big night at bat with four h	A15 1250	A
(OoP0MN02AB	3101	or their new post. There will be a stag dinner	Friday	evening at the Denver Country Club which will pr	A17 0310	A
(OoWJLJ02JB	3466	ers a proportion "p) are licensed drivers. Each	Friday	one manager is chosen by lot to stay late and lo	J19 0510	J
(OoWJFJ02AB	3608	ring-in ceremony will be held in his office next	Friday	. #SIMPKINS FROM SOMERSET# Mr& Simpkins is a resid	A19 1460	A
OoWFO01AB	3934	family. #STUDENTS HELP OUT# Houghton said 6 plm&	Friday	had been set for a canvass of all 480 homes in t	A21 0860	A
(OoP0FN03AB	3947	d another dance for Fuhrmann's 770 students this	Friday	night. "Fuhrmann's faculty is proud that this	A21 0990	A
OoWFO01AB	4041	Emory University's Board of Trustees announced	Friday	that it was prepared to accept students of any r	A22 0010	A
(OoWMO01AB	4047	by the 33-man Emory board at its annual meeting	Friday	noted that state taxing requirements at present	A22 0070	A
(OoWMO01AB	4073	rom its beginning", the trustees' statement said	Friday	, "Emory University has assumed as its primary co	A22 0320	A
(OoWFO01AB	4100	man was killed and two others injured at midnight	Friday	when the car they were riding slid into a utili	A22 0560	A
(OoWFO01AB	4111	enewed its demands for movie theater integration	Friday	and threatened picketing and "stand-ins" if nego	A22 0690	A
(OoP0MN02AB	4119	wn and art theater managers and owners, contacte	Friday	night for comment on the "OOAHR request, said t	A22 0770	A
OoWFO003AB	4121	dge of such a letter, and that it was not in the	Friday	mail. However, three of the managers did say tha	A22 0790	A
(OoWFO01AB	4130	ago. Charles A& Black, "OOAHR chairman, said	Friday	that three theater representatives had agreed to	A22 0880	A
OoWFO01AB	4132	led to show up. He declined to name the three.	Friday	's letters asked for a Nov& 15 meeting. Failure t	A22 0900	A
(OoWMO01AB	4139	prepared statement released by the student group	Friday	stated that "extensive research by "OOAHR into t	A22 0970	A
(OoWFO01AB	4201	it lost at least for a while", she said wistfully	Friday	. One of the largest crowds in the club's hist	A22 1580	A
(OoWFO01AB	4238	Grady County. That death occurred at 6:50 plm&	Friday	and was reported Sunday, the patrol said. #BURST	A22 1960	A
(OoWFO01AB	4248	ed, the state public welfare commission was told	Friday	. It is the only county in the state so far thi	A23 0040	A
(OoWFO01AB	4311	of assault and robbery, Portland detectives said	Friday	. Mrs& Lavaughn Huntley is accused of driving th	A23 0660	A

図 6 コーディング (Friday) の例

### 3) 分析 (analysis)

コンコードンスラインを符号化 (coding) したら、統計処理でクロス集計 (crosstabulations) とロジスティクス回帰 (logistic regression) をするために、'GoldVarb 2001' という「言語変異分析」のための統計処理のソフトにそれらのデータを入れる作業がまっている。その「言語変異分析」のための 'GoldVarb 2001' について詳細に扱った論文がある (「VARBRUL プログラムとは何か」松田 謙次郎、2006。オンライン: <https://core.ac.uk/download/pdf/49291434.pdf>)。

この論文は言語変異理論において、最近になって統計処理を行うためにさまざまなソフトが使われるようになってきて、それらの総称として「VARBRUL プログラム」というと概説している。それらのソフトに関する本格的な教科書は今まで存在せず、付属の説明書、並びにワークショップなどで学ぶ以外になかったと説いているが、しかし Paolillo (2002) が本格的な教科書が出版して使い方がやや容易になって来た感があるような書き方である。

そのように VARBRUL の中の GoldVarb の使い方はかなりの専門的な知識を必要とし、理解して使いこなすには相当な時間が掛かる。そこで、本論文執筆のための統計処理に関して、この GoldVarb に精通し自らの研究にも積極的に役立てている統計の専門家の助言に大いに依存した事を明記したい<sup>20</sup>。

「VARBRUL プログラム」とは、Cedergren and Sankoff (1974)、またその前年の Cedergren (1973) において初めて使用された言語変異分析用フリーウェアの名称である。日本ではあまり知られているとは言い難いが、欧米の変異理論研究論文では頻繁に登場し、論文の議論もその出力結果に基づいてなされることが多い。しかしながらそうした状況にある欧米においても、長い間このプログラムに関する本格的な教科書は存在せず、ソフト付属のマニュアルや概説書、また変異理論最大の学会である NWAV (New Ways of Analyzing Variation) を始めとする学会で付随して開催されるワークショップなどでその使用法を学んでいたのが実情であった。Paolillo (2002) の出版によって、変異理論界は、ようやく最近統計的背景や歴史までも含めた、VARBRUL プログラムの全貌を纏めた教科書を持つことになったと言っても過言ではない。

実際に現在使用されているソフトは、VARBRUL (MS-DOS)、Goldvarb 2.1 (Macintosh)、GoldVarb2001 (Windows) および Goldvarb X (Windows および Macintosh) とあるが、変異理論関連の論文では、一概に VARBRUL analysis と言うことが多い。<sup>1</sup> 現在、VARBRUL プログラムは以下のウェブサイトからダウンロードが可能である。

- アメリカ・ペンシルバニア大学言語データコンソーシアム (VARBRUL) [ftp://ftp.ldc.upenn.edu/pub/ldc/misc\\_sw/varbrul.tar.Z](ftp://ftp.ldc.upenn.edu/pub/ldc/misc_sw/varbrul.tar.Z)
- カナダ・ケベック大学モントリオール校数学研究センター (Goldvarb 2.1) [http://www.crm.umontreal.ca/~sankoff/GoldVarb\\_Eng.html](http://www.crm.umontreal.ca/~sankoff/GoldVarb_Eng.html)
- イギリス・ヨーク大学言語学科 (Goldvarb2001) <http://www.york.ac.uk/depts/lang/webstuff/goldvarb/>
- カナダ・トロント大学タリアモンテ教授ウェブサイト (Goldvarb X) <http://individual.utoronto.ca/tagliamonte/goldvarb.htm>

図7 Varbrul プログラムについて (松田 謙次郎、2006)

---

<sup>20</sup> R. Sigley 氏は長年この統計処理研究している研究者であり、この論文の情報処理に甚大な協力をいただいたことに対してここに深く謝意を述べたい。

### 3.4 分析結果

今まで述べたように 5 つのコーパスを使って「時を表す前置詞句の副詞類における前置詞の省略」を分析するのであるが、そのためにはいくつかの分析手順があり、その手順を総称して、Data extraction (データ抽出)、coding (コーディング (符号化))、analysis (分析) と呼んでそれぞれの作業を行った。そしてその分析手順によって検索された分析結果を次の三つに分類して詳述していく。

1. Weekday-headed adverbials :  
(週を key word とする副詞類)
2. Month-headed adverbials :  
(月を key word とする副詞類)
3. Year-headed adverbials :  
(年を key word とする副詞類)

#### 3.4.1 分析結果①: 'weekday-headed' adverbials (週を key word とする副詞類)

まず weekday-headed adverbials の統計処理の結果は以下のようにになっている。それぞれの factor group に分けて説明をする。注目すべき数値は、一番右端の欄にある factor weight で 0.5 以上であれば「時を表す前置詞句の副詞類における前置詞の省略」は favour (起きやすい) であることを示している。

5. Model for Zero-marking of Weekday-headed Adverbials			
Factor group (with significance relative to this model)	Ø / total	(% Ø)	Factor weight (over 0.5 favours Ø)
<b>1) Definiteness (p&lt;0.001)</b>			
Ø: no determiner	519/1203	(43.1%)	0.429
i: indefinite	3/ 46	(6.5%)	0.060
d: <i>the</i> + head	17/ 55	(30.9%)	0.329
t: deictic words + head	178/ 196	(90.8%)	0.920
j: indefinite without <i>a(n)</i>	16/ 19	(84.2%)	0.899
x: other	3/ 8	(37.5%)	0.202
<b>2) Corpus (p&lt;0.001)</b>			
Brown	263/ 372	(70.7%)	0.758
Frown	214/ 340	(62.9%)	0.687
LOB	81/ 252	(32.1%)	0.290
FLOB	78/ 259	(30.1%)	0.286
WWC	100/ 304	(32.9%)	0.320
<b>3) Preposition identity (p&lt;0.001)</b>			
O: <i>on</i>	729/1417	(51.4%)	0.562
X: other	7/ 110	(6.4%)	0.038

## 5. Model for Zero-marking of Weekday-headed Adverbials (cont)

<b>4) Type of head-noun</b>				(p<0.001)
W: day of week	575/1132	(50.8%)		0.577
D: other specific day	3/24	(12.5%)		0.131
P: part of day	151/349	(43.3%)		0.302
T: other	7/22	(31.8%)		0.344
<b>5) Total length of adverbial</b>				(p<0.001)
1 word:	333/759	(43.9%)		0.350
2 words:	256/450	(56.9%)		0.626
3 words:	75/174	(43.1%)		0.629
4 words:	39/70	(55.7%)		0.751
5 words:	10/28	(35.7%)		0.632
6 words:	14/22	(63.6%)		0.840
7 words:	4/13	(30.8%)		0.585
8+ words:	5/11	(45.5%)		0.764
<b>6) Postmodification</b>				(p<0.001)
Ø: none	687/1392	(49.4%)		0.527
P: other	49/135	(36.3%)		0.250

## 5. Model for Zero-marking of Weekday-headed Adverbials (cont)

<b>7) Text category</b>				(p=0.008)
A	Reportage	413/731	(56.5%)	0.556
B	Editorials	42/94	(44.7%)	0.490
C	Reviews	43/95	(45.3%)	0.499
D	Religion	16/43	(37.2%)	0.282
E	Hobbies	30/66	(45.5%)	0.548
F	Popular Lore	26/69	(37.7%)	0.566
G	Biography & belles lettres	31/107	(29.0%)	0.332
H	Official documents	7/25	(28.0%)	0.214
J	Academic writing	12/26	(46.2%)	0.501
K-R	Fiction	116/271	(42.8%)	0.459
<b>8) Position of adverbial (non-significant)</b>				(p=0.146)
<b>9) Premodifier (non-significant)</b>				(p=0.271)
<b>Overall (input weight)</b>		736/1527	(48.2%)	0.471

図8 Model for Zero-marking of 'Weekday-headed' Adverbials

### 分析結果：

次に'Weekday-headed' Adverbials の分析結果について、factor group ごとに説明をしていく。

1. **Definiteness** : 結果から、「不定冠詞」では極端に起こらずに (例、*on a beautiful day*)、「直示的」(例、*the day*)または「a(n)のない不定冠詞 (one)」(例、*one day*) の場合は起こる場合が多い。それをグラフに表すと以下ようになる。



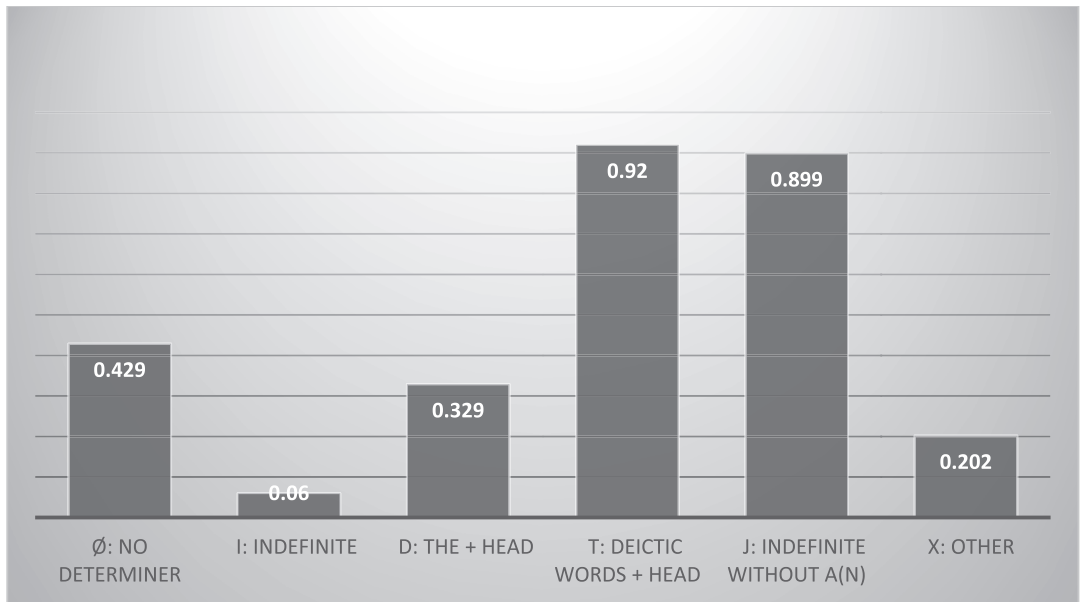


図9 Weekday-headed : Definiteness

2. **Corpus**: アメリカ英語 (Brown corpus 0.758, Frown corpus 0.687) はイギリス英語 (LOB corpus 0.290, FLOB corpus 0.286) やニュージーランド英語 (WWC 0.320) よりも省略が起こりやすい。

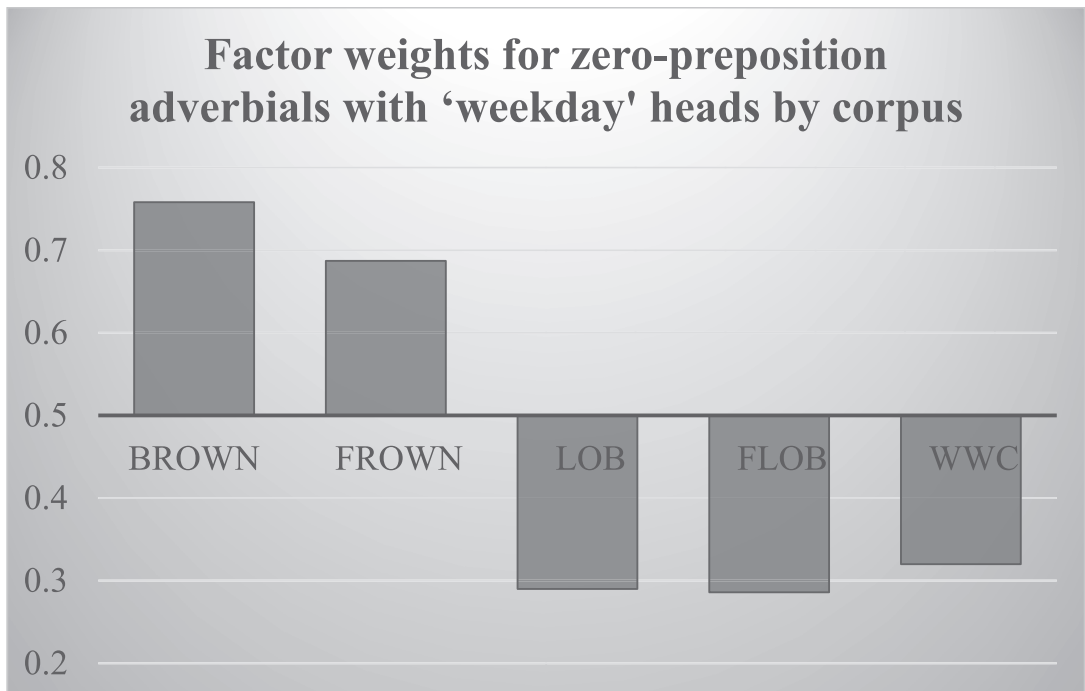


図10 Weekday-headed : Corpus

3. **Preposition identity** : 前置詞 on 以外の省略は起こらない (0.038 vs 0.562)。
4. **Type of head-noun** : weekday heads の方が省略が起こりやすい (0.577)。
5. **Total length of adverbial** : より長い副詞類の方が省略が起こりやすい。下にそのグラフを挙げる。

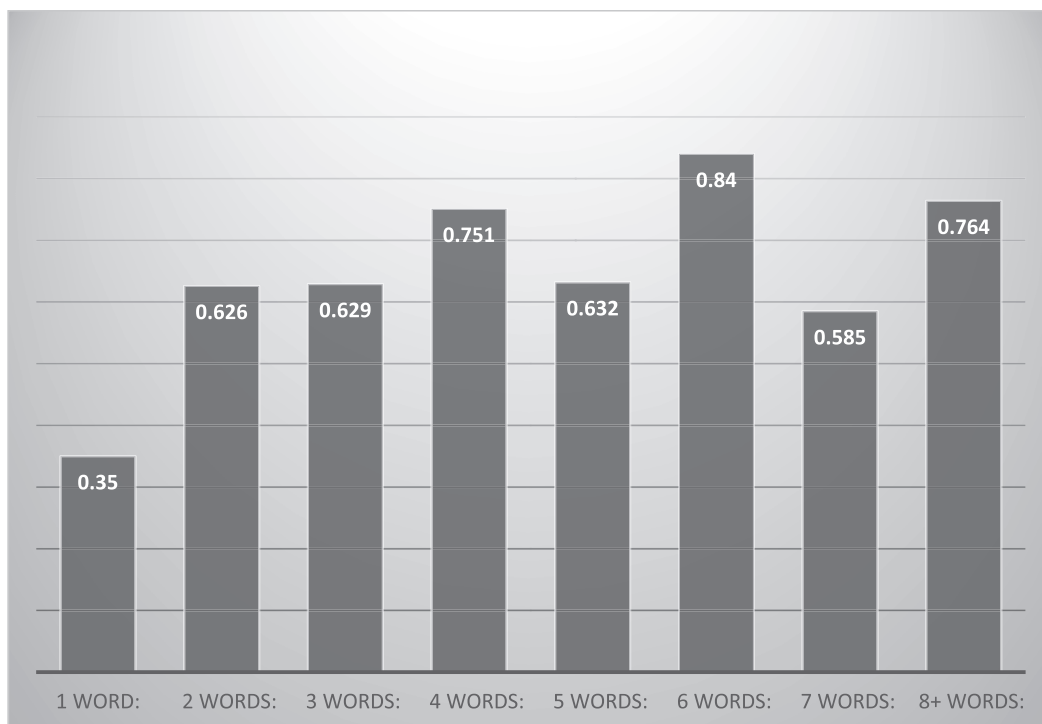


図 11 Weekday-headed: total length of adverbial

6. **Postmodification** : postmodified heads は省略が起こりにくい。
7. **Text category** : テキストカテゴリーにおいては、前置詞省略に大きな違いが見られる (図 12)。カテゴリーA (新聞記事)、カテゴリーE (技術娯楽)、カテゴリーF (一般実用) では省略が起こりやすいが、カテゴリーD (宗教)、カテゴリーG (書簡・伝記等)、カテゴリーH (雑 (政府文書等)) 等では省略は起こりにくいという結果がでた。この現象は、前に述べた Formality Index との関係において分析するとその関連性が見えてくる。しかし、最も Formality Index が less formal であった創作散文の K-R がやや省略の度合いが低い事の説明は今後の課題であろう。

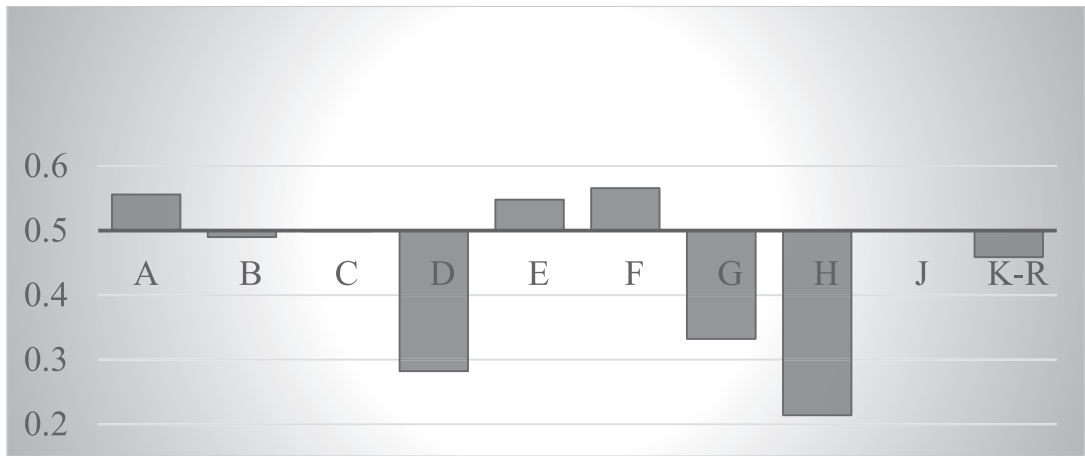


図 12 Weekday-headed: Text category

今度は、テキストカテゴリーにおける前置省略を frequency per 1,000 words で見てみる（図 13）とテキストカテゴリーの特徴が顕著に表れた結果を見る事ができる。テキストカテゴリーA（新聞記事）は、その文体、伝えたい目的が顕著に他のカテゴリーと一線を画していることと、重要な事は紙面の制限であり、伝えたい情報の時間的長さが短時間である事などが影響して、前置詞省略が顕著であるという事ができる。

また同時に、Formality Index の高いカテゴリーH（政府文書等）、カテゴリーJ（学術）が下の図で低いのは理解できるが、Formality Index の低いカテゴリーK-R（創作散文）が比較的低いのは、また別の分析が必要になってくるのだろう。

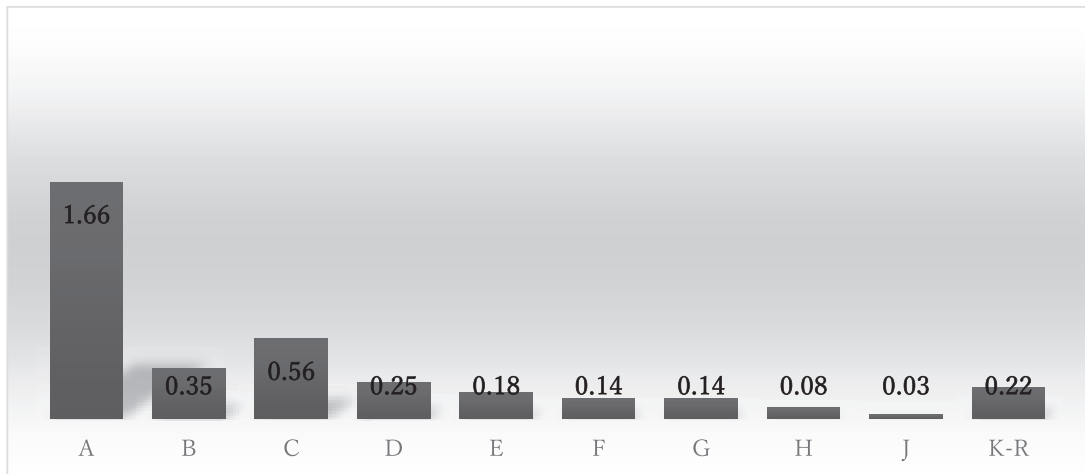


図 13 Weekday-headed: Text category (per 1,000 words)

8. **Position of adverbial** : これには大きな統計的な差異は見られなかった。

9. **Premodifier** : これも大きな統計的な差異は見られなかった。

ここで、‘weekday-headed’ dataset を総合的に整理してみるといくつかの結論を導き出すことができる。

‘weekday-headed’ dataset の結論：ここでは、‘weekday-headed’ dataset で出てきた結論を簡単にまとめてみる。

- ① 出された結果のうち、text category と corpus の関係に潜在的な違いを見ることが出来る。
- ② アメリカ英語とイギリス英語の違いは、text category によって大きな違いが出てくる。
- ③ 創作散文 (text category K-R) においては、大きな地域差は散見されない。
- ④ アメリカ英語の中では、Brown corpus の方が Frown corpus よりも省略の頻度が大きいという事が見られる。

### 3.4.2 分析結果②： ‘month-headed’ adverbials (月を key word とする副詞類)

次に ‘month-headed’ adverbials について出された結果をもとに、同様にそれぞれの factor group に分けて説明をする。それぞれの factor group に分けて説明するとき、出てきた悔過の順番はそれぞれ違うが同じような factor weight を注視しながら説明していく。この ‘month-headed’ adverbials も同じように注目すべき数値は、factor weight で 0.5 以上であれば「時を表す前置詞句の副詞類における前置詞の省略」は favour (起きやすい) であることは ‘weekday headed’ adverbials と同じである。

下に Model for Zero-marking of ‘Month-headed’ Adverbials の一覧表を挙げてその factor group ごとに説明していく。

Factor group (with significance relative to this model)	Ø / total	% Ø	Factor weight (over 0.5 favours Ø) (p<0.001)
<b>1) Preposition identity</b>			
o: <i>on</i>	455/1336	(34.1%)	0.856
i: <i>in</i>	181/1438	(12.6%)	0.405
X: other	14/658	(2.1%)	0.058
<b>2) Definiteness</b>			
Ø: no determiner	541/3277	(16.5%)	0.459
t: deictic words + head	108/115	(93.9%)	0.995
X: other	1/40	(2.5%)	0.157
<b>3) Position of adverbial</b>			
Final	444/2487	(17.9%)	0.507
Medial	33/193	(17.1%)	0.439
Initial	61/635	(9.6%)	0.272
Other	112/117	(95.7%)	0.994

## 6. Model for Zero-marking of Month-headed Adverbials (cont)

4) Text category				(p<0.001)
<b>A</b>	Reportage	148/724	(20.4%)	0.472
<b>B</b>	Editorials	154/324	(47.5%)	0.781
<b>C</b>	Reviews	38/104	(36.5%)	0.733
<b>D</b>	Religion	6/ 54	(11.1%)	0.401
<b>E</b>	Hobbies	47/311	(15.1%)	0.488
<b>F</b>	Popular Lore	40/341	(11.7%)	0.437
<b>G</b>	Biography & belles lettres	68/631	(10.8%)	0.338
<b>H</b>	Official documents	57/408	(14.0%)	0.444
<b>J</b>	Academic writing	62/384	(16.1%)	0.566
<b>K-R</b>	Fiction	30/151	(19.9%)	0.661
5) Corpus				(p<0.001)
<b>Brown</b>		188/698	(26.9%)	0.568
<b>Frown</b>		173/680	(25.4%)	0.650
<b>LOB</b>		84/635	(13.2%)	0.430
<b>FLOB</b>		85/623	(13.6%)	0.457
<b>WVC</b>		120/796	(15.1%)	0.399
6) Premodifier				(p=0.038)
<b>Ø: none</b>		531/3178	(16.7%)	0.487
<b>J: other</b>		119/254	(46.9%)	0.660

## 6. Model for Zero-marking of Month-headed Adverbials (cont)

7) Type of head noun				(p=0.061)
<b>T: date</b>		461/632	(28.2%)	----
<b>M: month</b>		189/1800	(10.5%)	----
8) Total length of adverbial				(p=0.209)
<b>1 word</b>		32/658	(4.9%)	----
<b>2 words</b>		411/1823	(22.5%)	----
<b>3 words:</b>		184/774	(23.8%)	----
<b>4 words</b>		13/ 81	(16.0%)	----
<b>5 words:</b>		9/ 73	(12.3%)	----
<b>6+ words:</b>		1/ 23	(4.3%)	----
9) Postmodification				(p=0.833)
<b>Ø: none</b>		606/2557	(23.7%)	----
<b>P: other</b>		44/875	(5.0%)	----
<b>Overall (input weight)</b>		650/3432	(18.9%)	0.085

図 14 Model for Zero-marking of 'Month-headed' Adverbials

分析結果：

1. Preposition identity：前置詞 in が好まれるのは当然の事であろう。

2. **Definiteness** : 直示的 (deictic words + head) 表現で起こるのは (0.995)、は‘weeday-headed’ adverbials と同様である。
3. **Position of adverbial** : 文頭で省略が起こるのは極端に少ないことが分かる。しかし、文中でも起こるが統計的には有意ではない。その点、文末で前置詞が省略されるのが多く有意の範囲を超えている。(注意したいのは、‘weekday-headed’では統計的に大きな差異はなかったが、この ‘month-headed’では、文頭、文中、文末で差異が見られた)。

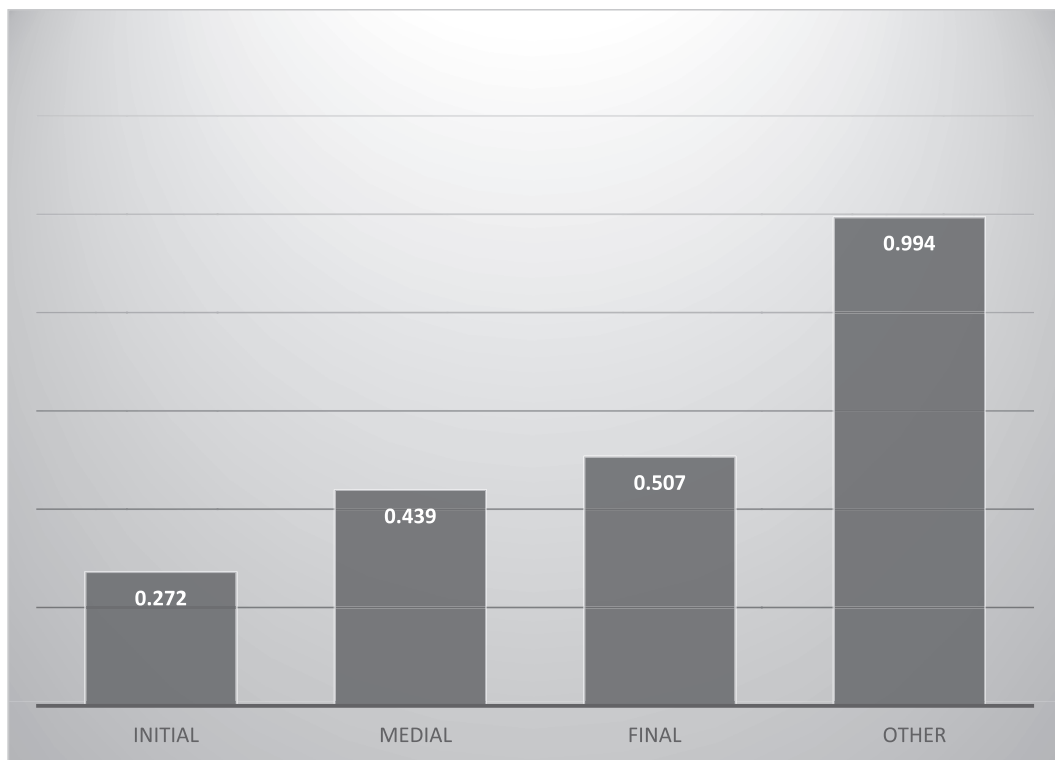


図 15 Month-headed: Position of adverbial

4. **Text category** : ここでも他と同様に text category にはさまざまな差異が見られる。カテゴリー B (新聞論説)、カテゴリー C (新聞書評等)、カテゴリー J (学術)、カテゴリー K-R (創作散文) では省略が favour で省略が起こるが、カテゴリー D (宗教) とカテゴリー G (書簡・伝記等) 等では少ないことが見て取れる。

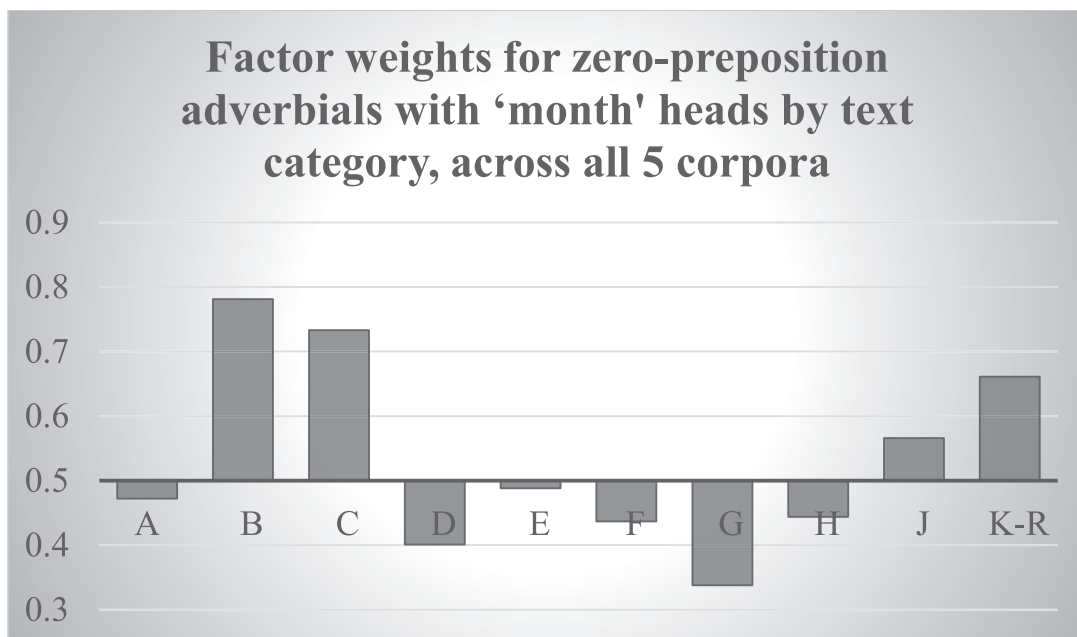


図 16 Month-headed: Text category

頻度によるテキストカテゴリー毎の frequency per 1,000 words を見てみると、weekday-headed と同様にカテゴリーA が省略が多いことが見て取れる。しかし、カテゴリーH が weekday-headed と違って、省略が高いのは興味ある現象である。

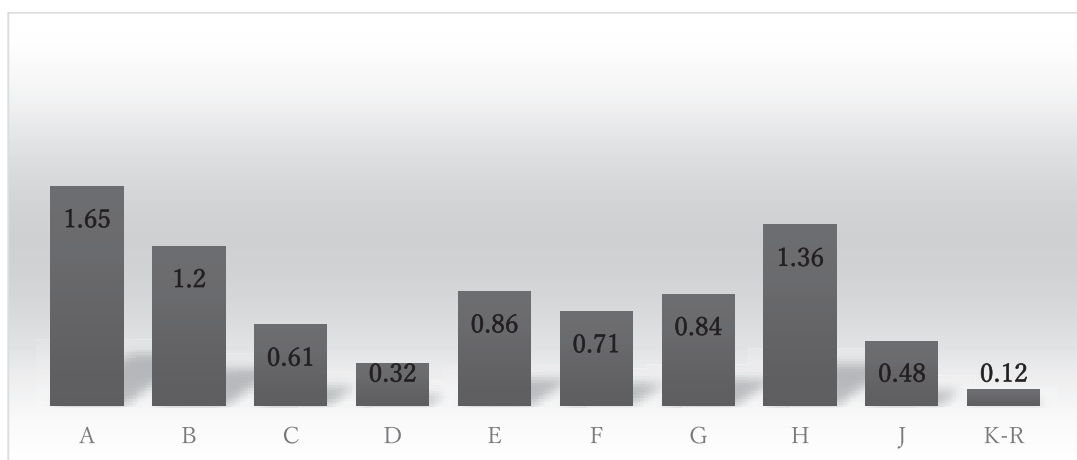


図 17 Month-headed: Text category (per 1,000 words)

5. **Corpus**: アメリカ英語 (Brown corpus, Brown corpus) では、イギリス英語 (LOB corpus, FLOB corpus) やニュージーランド英語 (WWC) と比較して省略が起こることが証明されている。それは、'weekday-headed adverbials' と同じ傾向であるが、実際の数値には隔たりがある。特に 1992 年の新しいアメリカ英語では、その傾向が顕著である。

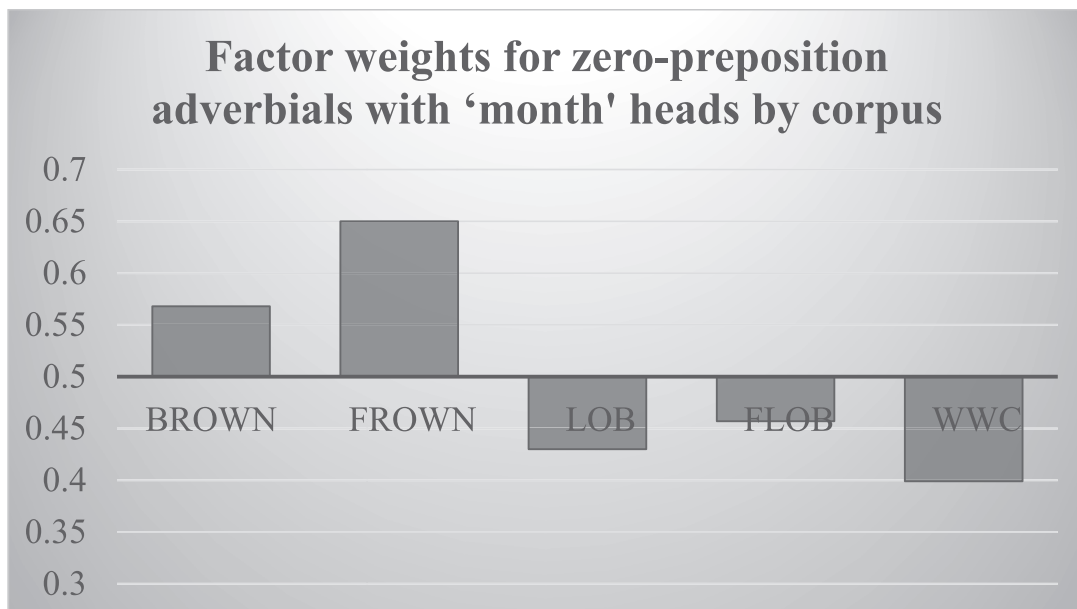


図 18 Month-headed: Corpus

6. **Premodifier** : 大きな差は生じていない。
7. **Type of head-noun** : date heads の方が month heads よりやや大きいことが分かる。
8. **Total length of adverbial** : 2 語、3 語の month-headed adverbials が比較的省略が起こりやすいことが分かるが、大きな差にはなっていない。
9. **Postmodification** : 大きな影響はみられない。

### 3.4.3 分析結果③ : 'year-headed' adverbials (年を key word とする副詞類)

次に 'year-headed' adverbials について出てきた結果をもとに、同様にそれぞれの factor group に分けて説明をする。この 'year-headed' adverbials も同様に注目すべき数値は、factor weight で 0.5 以上あれば「時を表す前置詞句の副詞類における前置詞の省略」は favour (起きやすい) であることは 'weekday headed' adverbials や 'month-headed' adverbials と同じである。

下に Model for Zero-marking of 'Year-headed' Adverbials の一覧表を挙げてその factor group ごとに説明していく。



## 7. Model for Zero-marking of Year-headed Adverbials

Factor group (with significance relative to this model)	Ø / total	% Ø	Factor weight (over 0.5 favours Ø)
<b>1) Preposition identity</b>			( <i>p</i> <0.001)
i: <i>in</i>	70/ 3998	1.8%	0.699
X: other	4/ 1569	0.3%	0.105
<b>2) Total length of adverbial</b>			( <i>p</i> <0.001)
1 word:	55/ 4707	1.2%	0.419
2 words:	2/ 76	2.6%	0.786
3 words or longer:	17/ 784	2.2%	0.862
<b>3) Corpus</b>			( <i>p</i> <0.001)
Brown	19/ 923	2.1%	0.665
Frown	6/ 1078	0.6%	0.337
LOB	17/ 984	1.7%	0.615
FLOB	13/ 1264	1.0%	0.411
WWC	19/ 1318	1.4%	0.517

## 7. Model for Zero-marking of Year-headed Adverbials (cont)

<b>4) Text category</b>			( <i>p</i> <0.001)
A (Reportage)	3/ 644	0.5%	0.263
B (Editorials)	2/ 269	0.7%	0.367
C (Reviews)	3/ 128	2.3%	0.707
D (Religion)	3/ 210	1.4%	0.592
E (Hobbies)	10/ 464	2.2%	0.672
F (Popular Lore)	8/ 797	1.0%	0.471
G (Biography & Essays)	17/ 1200	1.4%	0.568
H (Official)	5/ 636	0.8%	0.390
J (Academic writing)	17/ 1125	1.5%	0.533
K-R (Fiction)	6/ 94	6.4%	0.878
<b>5) Type of head-noun</b>			( <i>p</i> =0.001)
Year	73/ 5408	1.3%	0.519
Season	1/ 159	0.6%	0.073
<b>6) Position of adverbial in clause</b>			( <i>p</i> =0.018)
Initial	6/ 1193	0.5%	0.303
Medial	5/ 270	1.9%	0.608
Final	63/ 4104	1.5%	0.553
<b>7) Premodifier (non-significant)</b>			( <i>p</i> =0.695)
Overall (input weight)	74/ 5567	1.3%	0.006

図 19 Model for Zero-marking of 'Year-headed' Adverbials

### 分析結果：

全体的に前置詞省略は起こりえないという統計結果が出ている (input wait=0.006) が、他の factor でさまざまな結果が出ている。前に述べた 'weekday-headed' adverbials の input wait は 0.471、'month-headed' adverbials の input wait は 0.085 に比べて低い事が分かる。

1. **Preposition identity** : 当然で推測可能であるが、他の前置詞より *in* (0.699) で favour である。

2. **Total length of adverbial** : 長い adverbial NPs においては、前置詞両略が favour であり (0.862)、短い一語の場合は、省略は favour ではない (0.419)。
3. **Corpus** : この 'year-headed' adverbials' では、興味ある現象が見られ、アメリカ英語、イギリス英語とも 1961 年の英語は両方ともやや favour である (Brown; 0.665, LOB; 0.615) が、1991, 1992 年の新しい英語では、前置詞省略は favour ではない (Frown; 0.337, FLOB; 0.411) という現象はとても面白い結果であると言える。また、ニュージーランド英語はその中間でやや省略がおこるといったところだろう (WWC; 0.517)。

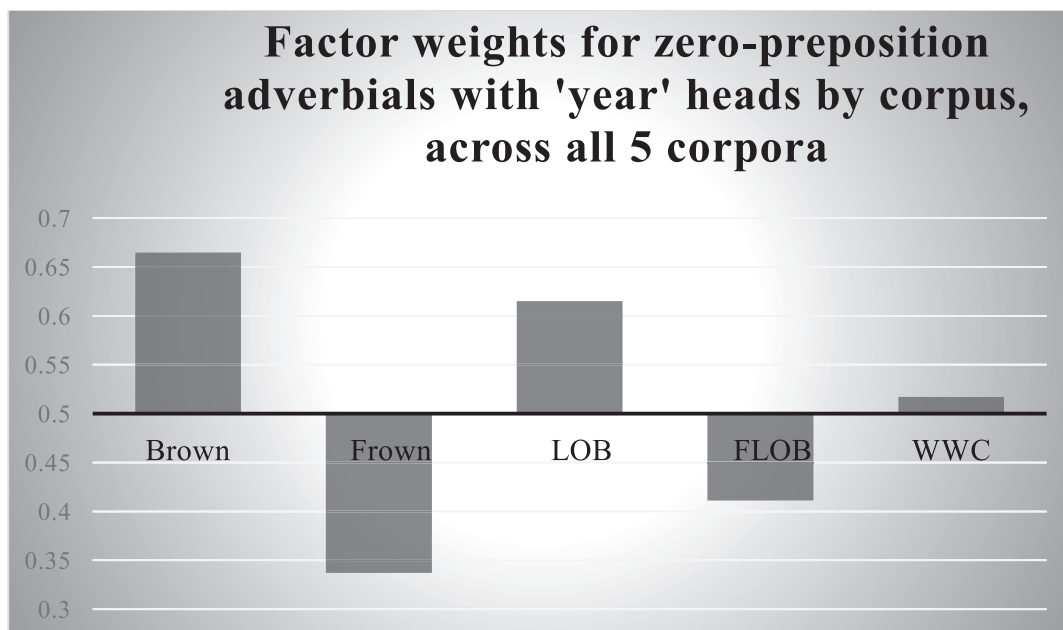


図 20 Year-headed: Corpus

4. **Text category** : いままで分析してきた他の Model (weekday-head, month-head) と同様この 'year-headed' adverbials' の text category にもさまざまな差異が見られる。カテゴリ-K-R (創作散文) で省略が最も起こり (0.878)、カテゴリ-C (新聞書評等; 0.707)、カテゴリ-E (技術娯楽; 0.672) 等が favour である。一方、カテゴリ-A (新聞記事; 0.263) とカテゴリ-B (新聞論説; 0.367)、カテゴリ-H (雑 (政府文書等)) 等では favour ではなかった。

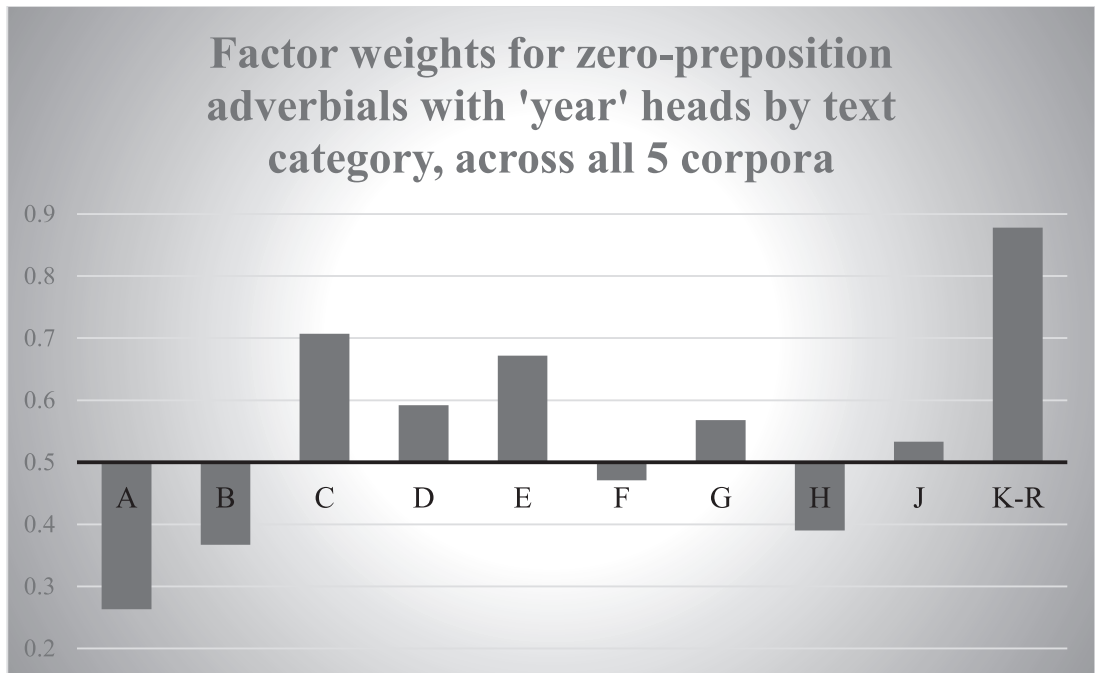


図 21 Year-headed: Text Category

頻度によるテキストカテゴリー毎の frequency per 1,000 words を見てみると、前の 2 つの Model と大きな差異が見られる。Formality Index の高いカテゴリーH、カテゴリーJ に高い値が見られ、カテゴリーF、カテゴリーG も高い値である。そして、カテゴリーA がそれに続くという現象が表出しているのが見て取れる。

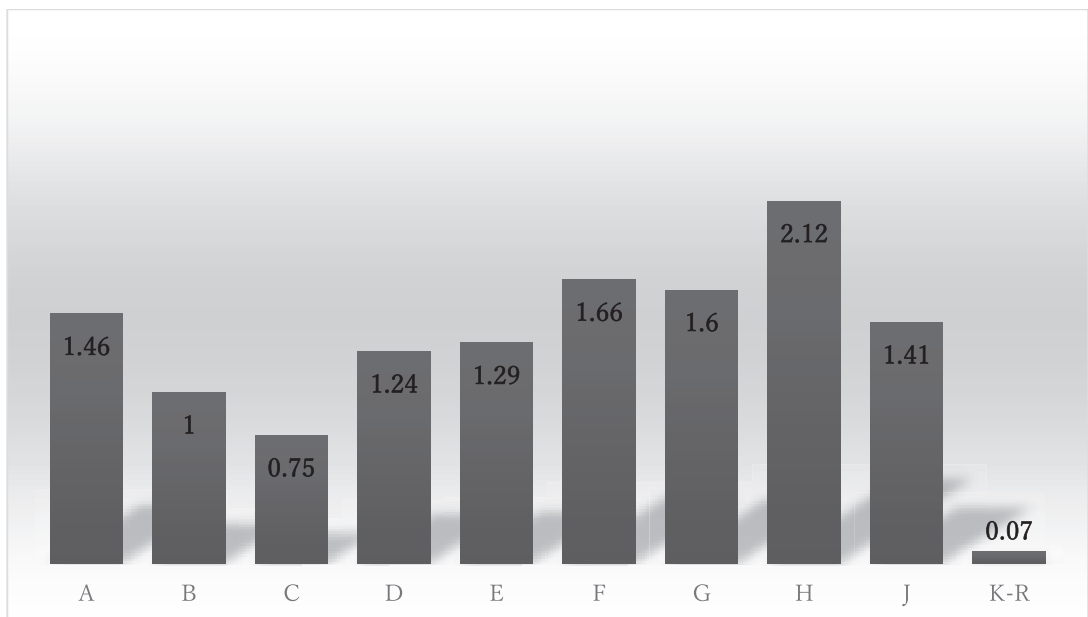


図 22 Year-headed: Text Category (per 1,000 words)

5. **Type of head-noun** : この factor においては、year の省略が favour であり (0.519)、season (0.073) は極端に少ない。
6. **Position of adverbial in clause** : 文の位置において、文中 (0.608) が最も省略が起こり、文末 (0.553) がそれに続いている。しかし、文頭では省略が favour ではない (0.303) 事が分かる。

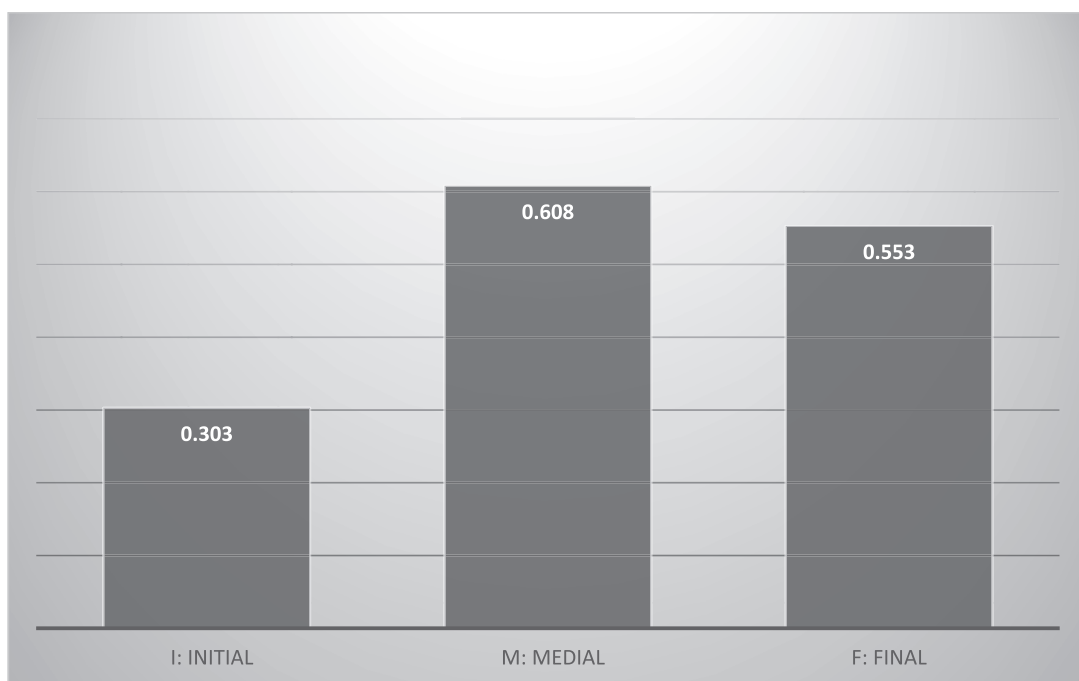


図 23 Year-headed: Position of Adverbial in Clause

7. **Premodifier** : 統計的には、significant ではない (0.006)。

## 第4章 結論

### 4.1 副詞類における前置詞省略のまとめと今後の課題

これまで GoldVarb を使って導かれた5つの Brown Family Corpora における「時を表す前置詞句の副詞類における前置詞に省略」の分析結果を仮説に従って総合的に概観してみると、次のような結論に到着するであろう。

1. **地域差によるヴァリエーション**：アメリカ英語は全般的に、「時を表す前置詞句の副詞類における前置詞に省略」が起りやすい傾向にある。例外は year-headed の前置詞省略で起りにくい点である。よって 1961 年 Brown Corpus は省略が起りやすい。しかし、イギリス英語では、省略が起りにくく、唯一 1961 年 LOB では year-headed で省略が起りやすい。ニュージーランド英語はほぼどれも省略が起らず、year-headed だけがややその傾向とは違う結果を見せている。
2. **テキストカテゴリー (genre) によるヴァリエーション**：Formality Index において Formality が低いと言われるカテゴリー K-R (創作散文) では省略が起こる (特に year-headed) が、逆に weekday-headed では省略が起こる傾向にないという結果が示された。しかし、Formality が高いと言われるカテゴリー J (学術) では、期待した省略 disfavour が表れていないのは特筆すべきことであろう。同じように Formality の高いカテゴリー H (雑 (政府文書等)) は省略が起こらないという典型的な結果である。またカテゴリー D (宗教) は概ね省略は起こらない (year-headed は例外)。新聞記事では、カテゴリー C (新聞書評等) では省略は起りにくいが、カテゴリー A (新聞記事) とカテゴリー B (新聞論説) は省略に複雑な結果が出ている。
3. **副詞類の文中の位置におけるヴァリエーション**：大きな言語学的 significance はないが、特に weekday-headed ではその傾向が強い。しかし、month-headed では文末がその傾向が強いのに対して、year-headed は文中の省略が起こる傾向にあることが分かる。

言語学的傾向として、これら5つの Brown Family Corpora を分析した結果、「時を表す前置詞句の副詞類における前置詞の省略」は、どの英語で、どのカテゴリーで、どの前置詞を省略するかは簡潔な表現で結論を述べるとすると、「期待される前置詞または頻度の高い前置詞」が省略されるという事は論を待たない傾向であると言える。

前述のように「前置詞が省略された時を表す前置詞句」は 概して informal な表現で省略が起こると次のように述べている。

On the whole, the sentence without the preposition tends to be more informal and more usual

(Quirk et al. (1985: 693).

さらに、Leech and Svartvik (1994: 84) は次のように述べて、

The preposition is also sometimes omitted directly before days of the week: I'll see you (on) Wednesday, then ... This omission is especially common in <informal, AmE>.

つまり informal なアメリカ英語で省略が起こると明言している。

しかし、この論文で地域差、テキストカテゴリー、文中の位置など Biber (1995) が使った register などという言語学的に具体的な制限を掛けていくとさらに細かい結果が出てくる事が証明された。またアメリカ英語はイギリス英語、ニュージーランド英語よりも「前置詞が省略された時を表す前置詞句」が顕著であることが証明された。

今後の課題として、カテゴリーA (新聞記事) において、その文体の特徴、つまり stylistic features を研究するに際し、機能語、特に前置詞等が省略されることは推測されるが、それをさらに詳細に検索し分析するためにより多くの corpus を分析して結論を出す必要があると感じられる。また大きな課題である formal と informal の文章の違いを明らかにするために更なる研究が必要であると感じる。

## References

- Aarts, J. and W. Meijs. (eds.). 1984. *Corpus Linguistics*. Amsterdam: Rodopi.
- Aijmer, K. and B. Altenberg. (eds.). 1991. *English Corpus Linguistics: Studies in Honour of Jan Svartvik*. London: Longman.
- Algeo, J. 1988. British and American grammatical differences. *International Journal of Lexicography* 1: 1-31.
- Altenberg, B. 1991. A bibliography of publications relating to English computer corpora. In Johansson, S. and A. Stenström. (eds.). *English Computer Corpora*. Berlin: Mouton de Gruyter.
- Bailey, N. 1730. *Dictionarium Britannicum*. London: T. Cox.
- Bailey, N. 1755. *A New Universal Etymological English Dictionary*. Revised by J. N. Scott. London: Osborne and Shipton.
- Bauer, L. 1993. *Manual of Information to Accompany The Wellington Corpus of Written New Zealand English*. Wellington: Victoria University of Wellington.
- Biber, D. 1995. *Dimensions of Register Variation: A Cross Linguistic Comparison*. Cambridge: Cambridge University Press.
- Biber, D., S. Conrad, and R. Reppen. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Biber, D. and R. Reppen. 2015. *The Cambridge Handbook of English Corpus Linguistics*. Cambridge: Cambridge University Press.
- Biber, D., S. Johansson, G. Leech, S. Conrad, and E. Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Cawdrey, R. 1604. *A Table Alphabeticall*. London: Edward Weaver.
- Coote, E. 1596. *The English Schoole-Master*. London.
- Crystal, D. 1987. *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge University Press. 風間喜代三/長谷川欣佑監訳『言語学百科事典』大修館書店 1992.
- Crystal, D. 1995. *The Cambridge Encyclopedia of the English Language*. Cambridge: Cambridge University Press.
- DeMaria, R. Jr. 1993. *The Life of Samuel Johnson*. Oxford and Cambridge: Blackwell.
- Deverson, T. and G. Kennedy. 2005. *The New Zealand Oxford Dictionary*. Australia: Oxford University Press.
- Francis, W. N. and H. Kučera. 1964. *Manual of Information to Accompany A Standard Corpus of Present-Day Edited American English, for use with Digital Computers*. Rhode Island: Brown University Press.
- Francis, W. N. 1992. Language corpora B. C. In Svartvik, J. (ed.). *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82 Stockholm, 4-8 August 1991*. Berlin:

- Mouton de Gruyter.
- Fries, C. C. 1940. *American English Grammar*. New York: Appleton-Century.
- Halliday, M. A. K. 1991. Corpus Studies and Probabilistic Grammar. In Aijmer, K. and B. Altenberg. (eds.). *English Corpus Linguistics: Studies in Honour of Jan Svartvik*. London: Longman. 30-43.
- Hawkins, Sir J. 1961. *The Life of Samuel Johnson LLD*. Edited, Abridged, and with an Introduction by Bertram H. Davis. New York: The Macmillan Co.
- Hoffmann, S. and R. Sigley. 2013. Approaching a linguistic variable: That-omission in mandative sentences. In Yamazaki, S. and R. Sigley. (eds.). *Approaching Language Variation through Corpora*. Bern: Peter Lang. 115-154.
- 石川慎一郎 2012. 『ベーシックコーパス言語学』ひつじ書房.
- 井上永幸、赤野一郎 (編) 2003. 『ウイズダム英和辞典』三省堂.
- Jespersen, O. 1914. *A Modern English Grammar on Historical Principles*. Part II Syntax, First Volume. Heidelberg.
- Jespersen, O. 1938. *En Sprogmands Levned*. Copenhagen: Gyldendal.
- Johansson, S. 1978. *Manual Information to Accompany The Lancaster-Oslo/Bergen Corpus of British English, for use with Digital Computers*. Oslo: University of Oslo.
- Johansson, S. 1991. Times change, and so do corpora. In Aijmer, K. and B. Altenberg. (eds.). *English Corpus Linguistics: Studies in Honour of Jan Svartvik*. London: Longman. 305-314.
- Johnson, S. 1755. *A Dictionary of the English Language*. London: Longman.
- 城生伯太郎 1990. 『言語学は科学である』情報センター出版局.
- 『ジーニアス英和大辞典』2001. 大修館.
- Kennedy, G. 1998. *An Introduction to Corpus Linguistics*. Harlow: Longman.
- Kennedy, G. and S. Yamazaki. 2000. The Influence of Maori on the New Zealand English Lexicon. In J. Kirk. (ed.). *Corpora Galore: Analyses and Techniques in Describing English*. Amsterdam: Rodopi. 33-44.
- Kennedy, G. 2003. *Structure and Meaning in English*. London: Longman.
- Kruisinga, E. 1925. *A Handbook of Present-Day English*. 4<sup>th</sup> ed. Utrecht: Kemink.
- Leech, G. 1992. Corpora and Theories of Linguistic Performance. In J. Svartvik. (ed.). *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82 Stockholm, 4-8 August 1991*. Berlin: Mouton de Gruyter. 105-22.
- Leech, G. and J. Svartvik. 1975. *A Communicative Grammar of English*. London: Longman.
- Leech, G. and J. Svartvik. 1994. *A Communicative Grammar of English*. 2<sup>nd</sup> ed. London: Longman.
- Leech, G., M. Hundt, C. Mair, and N. Smith. 2009. *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.
- Longman Dictionary of Contemporary English* (3<sup>rd</sup> Edition). 1995. London: Longman



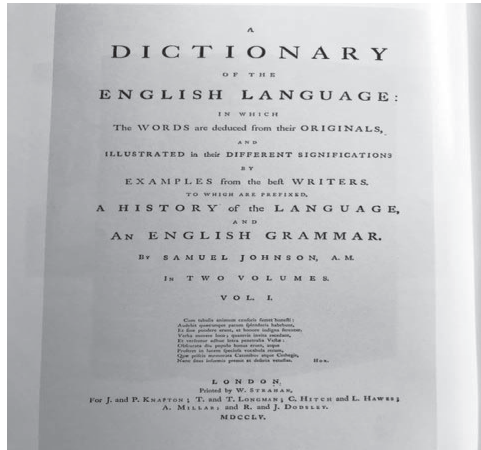
- 松田謙次郎 2006. 「VAEBRUL プログラムとは何か」 *Theoretical and applied linguistics at Kobe Shoin* : トークス Vol.6. 31-56. <https://core.ac.uk/download/pdf/49291434.pdf>.
- McEnery, T. and A. Hardie. 2012. *Corpus Linguistics*. Cambridge: Cambridge University Press.
- Murray, J. A. H. 1900. *The Evolution of English Lexicography*. Oxford: The Clarendon Press.
- Murray, J. A. H. 1928. *The Oxford English Dictionary: A New English Dictionary on Historical Principles*. Oxford: Oxford University Press.
- Murray, K. M. E. 1977. *Caught in the Web of Words: James Murray and the Oxford English Dictionary*. New Haven and London: Yale University Press.
- 永嶋大典 1983. 『ジョンソンの『英語辞典』』大修館.
- O’Keeffe, A. and M. McCarthy. (eds.). 2010. *The Routledge Handbook of Corpus Linguistics*. Oxford: Routledge.
- Oxford Advanced Learner’s Dictionary*. 1995. Oxford: Oxford University Press.
- Palmer, H. E. 1933. *Second Interim Report on English Collocations*. Tokyo: Institute for Research in English Teaching.
- Paolillo, J. C. 2002. *Analyzing Linguistic Variation: Statistical Models and Methods*. Stanford, CA: CSLI Publications.
- Phillips, E. 1706. *The New World of Words*. Revised by John Kersey. London: Phillips.
- Poutsma, H. 1928. *A Grammar of Late Modern English*. 2ed. Groningen: P. Noordoff.
- Quirk, R. and S. Greenbaum. 1973. *A University Grammar of English*. London: Longman.
- Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik. 1972. *A Grammar of Contemporary English*. London: Longman.
- Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Saito, T., J. Nakamura, and S. Yamazaki. (eds.). 2002. *English Corpus Linguistics in Japan*. Amsterdam: Rodopi.
- 齋藤俊雄、中村純作、赤野一郎（編）2005. 『英語コーパス言語学』（改訂新版）研究社.
- 齋藤俊雄、朝尾幸次郎、山崎俊次、新井洋一、梅咲敦子、塚本聡（共訳）2003. 『コーパス言語学：英語構造と用法の研究』南雲堂.
- Biber, D., S. Conrad, and R. Reppen. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Sigley, R. 1997. Text categories and where you can stick them: a crude formality index. *International Journal of Corpus Linguistics* 2. 199-237.
- Sigley, R. 2013. Assessing corpus comparability using a formality index: The case of the Brown/LOB clones. In Yamazaki, S. and R. Sigley. (eds.). *Approaching Language Variation through Corpora*. Bern: Peter Lang. 65-113.
- Sinclair, J. (ed.) 1987. *Collins COBUILD Advanced Learner’s English Dictionary*. Glasgow:

- HarperCollins.
- 『新英和大辞典』2002. 研究社.
- Sonoda, K. 2002. Omission of prepositions in time adverbials in present-day spoken AmE. *Bulletin of Nagasaki University School of Health Sciences*. 15(2): 19-25.
- Svartvik, J. (ed.). 1992. *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82 Stockholm, 4-8 August 1991*. Berlin: Mouton de Gruyter.
- Svartvik, J. 1992. Corpus Linguistics comes of age. In Svartvik, J. (ed.). *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82 Stockholm, 4-8 August 1991*. Berlin: Mouton de Gruyter. 7-13.
- Tagliamonte, S. A. 2006. *Analysing Sociolinguistic Variation*. New York: Cambridge University Press.
- Thorndike, E. L. 1921. *Teacher's Words Book*. New York: Columbia Teachers College.
- Thorndike, E. L. and I. Lorge. 1944. *A Teacher's Word Book of 30,000 words*. New York: Columbia Teachers College.
- Tognini-Bonelli, E. 2001. *Corpus Linguistics at Work*. Amsterdam: John Benjamins.
- Webster, N. 1934. *Webster's New International Dictionary of the English Language*. 2<sup>nd</sup> ed. Mass.: G. and C. Merriam Company.
- Yamazaki, S. 1996. The Use of Corpora before the Computer Age. 『大東文化大学語学教育研究論叢』. 第 13 号.
- 山崎俊次 2019. 「ニュージーランド英語におけるマオリ語の影響」、堀正広、赤野一郎(編) 『コーパスと英語研究』(英語コーパス研究シリーズ全 7 巻の第 1 巻) ひつじ書房. 146-63.
- Yamazaki, S. and R. Sigley. (eds.) 2013. *Approaching Language Variation through Corpora*. Bern: Peter Lang.
- 安井稔(編) 1996. 『コンサイス英文法辞典』三省堂.
- Zachrisson, R. E. 1913. *Pronunciation of English Vowels 1400 – 1700*. Goeteborg. Reprint of 1913 Edition. New York: AMS Press (1971).

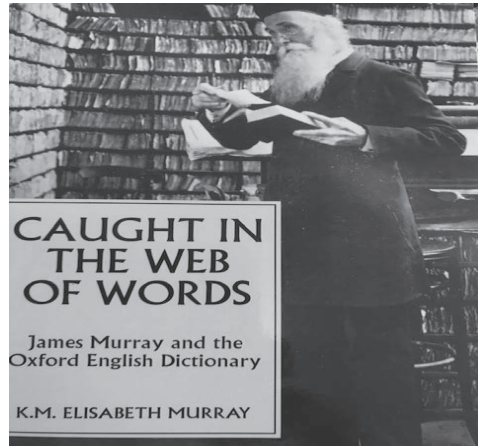
## Appendices

### (1) Books and Dictionaries

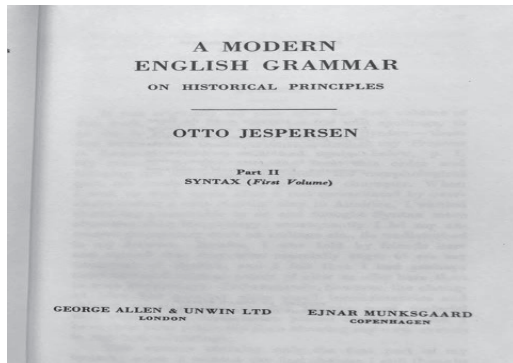
#### 1. Johnson's Dictionary (1755)



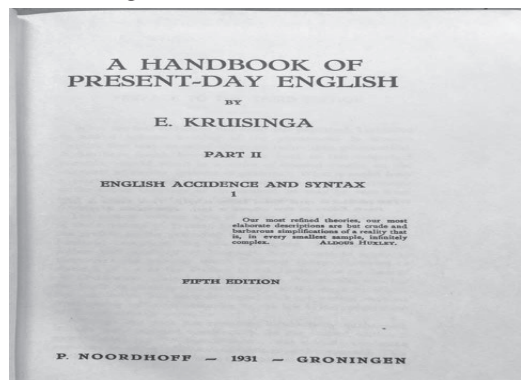
#### 2. James Murray (with slips of papers)



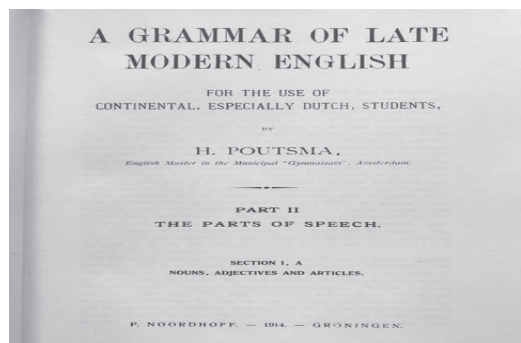
#### 3. Otto Jespersen's Grammar



#### 4. Kruisinga's Grammar



#### 5. Poutsma's Grammar





### (3) Pictures of Some Corpus Linguists

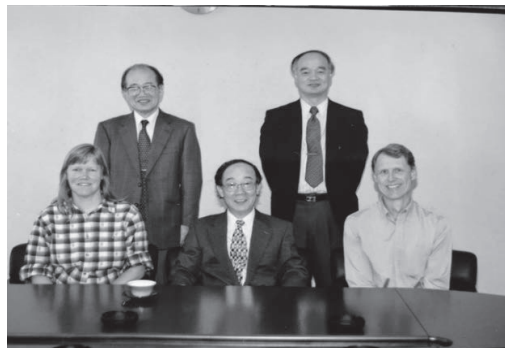
1. S. Greenbaum at SEU



2. G. Leech and G. Kennedy (右側の二人)



3. D. Biber and R. Reppen



4. G. Leech and T. Saito (右側の二人)



5. J. Svartvik



5. S. Johansson (右から 2 人目)



#### (4) [Annotated] Results of a first-pass step-up/step-down regression run on the "month" dataset.

• CELL CREATION • 2015/08/25 13:04:59  
.....

[The GoldVarb report starts with the name of the "token file" (coded dataset) used, and the "condition file" used to recode the tokens into a "cell file" kept in memory for modelling purposes. This run was performed on tokens retrieved from Brown, LOB, Frown, and FLOB; WWC was not included.]

Name of token file: month.tkn  
Name of condition file: month1.cnd

```
(  
; This is the set of recoding conditions used for a model of preposition  
omission in adverbials headed by "month" expressions in the Brown, LOB,  
Frown and FLOB corpora.  
; Text after a semicolon is not read by GoldVarb, and is used for  
explanatory comments.  
; For this run, the original data coding was slightly simplified after  
an initial inspection of frequencies to remove "knockout" environments  
in which no variation is observed. Such environments cannot be modelled  
separately in a logistic regression model, so they must either be  
excluded (if systematic), or combined with other low-frequency  
environments (if they could be accidental rather than systematic).  
; Column 1 is the dependent variable (whether the preposition is omitted  
or present).  
(1)  
; Column 2 is factor group #1 (type of preposition): prepositions other  
than on/in are mostly knockout environments in which omission is either  
not observed or very rare. These cannot be modelled separately, so are  
combined in this model as "Elsewhere".  
(2 (i (COL 2 i))  
 (o (COL 2 o))  
 (X (ELSEWHERE)))  
; Column 3 is factor group #2 (head of temporal noun phrase). The two  
lowest-frequency environments D=specified day (n=7 tokens) and W=weekday  
heads (n=43 tokens) were combined to avoid a knockout.  
(3 (W (COL 3 D)))  
; Column 4 is factor group #3 (definiteness).  
(4)  
; Column 5 is factor group #4 (position of adverbial within the sentence).  
(5)  
; Column 6 is factor group #5 (presence of a premodifier).  
(6)  
; Column 7 is factor group #6 (presence of a postmodifier).  
(7)  
; Column 8 is factor group #7 (length of adverbial phrase). Longer  
adverbials (over 6 words) were combined to avoid knockouts.  
(8 (+ (OR (COL 8 6) (COL 8 7) (COL 8 8) (COL 8 9)))  
; Column 9 is factor group #8 (the source text category). The subgenres  
of Fiction used in Brown and LOB are not distinguished in WWC, so these  
were all combined as one factor in all runs.  
(9 (K (OR (COL 9 L) (COL 9 M) (COL 9 N) (COL 9 P) (COL 9 R))))  
; Column 10 is factor group #9 (the source corpus).  
(10)  
)
```

[GoldVarb then reports the raw token frequencies for each factor group in turn:]

Number of cells: 741  
 Application value(s): 0P  
 Total no. of factors: 35

Group		0	P	Total	%
-----					
1 (2)		0	P		
X	N	13	477	490	18.4
	%	2.7	97.3		
o	N	379	720	1099	41.4
	%	34.5	65.5		
i	N	132	936	1068	40.2
	%	12.4	87.6		
Total	N	524	2133	2657	
	%	19.7	80.3		
-----					
2 (3)		0	P		
T	N	376	887	1263	47.5
	%	29.8	70.2		
W	N	9	41	50	1.9
	%	18.0	82.0		
M	N	139	1205	1344	50.6
	%	10.3	89.7		
Total	N	524	2133	2657	
	%	19.7	80.3		
-----					
3 (4)		0	P		
0	N	435	2129	2564	96.5
	%	17.0	83.0		
t	N	89	4	93	3.5
	%	95.7	4.3		
Total	N	524	2133	2657	
	%	19.7	80.3		
-----					
4 (5)		0	P		
F	N	424	1568	1992	75.0
	%	21.3	78.7		
M	N	24	114	138	5.2
	%	17.4	82.6		
I	N	76	451	527	19.8
	%	14.4	85.6		
Total	N	524	2133	2657	
	%	19.7	80.3		
-----					

5 (6)	0	P			
0	N	426	2035	2461	92.6
	%	17.3	82.7		
J	N	98	98	196	7.4
	%	50.0	50.0		
Total	N	524	2133	2657	
	%	19.7	80.3		
-----					
6 (7)	0	P			
0	N	521	2086	2607	98.1
	%	20.0	80.0		
P	N	3	47	50	1.9
	%	6.0	94.0		
Total	N	524	2133	2657	
	%	19.7	80.3		
-----					
7 (8)	0	P			
2	N	340	1044	1384	52.1
	%	24.6	75.4		
4	N	15	47	62	2.3
	%	24.2	75.8		
3	N	143	491	634	23.9
	%	22.6	77.4		
1	N	16	489	505	19.0
	%	3.2	96.8		
+	N	1	13	14	0.5
	%	7.1	92.9		
5	N	9	49	58	2.2
	%	15.5	84.5		
Total	N	524	2133	2657	
	%	19.7	80.3		
-----					



8 (9)		0	P		
A	N	131	406	537	20.2
	%	24.4	75.6		
B	N	108	131	239	9.0
	%	45.2	54.8		
C	N	35	56	91	3.4
	%	38.5	61.5		
D	N	2	30	32	1.2
	%	6.2	93.8		
E	N	41	187	228	8.6
	%	18.0	82.0		
F	N	34	226	260	9.8
	%	13.1	86.9		
G	N	56	458	514	19.3
	%	10.9	89.1		
H	N	52	293	345	13.0
	%	15.1	84.9		
J	N	38	246	284	10.7
	%	13.4	86.6		
K	N	27	100	127	4.8
	%	21.3	78.7		
Total	N	524	2133	2657	
	%	19.7	80.3		
-----					
9 (10)		0	P		
B	N	186	506	692	26.0
	%	26.9	73.1		
R	N	171	504	675	25.4
	%	25.3	74.7		
L	N	83	572	655	24.7
	%	12.7	87.3		
F	N	84	551	635	23.9
	%	13.2	86.8		
Total	N	524	2133	2657	
	%	19.7	80.3		
-----					
TOTAL	N	524	2133	2657	
	%	19.7	80.3		

Name of new cell file: month1.cel

[After GoldVarb has created a “cell file”, it can be used for a step-up/step-down regression run, which tries to identify a “best-fit” model of factor effects, containing only factor groups with statistically significant contributions to the model.]

• BINOMIAL VARBRUL • 2015/08/25 13:05:52  
.....

Name of cell file: month1.cel

Averaging by weighting factors.  
Threshold, step-up/down: 0.050001

Stepping up...

----- Level # 0 -----

[GoldVarb starts with a model assuming no factor effects at all, as a baseline.]

Run # 1, 1 cells:

Convergence at Iteration 2

Input 0.197

Log likelihood = -1319.246

----- Level # 1 -----

[On each level of the “step-up”, GoldVarb evaluates the contribution of each remaining factor group in turn to the model likelihood, and adds the factor group with the most significant further contribution (the largest increase in log-likelihood), until there are no further significant contributions.]

Run # 2, 3 cells:

Convergence at Iteration 6

Input 0.152

Group # 1 -- X: 0.132, o: 0.746, i: 0.440

Log likelihood = -1167.453 Significance = 0.000

Run # 3, 3 cells:

Convergence at Iteration 5

Input 0.178

Group # 2 -- T: 0.661, W: 0.503, M: 0.348

Log likelihood = -1239.551 Significance = 0.000

Run # 4, 2 cells:

Convergence at Iteration 6

Input 0.194

Group # 3 -- O: 0.459, t: 0.989

Log likelihood = -1183.994 Significance = 0.000

Run # 5, 3 cells:

Convergence at Iteration 4

Input 0.196

Group # 4 -- F: 0.527, M: 0.464, I: 0.410

Log likelihood = -1312.447 Significance = 0.001

Run # 6, 2 cells:

Convergence at Iteration 5

Input 0.190

Group # 5 -- O: 0.471, J: 0.810

Log likelihood = -1269.808 Significance = 0.000

Run # 7, 2 cells:

Convergence at Iteration 5

Input 0.196

Group # 6 -- O: 0.506, P: 0.208

Log likelihood = -1315.343 Significance = 0.007

Run # 8, 6 cells:  
Convergence at Iteration 6  
Input 0.167  
Group # 7 -- 2: 0.619, 4: 0.614, 3: 0.592, 1: 0.140, +: 0.277, 5: 0.478  
Log likelihood = -1243.983 Significance = 0.000

Run # 9, 10 cells:  
Convergence at Iteration 5  
Input 0.182  
Group # 8 -- A: 0.592, B: 0.787, D: 0.231, E: 0.496, F: 0.404, G: 0.355,  
H: 0.444, J: 0.410, K: 0.548, C: 0.737  
Log likelihood = -1239.995 Significance = 0.000

Run # 10, 4 cells:  
Convergence at Iteration 4  
Input 0.188  
Group # 9 -- B: 0.613, R: 0.593, L: 0.385, F: 0.397  
Log likelihood = -1281.868 Significance = 0.000

Add Group # 1 with factors Xoi

----- Level # 2 -----

Run # 11, 8 cells:  
No Convergence at Iteration 20  
Input 0.151  
Group # 1 -- X: 0.129, o: 0.693, i: 0.510  
Group # 2 -- T: 0.579, W: 0.365, M: 0.431  
Log likelihood = -1162.563 Significance = 0.009

Run # 12, 6 cells:  
Convergence at Iteration 8  
Input 0.126  
Group # 1 -- X: 0.156, o: 0.816, i: 0.318  
Group # 3 -- 0: 0.448, t: 0.997  
Log likelihood = -968.439 Significance = 0.000

Run # 13, 9 cells:  
Convergence at Iteration 6  
Input 0.150  
Group # 1 -- X: 0.129, o: 0.748, i: 0.439  
Group # 4 -- F: 0.531, M: 0.498, I: 0.384  
Log likelihood = -1157.739 Significance = 0.000

Run # 14, 6 cells:  
Convergence at Iteration 10  
Input 0.122  
Group # 1 -- X: 0.119, o: 0.821, i: 0.344  
Group # 5 -- 0: 0.446, J: 0.937  
Log likelihood = -1053.411 Significance = 0.000

Run # 15, 6 cells:  
Convergence at Iteration 6  
Input 0.151  
Group # 1 -- X: 0.131, o: 0.744, i: 0.442  
Group # 6 -- 0: 0.505, P: 0.263  
Log likelihood = -1165.483 Significance = 0.048

Run # 16, 18 cells:  
Convergence at Iteration 7  
Input 0.135  
Group # 1 -- X: 0.116, o: 0.724, i: 0.486  
Group # 7 -- 2: 0.605, 4: 0.621, 3: 0.501, 1: 0.204, +: 0.443, 5: 0.773  
Log likelihood = -1128.986 Significance = 0.000

Run # 17, 30 cells:  
Convergence at Iteration 6  
Input 0.136  
Group # 1 -- X: 0.117, o: 0.758, i: 0.438  
Group # 8 -- A: 0.589, B: 0.792, D: 0.238, E: 0.514, F: 0.393, G: 0.313,  
H: 0.437, J: 0.427, K: 0.640, C: 0.803  
Log likelihood = -1081.216 Significance = 0.000

Run # 18, 12 cells:  
Convergence at Iteration 6  
Input 0.146  
Group # 1 -- X: 0.135, o: 0.744, i: 0.438  
Group # 9 -- B: 0.594, R: 0.604, L: 0.389, F: 0.402  
Log likelihood = -1137.143 Significance = 0.000

Add Group # 3 with factors 0t

----- Level # 3 -----

Run # 19, 12 cells:  
No Convergence at Iteration 20  
Input 0.122  
Group # 1 -- X: 0.146, o: 0.728, i: 0.450  
Group # 2 -- T: 0.645, W: 0.434, M: 0.366  
Group # 3 -- 0: 0.447, t: 0.997  
Log likelihood = -961.033 Significance = 0.001

Run # 20, 16 cells:  
Convergence at Iteration 8  
Input 0.123  
Group # 1 -- X: 0.153, o: 0.819, i: 0.318  
Group # 3 -- 0: 0.447, t: 0.997  
Group # 4 -- F: 0.541, M: 0.430, I: 0.366  
Log likelihood = -956.529 Significance = 0.000

Run # 21, 11 cells:  
Convergence at Iteration 9  
Input 0.125  
Group # 1 -- X: 0.155, o: 0.821, i: 0.312  
Group # 3 -- 0: 0.453, t: 0.995  
Group # 5 -- 0: 0.488, J: 0.644  
Log likelihood = -966.973 Significance = 0.090

Run # 22, 10 cells:  
Convergence at Iteration 8  
Input 0.125  
Group # 1 -- X: 0.156, o: 0.815, i: 0.320  
Group # 3 -- 0: 0.448, t: 0.997  
Group # 6 -- 0: 0.505, P: 0.281  
Log likelihood = -967.323 Significance = 0.144

Run # 23, 26 cells:  
Convergence at Iteration 9  
Input 0.120  
Group # 1 -- X: 0.134, o: 0.813, i: 0.342  
Group # 3 -- 0: 0.450, t: 0.996  
Group # 7 -- 2: 0.549, 4: 0.609, 3: 0.463, 1: 0.364, +: 0.438, 5: 0.796  
Log likelihood = -957.738 Significance = 0.001

Run # 24, 44 cells:  
Convergence at Iteration 9  
Input 0.110  
Group # 1 -- X: 0.135, o: 0.833, i: 0.310  
Group # 3 -- O: 0.446, t: 0.997  
Group # 8 -- A: 0.529, B: 0.806, D: 0.286, E: 0.500, F: 0.400, G: 0.306,  
H: 0.432, J: 0.487, K: 0.702, C: 0.843  
Log likelihood = -889.788 Significance = 0.000

Run # 25, 21 cells:  
Convergence at Iteration 9  
Input 0.119  
Group # 1 -- X: 0.162, o: 0.817, i: 0.312  
Group # 3 -- O: 0.447, t: 0.997  
Group # 9 -- B: 0.586, R: 0.632, L: 0.366, F: 0.405  
Log likelihood = -937.609 Significance = 0.000

Add Group # 8 with factors ABDEFGHJKC

----- Level # 4 -----

Run # 26, 72 cells:  
No Convergence at Iteration 20  
Input 0.107  
Group # 1 -- X: 0.124, o: 0.758, i: 0.431  
Group # 2 -- T: 0.634, W: 0.438, M: 0.376  
Group # 3 -- O: 0.446, t: 0.998  
Group # 8 -- A: 0.532, B: 0.805, D: 0.290, E: 0.504, F: 0.408, G: 0.306,  
H: 0.424, J: 0.484, K: 0.705, C: 0.837  
Log likelihood = -884.127 Significance = 0.006

Run # 27, 112 cells:  
Convergence at Iteration 9  
Input 0.109  
Group # 1 -- X: 0.134, o: 0.835, i: 0.308  
Group # 3 -- O: 0.446, t: 0.998  
Group # 4 -- F: 0.535, M: 0.390, I: 0.397  
Group # 8 -- A: 0.525, B: 0.806, D: 0.289, E: 0.488, F: 0.407, G: 0.316,  
H: 0.426, J: 0.486, K: 0.709, C: 0.833  
Log likelihood = -882.476 Significance = 0.001

Run # 28, 65 cells:  
Convergence at Iteration 9  
Input 0.109  
Group # 1 -- X: 0.134, o: 0.838, i: 0.303  
Group # 3 -- O: 0.452, t: 0.995  
Group # 5 -- O: 0.487, J: 0.659  
Group # 8 -- A: 0.529, B: 0.808, D: 0.290, E: 0.501, F: 0.400, G: 0.304,  
H: 0.434, J: 0.490, K: 0.695, C: 0.841  
Log likelihood = -888.022 Significance = 0.064

Run # 29, 62 cells:  
Convergence at Iteration 9  
Input 0.109  
Group # 1 -- X: 0.133, o: 0.832, i: 0.312  
Group # 3 -- O: 0.446, t: 0.997  
Group # 6 -- O: 0.506, P: 0.229  
Group # 8 -- A: 0.527, B: 0.809, D: 0.284, E: 0.500, F: 0.401, G: 0.307,  
H: 0.431, J: 0.484, K: 0.702, C: 0.845  
Log likelihood = -888.078 Significance = 0.069

Run # 30, 147 cells:  
Convergence at Iteration 9  
Input 0.104  
Group # 1 -- X: 0.105, o: 0.809, i: 0.377  
Group # 3 -- O: 0.450, t: 0.996  
Group # 7 -- 2: 0.540, 4: 0.642, 3: 0.564, 1: 0.272, +: 0.435, 5: 0.797  
Group # 8 -- A: 0.547, B: 0.827, D: 0.275, E: 0.490, F: 0.400, G: 0.298,  
H: 0.419, J: 0.469, K: 0.715, C: 0.824  
Log likelihood = -878.329 Significance = 0.000

Run # 31, 149 cells:  
Convergence at Iteration 10  
Input 0.104  
Group # 1 -- X: 0.138, o: 0.835, i: 0.304  
Group # 3 -- O: 0.445, t: 0.998  
Group # 8 -- A: 0.520, B: 0.800, D: 0.285, E: 0.522, F: 0.423, G: 0.305,  
H: 0.427, J: 0.493, K: 0.697, C: 0.828  
Group # 9 -- B: 0.576, R: 0.623, L: 0.370, F: 0.420  
Log likelihood = -866.433 Significance = 0.000

Add Group # 9 with factors BRLF

----- Level # 5 -----

Run # 32, 207 cells:  
No Convergence at Iteration 20  
Input 0.101  
Group # 1 -- X: 0.128, o: 0.762, i: 0.422  
Group # 2 -- T: 0.631, W: 0.488, M: 0.377  
Group # 3 -- O: 0.445, t: 0.998  
Group # 8 -- A: 0.523, B: 0.800, D: 0.286, E: 0.524, F: 0.428, G: 0.305,  
H: 0.420, J: 0.492, K: 0.702, C: 0.821  
Group # 9 -- B: 0.573, R: 0.623, L: 0.372, F: 0.423  
Log likelihood = -862.115 Significance = 0.014

Run # 33, 314 cells:  
Convergence at Iteration 10  
Input 0.102  
Group # 1 -- X: 0.135, o: 0.838, i: 0.302  
Group # 3 -- O: 0.444, t: 0.998  
Group # 4 -- F: 0.543, M: 0.392, I: 0.368  
Group # 8 -- A: 0.513, B: 0.800, D: 0.281, E: 0.509, F: 0.439, G: 0.316,  
H: 0.418, J: 0.494, K: 0.705, C: 0.811  
Group # 9 -- B: 0.593, R: 0.627, L: 0.367, F: 0.402  
Log likelihood = -855.720 Significance = 0.000

Run # 34, 196 cells:  
Convergence at Iteration 8  
Input 0.104  
Group # 1 -- X: 0.137, o: 0.839, i: 0.298  
Group # 3 -- O: 0.450, t: 0.996  
Group # 5 -- O: 0.489, J: 0.636  
Group # 8 -- A: 0.519, B: 0.802, D: 0.288, E: 0.524, F: 0.422, G: 0.303,  
H: 0.429, J: 0.496, K: 0.693, C: 0.827  
Group # 9 -- B: 0.576, R: 0.621, L: 0.370, F: 0.423  
Log likelihood = -865.186 Significance = 0.120

Run # 35, 182 cells:  
Convergence at Iteration 10  
Input 0.103  
Group # 1 -- X: 0.136, o: 0.835, i: 0.306  
Group # 3 -- 0: 0.445, t: 0.998  
Group # 6 -- 0: 0.506, P: 0.216  
Group # 8 -- A: 0.517, B: 0.803, D: 0.283, E: 0.523, F: 0.424, G: 0.306,  
H: 0.426, J: 0.491, K: 0.698, C: 0.829  
Group # 9 -- B: 0.579, R: 0.622, L: 0.369, F: 0.420  
Log likelihood = -864.564 Significance = 0.054

Run # 36, 369 cells:  
Convergence at Iteration 9  
Input 0.099  
Group # 1 -- X: 0.111, o: 0.809, i: 0.371  
Group # 3 -- 0: 0.449, t: 0.997  
Group # 7 -- 2: 0.546, 4: 0.668, 3: 0.560, 1: 0.262, +: 0.487, 5: 0.779  
Group # 8 -- A: 0.537, B: 0.821, D: 0.268, E: 0.510, F: 0.419, G: 0.298,  
H: 0.417, J: 0.476, K: 0.718, C: 0.808  
Group # 9 -- B: 0.571, R: 0.631, L: 0.370, F: 0.417  
Log likelihood = -854.568 Significance = 0.000

Add Group # 4 with factors FMI

----- Level # 6 -----

Run # 37, 384 cells:  
No Convergence at Iteration 20  
Input 0.099  
Group # 1 -- X: 0.126, o: 0.766, i: 0.419  
Group # 2 -- T: 0.631, W: 0.477, M: 0.377  
Group # 3 -- 0: 0.443, t: 0.998  
Group # 4 -- F: 0.544, M: 0.387, I: 0.366  
Group # 8 -- A: 0.516, B: 0.800, D: 0.283, E: 0.509, F: 0.446, G: 0.316,  
H: 0.411, J: 0.493, K: 0.710, C: 0.803  
Group # 9 -- B: 0.589, R: 0.626, L: 0.369, F: 0.404  
Log likelihood = -851.242 Significance = 0.012

Run # 38, 371 cells:  
Convergence at Iteration 8  
Input 0.101  
Group # 1 -- X: 0.135, o: 0.842, i: 0.296  
Group # 3 -- 0: 0.449, t: 0.997  
Group # 4 -- F: 0.543, M: 0.393, I: 0.367  
Group # 5 -- 0: 0.489, J: 0.637  
Group # 8 -- A: 0.513, B: 0.802, D: 0.284, E: 0.510, F: 0.438, G: 0.314,  
H: 0.420, J: 0.497, K: 0.701, C: 0.810  
Group # 9 -- B: 0.593, R: 0.625, L: 0.366, F: 0.405  
Log likelihood = -854.445 Significance = 0.115

Run # 39, 350 cells:  
Convergence at Iteration 10  
Input 0.101  
Group # 1 -- X: 0.134, o: 0.837, i: 0.304  
Group # 3 -- 0: 0.444, t: 0.998  
Group # 4 -- F: 0.543, M: 0.389, I: 0.369  
Group # 6 -- 0: 0.506, P: 0.218  
Group # 8 -- A: 0.511, B: 0.803, D: 0.278, E: 0.509, F: 0.441, G: 0.317,  
H: 0.417, J: 0.492, K: 0.706, C: 0.813  
Group # 9 -- B: 0.595, R: 0.625, L: 0.366, F: 0.402  
Log likelihood = -853.911 Significance = 0.060

Run # 40, 615 cells:  
Convergence at Iteration 9  
Input 0.096  
Group # 1 -- X: 0.110, o: 0.811, i: 0.369  
Group # 3 -- O: 0.448, t: 0.997  
Group # 4 -- F: 0.544, M: 0.400, I: 0.365  
Group # 7 -- 2: 0.548, 4: 0.660, 3: 0.561, 1: 0.258, +: 0.503, 5: 0.771  
Group # 8 -- A: 0.530, B: 0.821, D: 0.265, E: 0.497, F: 0.436, G: 0.309,  
H: 0.407, J: 0.477, K: 0.729, C: 0.788  
Group # 9 -- B: 0.588, R: 0.634, L: 0.366, F: 0.400  
Log likelihood = -843.775 Significance = 0.000

Add Group # 7 with factors 2431+5

----- Level # 7 -----

Run # 41, 683 cells:  
No Convergence at Iteration 20  
Input 0.094  
Group # 1 -- X: 0.103, o: 0.757, i: 0.457  
Group # 2 -- T: 0.596, W: 0.412, M: 0.412  
Group # 3 -- O: 0.447, t: 0.997  
Group # 4 -- F: 0.544, M: 0.394, I: 0.364  
Group # 7 -- 2: 0.543, 4: 0.690, 3: 0.565, 1: 0.265, +: 0.488, 5: 0.745  
Group # 8 -- A: 0.536, B: 0.820, D: 0.265, E: 0.496, F: 0.440, G: 0.308,  
H: 0.400, J: 0.476, K: 0.729, C: 0.785  
Group # 9 -- B: 0.586, R: 0.631, L: 0.370, F: 0.401  
Log likelihood = -840.421 Significance = 0.038

Run # 42, 669 cells:  
Convergence at Iteration 10  
Input 0.096  
Group # 1 -- X: 0.110, o: 0.814, i: 0.363  
Group # 3 -- O: 0.449, t: 0.996  
Group # 4 -- F: 0.544, M: 0.400, I: 0.365  
Group # 5 -- O: 0.496, J: 0.554  
Group # 7 -- 2: 0.546, 4: 0.659, 3: 0.558, 1: 0.265, +: 0.498, 5: 0.774  
Group # 8 -- A: 0.529, B: 0.821, D: 0.267, E: 0.497, F: 0.436, G: 0.308,  
H: 0.408, J: 0.479, K: 0.726, C: 0.788  
Group # 9 -- B: 0.588, R: 0.633, L: 0.366, F: 0.402  
Log likelihood = -843.582 Significance = 0.545

Run # 43, 627 cells:  
Convergence at Iteration 11  
Input 0.093  
Group # 1 -- X: 0.095, o: 0.810, i: 0.387  
Group # 3 -- O: 0.448, t: 0.997  
Group # 4 -- F: 0.543, M: 0.398, I: 0.366  
Group # 6 -- O: 0.510, P: 0.115  
Group # 7 -- 2: 0.544, 4: 0.686, 3: 0.568, 1: 0.241, +: 0.692, 5: 0.842  
Group # 8 -- A: 0.531, B: 0.828, D: 0.259, E: 0.497, F: 0.436, G: 0.308,  
H: 0.404, J: 0.469, K: 0.734, C: 0.786  
Group # 9 -- B: 0.591, R: 0.633, L: 0.366, F: 0.398  
Log likelihood = -838.909 Significance = 0.003

Add Group # 6 with factors 0P



----- Level # 8 -----

Run # 44, 693 cells:  
No Convergence at Iteration 20  
Input 0.091  
Group # 1 -- X: 0.090, o: 0.764, i: 0.464  
Group # 2 -- T: 0.583, W: 0.386, M: 0.426  
Group # 3 -- O: 0.447, t: 0.997  
Group # 4 -- F: 0.544, M: 0.392, I: 0.366  
Group # 6 -- O: 0.510, P: 0.118  
Group # 7 -- 2: 0.540, 4: 0.720, 3: 0.575, 1: 0.245, +: 0.676, 5: 0.821  
Group # 8 -- A: 0.538, B: 0.828, D: 0.258, E: 0.496, F: 0.439, G: 0.307,  
H: 0.396, J: 0.467, K: 0.734, C: 0.784  
Group # 9 -- B: 0.589, R: 0.630, L: 0.371, F: 0.398  
Log likelihood = -835.730 Significance = 0.044

Run # 45, 680 cells:  
Convergence at Iteration 11  
Input 0.093  
Group # 1 -- X: 0.095, o: 0.813, i: 0.383  
Group # 3 -- O: 0.449, t: 0.996  
Group # 4 -- F: 0.543, M: 0.398, I: 0.366  
Group # 5 -- O: 0.497, J: 0.541  
Group # 6 -- O: 0.510, P: 0.115  
Group # 7 -- 2: 0.543, 4: 0.684, 3: 0.566, 1: 0.247, +: 0.690, 5: 0.844  
Group # 8 -- A: 0.530, B: 0.828, D: 0.261, E: 0.498, F: 0.436, G: 0.308,  
H: 0.405, J: 0.471, K: 0.732, C: 0.785  
Group # 9 -- B: 0.591, R: 0.632, L: 0.366, F: 0.399  
Log likelihood = -838.806 Significance = 0.662

Add Group # 2 with factors TWM

----- Level # 9 -----

Run # 46, 741 cells:  
No Convergence at Iteration 20  
Input 0.091  
Group # 1 -- X: 0.090, o: 0.765, i: 0.462  
Group # 2 -- T: 0.587, W: 0.390, M: 0.422  
Group # 3 -- O: 0.449, t: 0.997  
Group # 4 -- F: 0.544, M: 0.392, I: 0.366  
Group # 5 -- O: 0.496, J: 0.556  
Group # 6 -- O: 0.510, P: 0.119  
Group # 7 -- 2: 0.537, 4: 0.718, 3: 0.572, 1: 0.253, +: 0.672, 5: 0.821  
Group # 8 -- A: 0.537, B: 0.828, D: 0.260, E: 0.497, F: 0.440, G: 0.306,  
H: 0.397, J: 0.470, K: 0.730, C: 0.783  
Group # 9 -- B: 0.589, R: 0.629, L: 0.371, F: 0.400  
Log likelihood = -835.546 Significance = 0.557

No remaining groups significant

[Note that the contribution of the one remaining factor group to the model log-likelihood is assessed as having a significance of  $p=0.557$ , and so it is not added.]

Groups selected while stepping up: 1 3 8 9 4 7 6 2  
Best stepping up run: #44

[On the step-down, GoldVarb starts with a model containing every coded factor group, and then removes the least significant contribution at each step until all remaining factor groups have a significant contribution to the model. In this particular case, the endpoint (model #52) is the same as model #44 from the step-up runs, though this is not always guaranteed.]

Stepping down...

----- Level # 9 -----

Run # 47, 741 cells:

No Convergence at Iteration 20

Input 0.091

Group # 1 -- X: 0.090, o: 0.765, i: 0.462

Group # 2 -- T: 0.587, W: 0.390, M: 0.422

Group # 3 -- O: 0.449, t: 0.997

Group # 4 -- F: 0.544, M: 0.392, I: 0.366

Group # 5 -- O: 0.496, J: 0.556

Group # 6 -- O: 0.510, P: 0.119

Group # 7 -- 2: 0.537, 4: 0.718, 3: 0.572, 1: 0.253, +: 0.672, 5: 0.821

Group # 8 -- A: 0.537, B: 0.828, D: 0.260, E: 0.497, F: 0.440, G: 0.306, H: 0.397, J: 0.470, K: 0.730, C: 0.783

Group # 9 -- B: 0.589, R: 0.629, L: 0.371, F: 0.400

Log likelihood = -835.546

----- Level # 8 -----

Run # 48, 571 cells:

Convergence at Iteration 11

Input 0.115

Group # 2 -- T: 0.734, W: 0.627, M: 0.274

Group # 3 -- O: 0.448, t: 0.997

Group # 4 -- F: 0.540, M: 0.410, I: 0.374

Group # 5 -- O: 0.496, J: 0.552

Group # 6 -- O: 0.506, P: 0.228

Group # 7 -- 2: 0.547, 4: 0.651, 3: 0.592, 1: 0.276, +: 0.316, 5: 0.341

Group # 8 -- A: 0.531, B: 0.824, D: 0.284, E: 0.492, F: 0.473, G: 0.336, H: 0.367, J: 0.469, K: 0.740, C: 0.678

Group # 9 -- B: 0.588, R: 0.636, L: 0.354, F: 0.410

Log likelihood = -907.220 Significance = 0.000

Run # 49, 680 cells:

Convergence at Iteration 11

Input 0.093

Group # 1 -- X: 0.095, o: 0.813, i: 0.383

Group # 3 -- O: 0.449, t: 0.996

Group # 4 -- F: 0.543, M: 0.398, I: 0.366

Group # 5 -- O: 0.497, J: 0.541

Group # 6 -- O: 0.510, P: 0.115

Group # 7 -- 2: 0.543, 4: 0.684, 3: 0.566, 1: 0.247, +: 0.690, 5: 0.844

Group # 8 -- A: 0.530, B: 0.828, D: 0.261, E: 0.498, F: 0.436, G: 0.308, H: 0.405, J: 0.471, K: 0.732, C: 0.785

Group # 9 -- B: 0.591, R: 0.632, L: 0.366, F: 0.399

Log likelihood = -838.806 Significance = 0.041

Run # 50, 719 cells:

No Convergence at Iteration 20

Input 0.092

Group # 1 -- X: 0.077, o: 0.747, i: 0.506

Group # 2 -- T: 0.601, W: 0.396, M: 0.409

Group # 4 -- F: 0.534, M: 0.457, I: 0.386

Group # 5 -- O: 0.451, J: 0.922

Group # 6 -- O: 0.509, P: 0.135

Group # 7 -- 2: 0.539, 4: 0.705, 3: 0.570, 1: 0.250, +: 0.614, 5: 0.847

Group # 8 -- A: 0.575, B: 0.825, D: 0.262, E: 0.512, F: 0.425, G: 0.293, H: 0.421, J: 0.457, K: 0.643, C: 0.760

Group # 9 -- B: 0.582, R: 0.596, L: 0.383, F: 0.430

Log likelihood = -921.148 Significance = 0.000

Run # 51, 495 cells:  
No Convergence at Iteration 20  
Input 0.093  
Group # 1 -- X: 0.090, o: 0.763, i: 0.465  
Group # 2 -- T: 0.589, W: 0.399, M: 0.420  
Group # 3 -- O: 0.450, t: 0.996  
Group # 5 -- O: 0.495, J: 0.558  
Group # 6 -- O: 0.510, P: 0.114  
Group # 7 -- 2: 0.535, 4: 0.725, 3: 0.569, 1: 0.258, +: 0.656, 5: 0.832  
Group # 8 -- A: 0.544, B: 0.828, D: 0.262, E: 0.511, F: 0.421, G: 0.296,  
H: 0.407, J: 0.469, K: 0.719, C: 0.803  
Group # 9 -- B: 0.571, R: 0.626, L: 0.375, F: 0.418  
Log likelihood = -846.201 Significance = 0.000

Run # 52, 693 cells:  
No Convergence at Iteration 20  
Input 0.091  
Group # 1 -- X: 0.090, o: 0.764, i: 0.464  
Group # 2 -- T: 0.583, W: 0.386, M: 0.426  
Group # 3 -- O: 0.447, t: 0.997  
Group # 4 -- F: 0.544, M: 0.392, I: 0.366  
Group # 6 -- O: 0.510, P: 0.118  
Group # 7 -- 2: 0.540, 4: 0.720, 3: 0.575, 1: 0.245, +: 0.676, 5: 0.821  
Group # 8 -- A: 0.538, B: 0.828, D: 0.258, E: 0.496, F: 0.439, G: 0.307,  
H: 0.396, J: 0.467, K: 0.734, C: 0.784  
Group # 9 -- B: 0.589, R: 0.630, L: 0.371, F: 0.398  
Log likelihood = -835.730 Significance = 0.557

Run # 53, 732 cells:  
No Convergence at Iteration 20  
Input 0.094  
Group # 1 -- X: 0.103, o: 0.759, i: 0.454  
Group # 2 -- T: 0.602, W: 0.417, M: 0.407  
Group # 3 -- O: 0.449, t: 0.996  
Group # 4 -- F: 0.544, M: 0.394, I: 0.364  
Group # 5 -- O: 0.494, J: 0.575  
Group # 7 -- 2: 0.540, 4: 0.689, 3: 0.561, 1: 0.276, +: 0.480, 5: 0.746  
Group # 8 -- A: 0.535, B: 0.820, D: 0.267, E: 0.497, F: 0.441, G: 0.307,  
H: 0.401, J: 0.479, K: 0.725, C: 0.784  
Group # 9 -- B: 0.586, R: 0.630, L: 0.370, F: 0.403  
Log likelihood = -840.084 Significance = 0.005

Run # 54, 476 cells:  
No Convergence at Iteration 20  
Input 0.097  
Group # 1 -- X: 0.124, o: 0.765, i: 0.422  
Group # 2 -- T: 0.640, W: 0.485, M: 0.369  
Group # 3 -- O: 0.449, t: 0.997  
Group # 4 -- F: 0.544, M: 0.385, I: 0.367  
Group # 5 -- O: 0.487, J: 0.658  
Group # 6 -- O: 0.506, P: 0.219  
Group # 8 -- A: 0.513, B: 0.805, D: 0.285, E: 0.511, F: 0.447, G: 0.314,  
H: 0.412, J: 0.494, K: 0.706, C: 0.801  
Group # 9 -- B: 0.592, R: 0.622, L: 0.368, F: 0.407  
Log likelihood = -847.775 Significance = 0.000

Run # 55, 257 cells:  
No Convergence at Iteration 20  
Input 0.105  
Group # 1 -- X: 0.115, o: 0.753, i: 0.447  
Group # 2 -- T: 0.616, W: 0.454, M: 0.392  
Group # 3 -- O: 0.450, t: 0.996  
Group # 4 -- F: 0.549, M: 0.430, I: 0.339  
Group # 5 -- O: 0.492, J: 0.597  
Group # 6 -- O: 0.508, P: 0.168  
Group # 7 -- 2: 0.546, 4: 0.667, 3: 0.465, 1: 0.358, +: 0.588, 5: 0.806  
Group # 9 -- B: 0.594, R: 0.642, L: 0.359, F: 0.392  
Log likelihood = -902.388 Significance = 0.000

Run # 56, 390 cells:  
No Convergence at Iteration 20  
Input 0.096  
Group # 1 -- X: 0.085, o: 0.761, i: 0.474  
Group # 2 -- T: 0.598, W: 0.350, M: 0.413  
Group # 3 -- O: 0.452, t: 0.995  
Group # 4 -- F: 0.536, M: 0.391, I: 0.395  
Group # 5 -- O: 0.492, J: 0.603  
Group # 6 -- O: 0.509, P: 0.126  
Group # 7 -- 2: 0.529, 4: 0.707, 3: 0.574, 1: 0.268, +: 0.621, 5: 0.835  
Group # 8 -- A: 0.551, B: 0.833, D: 0.278, E: 0.483, F: 0.414, G: 0.304,  
H: 0.402, J: 0.462, K: 0.720, C: 0.811  
Log likelihood = -860.725 Significance = 0.000

Cut Group # 5 with factors OJ

----- Level # 7 -----

[The last level of the step-down runs is of interest because GoldVarb is now assessing the significance of the contribution of each factor group contained within the best-fit model; hence we can say the contribution of factor group #1 has a significance of  $p < 0.001$ ; #2,  $p = 0.044$ ; #3,  $p < 0.001$ ; #4,  $p < 0.001$ ; #6,  $p = 0.004$ ; #7,  $p < 0.001$ ; #8,  $p < 0.001$ ; #9,  $p < 0.001$ .]

Run # 57, 528 cells:  
Convergence at Iteration 10  
Input 0.115  
Group # 2 -- T: 0.729, W: 0.621, M: 0.279  
Group # 3 -- O: 0.446, t: 0.997  
Group # 4 -- F: 0.540, M: 0.409, I: 0.374  
Group # 6 -- O: 0.506, P: 0.224  
Group # 7 -- 2: 0.549, 4: 0.653, 3: 0.595, 1: 0.268, +: 0.319, 5: 0.343  
Group # 8 -- A: 0.532, B: 0.824, D: 0.282, E: 0.492, F: 0.473, G: 0.336,  
H: 0.366, J: 0.467, K: 0.743, C: 0.679  
Group # 9 -- B: 0.588, R: 0.637, L: 0.354, F: 0.409  
Log likelihood = -907.400 Significance = 0.000

Run # 58, 627 cells:  
Convergence at Iteration 11  
Input 0.093  
Group # 1 -- X: 0.095, o: 0.810, i: 0.387  
Group # 3 -- O: 0.448, t: 0.997  
Group # 4 -- F: 0.543, M: 0.398, I: 0.366  
Group # 6 -- O: 0.510, P: 0.115  
Group # 7 -- 2: 0.544, 4: 0.686, 3: 0.568, 1: 0.241, +: 0.692, 5: 0.842  
Group # 8 -- A: 0.531, B: 0.828, D: 0.259, E: 0.497, F: 0.436, G: 0.308,  
H: 0.404, J: 0.469, K: 0.734, C: 0.786  
Group # 9 -- B: 0.591, R: 0.633, L: 0.366, F: 0.398  
Log likelihood = -838.909 Significance = 0.044

Run # 59, 653 cells:  
Convergence at Iteration 12  
Input 0.105  
Group # 1 -- X: 0.079, o: 0.690, i: 0.576  
Group # 2 -- T: 0.525, W: 0.311, M: 0.484  
Group # 4 -- F: 0.532, M: 0.494, I: 0.384  
Group # 6 -- O: 0.510, P: 0.103  
Group # 7 -- 2: 0.592, 4: 0.748, 3: 0.638, 1: 0.111, +: 0.712, 5: 0.840  
Group # 8 -- A: 0.615, B: 0.831, D: 0.200, E: 0.507, F: 0.414, G: 0.303,  
H: 0.399, J: 0.380, K: 0.688, C: 0.756  
Group # 9 -- B: 0.586, R: 0.609, L: 0.394, F: 0.400  
Log likelihood = -987.526 Significance = 0.000

Run # 60, 454 cells:  
No Convergence at Iteration 20  
Input 0.093  
Group # 1 -- X: 0.089, o: 0.762, i: 0.468  
Group # 2 -- T: 0.585, W: 0.394, M: 0.424  
Group # 3 -- O: 0.448, t: 0.997  
Group # 6 -- O: 0.510, P: 0.112  
Group # 7 -- 2: 0.537, 4: 0.728, 3: 0.573, 1: 0.250, +: 0.659, 5: 0.832  
Group # 8 -- A: 0.545, B: 0.828, D: 0.260, E: 0.510, F: 0.421, G: 0.296,  
H: 0.406, J: 0.466, K: 0.722, C: 0.804  
Group # 9 -- B: 0.572, R: 0.627, L: 0.375, F: 0.417  
Log likelihood = -846.401 Significance = 0.000

Run # 61, 683 cells:  
No Convergence at Iteration 20  
Input 0.094  
Group # 1 -- X: 0.103, o: 0.757, i: 0.457  
Group # 2 -- T: 0.596, W: 0.412, M: 0.412  
Group # 3 -- O: 0.447, t: 0.997  
Group # 4 -- F: 0.544, M: 0.394, I: 0.364  
Group # 7 -- 2: 0.543, 4: 0.690, 3: 0.565, 1: 0.265, +: 0.488, 5: 0.745  
Group # 8 -- A: 0.536, B: 0.820, D: 0.265, E: 0.496, F: 0.440, G: 0.308,  
H: 0.400, J: 0.476, K: 0.729, C: 0.785  
Group # 9 -- B: 0.586, R: 0.631, L: 0.370, F: 0.401  
Log likelihood = -840.421 Significance = 0.004

Run # 62, 421 cells:  
No Convergence at Iteration 20  
Input 0.098  
Group # 1 -- X: 0.124, o: 0.765, i: 0.422  
Group # 2 -- T: 0.631, W: 0.475, M: 0.378  
Group # 3 -- O: 0.443, t: 0.998  
Group # 4 -- F: 0.544, M: 0.384, I: 0.367  
Group # 6 -- O: 0.506, P: 0.216  
Group # 8 -- A: 0.514, B: 0.803, D: 0.281, E: 0.510, F: 0.447, G: 0.316,  
H: 0.410, J: 0.491, K: 0.711, C: 0.804  
Group # 9 -- B: 0.592, R: 0.625, L: 0.368, F: 0.404  
Log likelihood = -849.420 Significance = 0.000

Run # 63, 226 cells:  
No Convergence at Iteration 20  
Input 0.105  
Group # 1 -- X: 0.115, o: 0.751, i: 0.450  
Group # 2 -- T: 0.610, W: 0.447, M: 0.398  
Group # 3 -- O: 0.447, t: 0.997  
Group # 4 -- F: 0.549, M: 0.431, I: 0.339  
Group # 6 -- O: 0.508, P: 0.165  
Group # 7 -- 2: 0.549, 4: 0.670, 3: 0.469, 1: 0.344, +: 0.592, 5: 0.805  
Group # 9 -- B: 0.594, R: 0.644, L: 0.359, F: 0.390  
Log likelihood = -902.978 Significance = 0.000

Run # 64, 349 cells:  
 No Convergence at Iteration 20  
 Input 0.097  
 Group # 1 -- X: 0.085, o: 0.759, i: 0.477  
 Group # 2 -- T: 0.591, W: 0.342, M: 0.421  
 Group # 3 -- O: 0.449, t: 0.997  
 Group # 4 -- F: 0.536, M: 0.391, I: 0.395  
 Group # 6 -- O: 0.509, P: 0.123  
 Group # 7 -- 2: 0.533, 4: 0.709, 3: 0.580, 1: 0.254, +: 0.629, 5: 0.834  
 Group # 8 -- A: 0.552, B: 0.833, D: 0.275, E: 0.482, F: 0.413, G: 0.304,  
 H: 0.400, J: 0.457, K: 0.726, C: 0.813  
 Log likelihood = -861.373 Significance = 0.000

All remaining groups significant

Groups eliminated while stepping down: 5  
 Best stepping up run: #44  
 Best stepping down run: #52

[A second step-up/step-down regression run was then performed after recoding the corpus factor group to check the contributions of the factors of corpus regional variety (AmE vs. BrE) and age (older vs. newer corpora).]

```

•          CELL          CREATION          •          2015/08/25          13:58:39
•.....
Name of token file: month.tkn
Name of condition file: month2.cnd
(
(1)
; factor group 2: prepositions other than on/in mostly knockouts, so
combined.
(2 (i (COL 2 i))
(o (COL 2 o))
(X (ELSEWHERE)))
; factor group 3: D=specified day(n=7) combined with W to avoid knockout
(3 (W (COL 3 D)))
(4)
(5)
(6)
(7)
; factor group 8: longer adverbials combined to avoid knockouts.
(8 (+ (OR (COL 8 6) (COL 8 7) (COL 8 8) (COL 8 9))))
(9 (K (OR (COL 9 L) (COL 9 M) (COL 9 N) (COL 9 P) (COL 9 R))))
; corpus distinctions recoded to generalise across time difference and
regional difference.
; factor group #9 is the corpus age (older vs. newer).
(10 (o (OR (COL 10 L) (COL 10 B)))
(n (OR (COL 10 F) (COL 10 R))))
; factor group #10 is the corpus variety (American vs. British English).
(10 (a (OR (COL 10 B) (COL 10 R)))
(b (OR (COL 10 L) (COL 10 F))))
)

```

[The raw token frequencies for each factor group in turn are exactly the same as for the first series of runs (and so are not repeated here) until we get to the last two factor groups.]

Number of cells: 741  
 Application value(s): 0P  
 Total no. of factors: 35

Group		0	P	Total	%
[...results for factor groups 1-8 exactly as above...]					
9 (10)		0	P		
o	N	269	1078	1347	50.7
	%	20.0	80.0		
n	N	255	1055	1310	49.3
	%	19.5	80.5		
Total	N	524	2133	2657	
	%	19.7	80.3		
10 (10)		0	P		
a	N	357	1010	1367	51.4
	%	26.1	73.9		
b	N	167	1123	1290	48.6
	%	12.9	87.1		
Total	N	524	2133	2657	
	%	19.7	80.3		
TOTAL	N	524	2133	2657	
	%	19.7	80.3		

Name of new cell file: month2.cel

• BINOMIAL VARBRUL • 2015/08/25 13:59:08  
 .....  
 Name of cell file: month2.cel

Averaging by weighting factors.  
 Threshold, step-up/down: 0.050001

Stepping up...

----- Level # 0 -----

Run # 1, 1 cells:  
 Convergence at Iteration 2  
 Input 0.197  
 Log likelihood = -1319.246

----- Level # 1 -----

Run # 2, 3 cells:  
 Convergence at Iteration 6  
 Input 0.152  
 Group # 1 -- X: 0.132, o: 0.746, i: 0.440  
 Log likelihood = -1167.453 Significance = 0.000

Run # 3, 3 cells:  
Convergence at Iteration 5  
Input 0.178  
Group # 2 -- T: 0.661, W: 0.503, M: 0.348  
Log likelihood = -1239.551 Significance = 0.000

Run # 4, 2 cells:  
Convergence at Iteration 6  
Input 0.194  
Group # 3 -- 0: 0.459, t: 0.989  
Log likelihood = -1183.994 Significance = 0.000

Run # 5, 3 cells:  
Convergence at Iteration 4  
Input 0.196  
Group # 4 -- F: 0.527, M: 0.464, I: 0.410  
Log likelihood = -1312.447 Significance = 0.001

Run # 6, 2 cells:  
Convergence at Iteration 5  
Input 0.190  
Group # 5 -- 0: 0.471, J: 0.810  
Log likelihood = -1269.808 Significance = 0.000

Run # 7, 2 cells:  
Convergence at Iteration 5  
Input 0.196  
Group # 6 -- 0: 0.506, P: 0.208  
Log likelihood = -1315.343 Significance = 0.007

Run # 8, 6 cells:  
Convergence at Iteration 6  
Input 0.167  
Group # 7 -- 2: 0.619, 4: 0.614, 3: 0.592, 1: 0.140, +: 0.277, 5: 0.478  
Log likelihood = -1243.983 Significance = 0.000

Run # 9, 10 cells:  
Convergence at Iteration 5  
Input 0.182  
Group # 8 -- A: 0.592, B: 0.787, D: 0.231, E: 0.496, F: 0.404, G: 0.355,  
H: 0.444, J: 0.410, K: 0.548, C: 0.737  
Log likelihood = -1239.995 Significance = 0.000

Run # 10, 2 cells:  
Convergence at Iteration 2  
Input 0.197  
Group # 9 -- o: 0.503, n: 0.496  
Log likelihood = -1319.193 Significance = 0.751

Run # 11, 2 cells:  
Convergence at Iteration 4  
Input 0.188  
Group #10 -- a: 0.603, b: 0.391  
Log likelihood = -1282.124 Significance = 0.000

Add Group # 1 with factors Xoi



----- Level # 2 -----

Run # 12, 8 cells:  
No Convergence at Iteration 20  
Input 0.151  
Group # 1 -- X: 0.129, o: 0.693, i: 0.510  
Group # 2 -- T: 0.579, W: 0.365, M: 0.431  
Log likelihood = -1162.563 Significance = 0.009

Run # 13, 6 cells:  
Convergence at Iteration 8  
Input 0.126  
Group # 1 -- X: 0.156, o: 0.816, i: 0.318  
Group # 3 -- O: 0.448, t: 0.997  
Log likelihood = -968.439 Significance = 0.000

Run # 14, 9 cells:  
Convergence at Iteration 6  
Input 0.150  
Group # 1 -- X: 0.129, o: 0.748, i: 0.439  
Group # 4 -- F: 0.531, M: 0.498, I: 0.384  
Log likelihood = -1157.739 Significance = 0.000

Run # 15, 6 cells:  
Convergence at Iteration 10  
Input 0.122  
Group # 1 -- X: 0.119, o: 0.821, i: 0.344  
Group # 5 -- O: 0.446, J: 0.937  
Log likelihood = -1053.411 Significance = 0.000

Run # 16, 6 cells:  
Convergence at Iteration 6  
Input 0.151  
Group # 1 -- X: 0.131, o: 0.744, i: 0.442  
Group # 6 -- O: 0.505, P: 0.263  
Log likelihood = -1165.483 Significance = 0.048

Run # 17, 18 cells:  
Convergence at Iteration 7  
Input 0.135  
Group # 1 -- X: 0.116, o: 0.724, i: 0.486  
Group # 7 -- 2: 0.605, 4: 0.621, 3: 0.501, 1: 0.204, +: 0.443, 5: 0.773  
Log likelihood = -1128.986 Significance = 0.000

Run # 18, 30 cells:  
Convergence at Iteration 6  
Input 0.136  
Group # 1 -- X: 0.117, o: 0.758, i: 0.438  
Group # 8 -- A: 0.589, B: 0.792, D: 0.238, E: 0.514, F: 0.393, G: 0.313,  
H: 0.437, J: 0.427, K: 0.640, C: 0.803  
Log likelihood = -1081.216 Significance = 0.000

Run # 19, 6 cells:  
Convergence at Iteration 6  
Input 0.152  
Group # 1 -- X: 0.132, o: 0.746, i: 0.439  
Group # 9 -- o: 0.496, n: 0.504  
Log likelihood = -1167.398 Significance = 0.745

Run # 20, 6 cells:  
Convergence at Iteration 6  
Input 0.146  
Group # 1 -- X: 0.135, o: 0.744, i: 0.439  
Group #10 -- a: 0.599, b: 0.396  
Log likelihood = -1137.244 Significance = 0.000

Add Group # 3 with factors 0t

----- Level # 3 -----

Run # 21, 12 cells:  
No Convergence at Iteration 20  
Input 0.122  
Group # 1 -- X: 0.146, o: 0.728, i: 0.450  
Group # 2 -- T: 0.645, W: 0.434, M: 0.366  
Group # 3 -- 0: 0.447, t: 0.997  
Log likelihood = -961.033 Significance = 0.001

Run # 22, 16 cells:  
Convergence at Iteration 8  
Input 0.123  
Group # 1 -- X: 0.153, o: 0.819, i: 0.318  
Group # 3 -- 0: 0.447, t: 0.997  
Group # 4 -- F: 0.541, M: 0.430, I: 0.366  
Log likelihood = -956.529 Significance = 0.000

Run # 23, 11 cells:  
Convergence at Iteration 9  
Input 0.125  
Group # 1 -- X: 0.155, o: 0.821, i: 0.312  
Group # 3 -- 0: 0.453, t: 0.995  
Group # 5 -- 0: 0.488, J: 0.644  
Log likelihood = -966.973 Significance = 0.090

Run # 24, 10 cells:  
Convergence at Iteration 8  
Input 0.125  
Group # 1 -- X: 0.156, o: 0.815, i: 0.320  
Group # 3 -- 0: 0.448, t: 0.997  
Group # 6 -- 0: 0.505, P: 0.281  
Log likelihood = -967.323 Significance = 0.144

Run # 25, 26 cells:  
Convergence at Iteration 9  
Input 0.120  
Group # 1 -- X: 0.134, o: 0.813, i: 0.342  
Group # 3 -- 0: 0.450, t: 0.996  
Group # 7 -- 2: 0.549, 4: 0.609, 3: 0.463, 1: 0.364, +: 0.438, 5: 0.796  
Log likelihood = -957.738 Significance = 0.001

Run # 26, 44 cells:  
Convergence at Iteration 9  
Input 0.110  
Group # 1 -- X: 0.135, o: 0.833, i: 0.310  
Group # 3 -- 0: 0.446, t: 0.997  
Group # 8 -- A: 0.529, B: 0.806, D: 0.286, E: 0.500, F: 0.400, G: 0.306,  
H: 0.432, J: 0.487, K: 0.702, C: 0.843  
Log likelihood = -889.788 Significance = 0.000

Run # 27, 12 cells:  
Convergence at Iteration 8  
Input 0.126  
Group # 1 -- X: 0.158, o: 0.818, i: 0.315  
Group # 3 -- 0: 0.447, t: 0.997  
Group # 9 -- o: 0.481, n: 0.519  
Log likelihood = -967.530 Significance = 0.183

Run # 28, 12 cells:  
Convergence at Iteration 8  
Input 0.119  
Group # 1 -- X: 0.160, o: 0.816, i: 0.316  
Group # 3 -- 0: 0.447, t: 0.997  
Group #10 -- a: 0.608, b: 0.386  
Log likelihood = -938.872 Significance = 0.000

Add Group # 8 with factors ABDEFGHJKC

----- Level # 4 -----

Run # 29, 72 cells:  
No Convergence at Iteration 20  
Input 0.107  
Group # 1 -- X: 0.124, o: 0.758, i: 0.431  
Group # 2 -- T: 0.634, W: 0.438, M: 0.376  
Group # 3 -- 0: 0.446, t: 0.998  
Group # 8 -- A: 0.532, B: 0.805, D: 0.290, E: 0.504, F: 0.408, G: 0.306,  
H: 0.424, J: 0.484, K: 0.705, C: 0.837  
Log likelihood = -884.127 Significance = 0.006

Run # 30, 112 cells:  
Convergence at Iteration 9  
Input 0.109  
Group # 1 -- X: 0.134, o: 0.835, i: 0.308  
Group # 3 -- 0: 0.446, t: 0.998  
Group # 4 -- F: 0.535, M: 0.390, I: 0.397  
Group # 8 -- A: 0.525, B: 0.806, D: 0.289, E: 0.488, F: 0.407, G: 0.316,  
H: 0.426, J: 0.486, K: 0.709, C: 0.833  
Log likelihood = -882.476 Significance = 0.001

Run # 31, 65 cells:  
Convergence at Iteration 9  
Input 0.109  
Group # 1 -- X: 0.134, o: 0.838, i: 0.303  
Group # 3 -- 0: 0.452, t: 0.995  
Group # 5 -- 0: 0.487, J: 0.659  
Group # 8 -- A: 0.529, B: 0.808, D: 0.290, E: 0.501, F: 0.400, G: 0.304,  
H: 0.434, J: 0.490, K: 0.695, C: 0.841  
Log likelihood = -888.022 Significance = 0.064

Run # 32, 62 cells:  
Convergence at Iteration 9  
Input 0.109  
Group # 1 -- X: 0.133, o: 0.832, i: 0.312  
Group # 3 -- 0: 0.446, t: 0.997  
Group # 6 -- 0: 0.506, P: 0.229  
Group # 8 -- A: 0.527, B: 0.809, D: 0.284, E: 0.500, F: 0.401, G: 0.307,  
H: 0.431, J: 0.484, K: 0.702, C: 0.845  
Log likelihood = -888.078 Significance = 0.069

Run # 33, 147 cells:  
Convergence at Iteration 9  
Input 0.104  
Group # 1 -- X: 0.105, o: 0.809, i: 0.377  
Group # 3 -- o: 0.450, t: 0.996  
Group # 7 -- 2: 0.540, 4: 0.642, 3: 0.564, 1: 0.272, +: 0.435, 5: 0.797  
Group # 8 -- A: 0.547, B: 0.827, D: 0.275, E: 0.490, F: 0.400, G: 0.298,  
H: 0.419, J: 0.469, K: 0.715, C: 0.824  
Log likelihood = -878.329 Significance = 0.000

Run # 34, 79 cells:  
Convergence at Iteration 9  
Input 0.110  
Group # 1 -- X: 0.135, o: 0.835, i: 0.307  
Group # 3 -- o: 0.446, t: 0.997  
Group # 8 -- A: 0.532, B: 0.803, D: 0.284, E: 0.506, F: 0.396, G: 0.303,  
H: 0.437, J: 0.486, K: 0.698, C: 0.849  
Group # 9 -- o: 0.477, n: 0.524  
Log likelihood = -888.622 Significance = 0.135

Run # 35, 81 cells:  
Convergence at Iteration 10  
Input 0.105  
Group # 1 -- X: 0.137, o: 0.834, i: 0.307  
Group # 3 -- o: 0.445, t: 0.998  
Group # 8 -- A: 0.517, B: 0.804, D: 0.287, E: 0.516, F: 0.426, G: 0.309,  
H: 0.422, J: 0.494, K: 0.701, C: 0.821  
Group #10 -- a: 0.599, b: 0.395  
Log likelihood = -867.746 Significance = 0.000

Add Group # 10 with factors ab

----- Level # 5 -----

Run # 36, 121 cells:  
No Convergence at Iteration 20  
Input 0.102  
Group # 1 -- X: 0.127, o: 0.763, i: 0.422  
Group # 2 -- T: 0.627, W: 0.478, M: 0.382  
Group # 3 -- o: 0.445, t: 0.998  
Group # 8 -- A: 0.520, B: 0.804, D: 0.288, E: 0.518, F: 0.431, G: 0.309,  
H: 0.415, J: 0.492, K: 0.705, C: 0.814  
Group #10 -- a: 0.597, b: 0.397  
Log likelihood = -863.528 Significance = 0.016

Run # 37, 190 cells:  
Convergence at Iteration 10  
Input 0.102  
Group # 1 -- X: 0.135, o: 0.837, i: 0.304  
Group # 3 -- o: 0.444, t: 0.998  
Group # 4 -- F: 0.544, M: 0.387, I: 0.365  
Group # 8 -- A: 0.511, B: 0.803, D: 0.283, E: 0.504, F: 0.442, G: 0.319,  
H: 0.414, J: 0.495, K: 0.708, C: 0.806  
Group #10 -- a: 0.609, b: 0.384  
Log likelihood = -856.398 Significance = 0.000

Run # 38, 115 cells:  
Convergence at Iteration 8  
Input 0.104  
Group # 1 -- X: 0.136, o: 0.838, i: 0.301  
Group # 3 -- 0: 0.450, t: 0.996  
Group # 5 -- 0: 0.489, J: 0.636  
Group # 8 -- A: 0.517, B: 0.806, D: 0.290, E: 0.518, F: 0.426, G: 0.307,  
H: 0.424, J: 0.496, K: 0.696, C: 0.820  
Group #10 -- a: 0.598, b: 0.396  
Log likelihood = -866.492 Significance = 0.119

Run # 39, 108 cells:  
Convergence at Iteration 10  
Input 0.104  
Group # 1 -- X: 0.136, o: 0.833, i: 0.309  
Group # 3 -- 0: 0.445, t: 0.998  
Group # 6 -- 0: 0.506, P: 0.208  
Group # 8 -- A: 0.514, B: 0.807, D: 0.284, E: 0.517, F: 0.428, G: 0.310,  
H: 0.422, J: 0.491, K: 0.701, C: 0.823  
Group #10 -- a: 0.600, b: 0.394  
Log likelihood = -865.731 Significance = 0.046

Run # 40, 235 cells:  
Convergence at Iteration 10  
Input 0.099  
Group # 1 -- X: 0.110, o: 0.806, i: 0.375  
Group # 3 -- 0: 0.449, t: 0.997  
Group # 7 -- 2: 0.544, 4: 0.668, 3: 0.564, 1: 0.264, +: 0.477, 5: 0.767  
Group # 8 -- A: 0.534, B: 0.826, D: 0.271, E: 0.505, F: 0.422, G: 0.302,  
H: 0.411, J: 0.475, K: 0.722, C: 0.800  
Group #10 -- a: 0.600, b: 0.394  
Log likelihood = -856.268 Significance = 0.000

Run # 41, 149 cells:  
Convergence at Iteration 10  
Input 0.104  
Group # 1 -- X: 0.138, o: 0.835, i: 0.304  
Group # 3 -- 0: 0.445, t: 0.998  
Group # 8 -- A: 0.520, B: 0.800, D: 0.285, E: 0.522, F: 0.423, G: 0.305,  
H: 0.426, J: 0.493, K: 0.697, C: 0.828  
Group # 9 -- o: 0.475, n: 0.525  
Group #10 -- a: 0.600, b: 0.395  
Log likelihood = -866.435 Significance = 0.108

Add Group # 4 with factors FMI

----- Level # 6 -----

Run # 42, 244 cells:  
No Convergence at Iteration 20  
Input 0.099  
Group # 1 -- X: 0.125, o: 0.766, i: 0.418  
Group # 2 -- T: 0.628, W: 0.469, M: 0.380  
Group # 3 -- 0: 0.443, t: 0.998  
Group # 4 -- F: 0.545, M: 0.382, I: 0.364  
Group # 8 -- A: 0.514, B: 0.803, D: 0.285, E: 0.505, F: 0.448, G: 0.319,  
H: 0.407, J: 0.494, K: 0.712, C: 0.797  
Group #10 -- a: 0.607, b: 0.386  
Log likelihood = -851.967 Significance = 0.013

Run # 43, 237 cells:  
Convergence at Iteration 8  
Input 0.101  
Group # 1 -- X: 0.134, o: 0.841, i: 0.298  
Group # 3 -- o: 0.449, t: 0.997  
Group # 4 -- F: 0.544, M: 0.388, I: 0.365  
Group # 5 -- o: 0.489, J: 0.637  
Group # 8 -- A: 0.511, B: 0.804, D: 0.287, E: 0.505, F: 0.441, G: 0.318,  
H: 0.416, J: 0.497, K: 0.703, C: 0.804  
Group #10 -- a: 0.608, b: 0.385  
Log likelihood = -855.117 Significance = 0.114

Run # 44, 222 cells:  
Convergence at Iteration 10  
Input 0.101  
Group # 1 -- X: 0.133, o: 0.836, i: 0.306  
Group # 3 -- o: 0.444, t: 0.998  
Group # 4 -- F: 0.544, M: 0.385, I: 0.366  
Group # 6 -- o: 0.506, P: 0.213  
Group # 8 -- A: 0.509, B: 0.806, D: 0.281, E: 0.505, F: 0.444, G: 0.320,  
H: 0.414, J: 0.492, K: 0.708, C: 0.807  
Group #10 -- a: 0.610, b: 0.384  
Log likelihood = -854.495 Significance = 0.052

Run # 45, 420 cells:  
Convergence at Iteration 9  
Input 0.096  
Group # 1 -- X: 0.110, o: 0.809, i: 0.372  
Group # 3 -- o: 0.448, t: 0.997  
Group # 4 -- F: 0.545, M: 0.394, I: 0.362  
Group # 7 -- 2: 0.547, 4: 0.659, 3: 0.564, 1: 0.259, +: 0.496, 5: 0.762  
Group # 8 -- A: 0.528, B: 0.824, D: 0.269, E: 0.493, F: 0.438, G: 0.312,  
H: 0.403, J: 0.477, K: 0.731, C: 0.781  
Group #10 -- a: 0.610, b: 0.383  
Log likelihood = -844.733 Significance = 0.000

Run # 46, 314 cells:  
Convergence at Iteration 10  
Input 0.102  
Group # 1 -- X: 0.135, o: 0.838, i: 0.302  
Group # 3 -- o: 0.444, t: 0.998  
Group # 4 -- F: 0.543, M: 0.392, I: 0.368  
Group # 8 -- A: 0.513, B: 0.800, D: 0.281, E: 0.509, F: 0.439, G: 0.316,  
H: 0.418, J: 0.494, K: 0.705, C: 0.811  
Group # 9 -- o: 0.482, n: 0.518  
Group #10 -- a: 0.610, b: 0.384  
Log likelihood = -855.721 Significance = 0.251

Add Group # 7 with factors 2431+5

----- Level # 7 -----

Run # 47, 481 cells:  
No Convergence at Iteration 20  
Input 0.094  
Group # 1 -- X: 0.102, o: 0.757, i: 0.457  
Group # 2 -- T: 0.594, W: 0.401, M: 0.415  
Group # 3 -- o: 0.447, t: 0.997  
Group # 4 -- F: 0.545, M: 0.388, I: 0.361  
Group # 7 -- 2: 0.542, 4: 0.690, 3: 0.568, 1: 0.265, +: 0.481, 5: 0.736  
Group # 8 -- A: 0.534, B: 0.824, D: 0.268, E: 0.493, F: 0.442, G: 0.311,  
H: 0.396, J: 0.476, K: 0.731, C: 0.778  
Group #10 -- a: 0.608, b: 0.386  
Log likelihood = -841.281 Significance = 0.035

Run # 48, 470 cells:  
Convergence at Iteration 10  
Input 0.096  
Group # 1 -- X: 0.110, o: 0.812, i: 0.366  
Group # 3 -- 0: 0.449, t: 0.996  
Group # 4 -- F: 0.545, M: 0.394, I: 0.362  
Group # 5 -- 0: 0.495, J: 0.557  
Group # 7 -- 2: 0.545, 4: 0.658, 3: 0.561, 1: 0.266, +: 0.491, 5: 0.765  
Group # 8 -- A: 0.527, B: 0.824, D: 0.271, E: 0.493, F: 0.438, G: 0.311,  
H: 0.404, J: 0.479, K: 0.728, C: 0.780  
Group #10 -- a: 0.610, b: 0.384  
Log likelihood = -844.524 Significance = 0.524

Run # 49, 436 cells:  
Convergence at Iteration 11  
Input 0.093  
Group # 1 -- X: 0.095, o: 0.809, i: 0.390  
Group # 3 -- 0: 0.448, t: 0.997  
Group # 4 -- F: 0.545, M: 0.392, I: 0.363  
Group # 6 -- 0: 0.510, P: 0.112  
Group # 7 -- 2: 0.543, 4: 0.686, 3: 0.571, 1: 0.242, +: 0.685, 5: 0.837  
Group # 8 -- A: 0.529, B: 0.831, D: 0.262, E: 0.493, F: 0.438, G: 0.311,  
H: 0.400, J: 0.469, K: 0.737, C: 0.779  
Group #10 -- a: 0.612, b: 0.382  
Log likelihood = -839.723 Significance = 0.003

Run # 50, 615 cells:  
Convergence at Iteration 10  
Input 0.096  
Group # 1 -- X: 0.110, o: 0.811, i: 0.369  
Group # 3 -- 0: 0.448, t: 0.997  
Group # 4 -- F: 0.544, M: 0.400, I: 0.365  
Group # 7 -- 2: 0.548, 4: 0.659, 3: 0.561, 1: 0.258, +: 0.501, 5: 0.770  
Group # 8 -- A: 0.530, B: 0.821, D: 0.265, E: 0.497, F: 0.435, G: 0.309,  
H: 0.408, J: 0.477, K: 0.729, C: 0.788  
Group # 9 -- o: 0.479, n: 0.522  
Group #10 -- a: 0.611, b: 0.383  
Log likelihood = -843.794 Significance = 0.178

Add Group # 6 with factors 0P

----- Level # 8 -----

Run # 51, 495 cells:  
No Convergence at Iteration 20  
Input 0.092  
Group # 1 -- X: 0.089, o: 0.764, i: 0.464  
Group # 2 -- T: 0.581, W: 0.376, M: 0.429  
Group # 3 -- 0: 0.447, t: 0.997  
Group # 4 -- F: 0.545, M: 0.387, I: 0.363  
Group # 6 -- 0: 0.510, P: 0.115  
Group # 7 -- 2: 0.539, 4: 0.722, 3: 0.578, 1: 0.245, +: 0.670, 5: 0.816  
Group # 8 -- A: 0.536, B: 0.831, D: 0.261, E: 0.493, F: 0.441, G: 0.309,  
H: 0.393, J: 0.467, K: 0.736, C: 0.778  
Group #10 -- a: 0.609, b: 0.385  
Log likelihood = -836.425 Significance = 0.040

Run # 52, 485 cells:  
Convergence at Iteration 12  
Input 0.093  
Group # 1 -- X: 0.095, o: 0.811, i: 0.386  
Group # 3 -- 0: 0.449, t: 0.996  
Group # 4 -- F: 0.545, M: 0.392, I: 0.363  
Group # 5 -- 0: 0.497, J: 0.543  
Group # 6 -- 0: 0.510, P: 0.113  
Group # 7 -- 2: 0.542, 4: 0.685, 3: 0.569, 1: 0.248, +: 0.683, 5: 0.838  
Group # 8 -- A: 0.528, B: 0.831, D: 0.264, E: 0.494, F: 0.438, G: 0.311,  
H: 0.401, J: 0.471, K: 0.734, C: 0.778  
Group #10 -- a: 0.611, b: 0.382  
Log likelihood = -839.609 Significance = 0.648

Run # 53, 627 cells:  
Convergence at Iteration 11  
Input 0.093  
Group # 1 -- X: 0.095, o: 0.810, i: 0.387  
Group # 3 -- 0: 0.448, t: 0.997  
Group # 4 -- F: 0.543, M: 0.398, I: 0.366  
Group # 6 -- 0: 0.510, P: 0.114  
Group # 7 -- 2: 0.544, 4: 0.685, 3: 0.569, 1: 0.241, +: 0.692, 5: 0.842  
Group # 8 -- A: 0.531, B: 0.828, D: 0.259, E: 0.497, F: 0.435, G: 0.308,  
H: 0.404, J: 0.469, K: 0.735, C: 0.786  
Group # 9 -- o: 0.480, n: 0.520  
Group #10 -- a: 0.612, b: 0.382  
Log likelihood = -838.923 Significance = 0.207

Add Group # 2 with factors TWM

----- Level # 9 -----

Run # 54, 542 cells:  
No Convergence at Iteration 20  
Input 0.091  
Group # 1 -- X: 0.089, o: 0.765, i: 0.463  
Group # 2 -- T: 0.585, W: 0.380, M: 0.424  
Group # 3 -- 0: 0.449, t: 0.996  
Group # 4 -- F: 0.545, M: 0.387, I: 0.363  
Group # 5 -- 0: 0.495, J: 0.557  
Group # 6 -- 0: 0.510, P: 0.116  
Group # 7 -- 2: 0.536, 4: 0.720, 3: 0.575, 1: 0.253, +: 0.666, 5: 0.816  
Group # 8 -- A: 0.536, B: 0.831, D: 0.263, E: 0.494, F: 0.442, G: 0.309,  
H: 0.394, J: 0.470, K: 0.732, C: 0.776  
Group #10 -- a: 0.608, b: 0.385  
Log likelihood = -836.227 Significance = 0.539



Run # 55, 693 cells:  
No Convergence at Iteration 20  
Input 0.091  
Group # 1 -- X: 0.089, o: 0.764, i: 0.464  
Group # 2 -- T: 0.583, W: 0.386, M: 0.426  
Group # 3 -- O: 0.447, t: 0.997  
Group # 4 -- F: 0.544, M: 0.393, I: 0.366  
Group # 6 -- O: 0.510, P: 0.117  
Group # 7 -- 2: 0.540, 4: 0.719, 3: 0.575, 1: 0.245, +: 0.676, 5: 0.820  
Group # 8 -- A: 0.538, B: 0.828, D: 0.258, E: 0.497, F: 0.439, G: 0.307,  
H: 0.396, J: 0.467, K: 0.734, C: 0.784  
Group # 9 -- o: 0.481, n: 0.519  
Group #10 -- a: 0.609, b: 0.384  
Log likelihood = -835.753 Significance = 0.252

No remaining groups significant

Groups selected while stepping up: 1 3 8 10 4 7 6 2  
Best stepping up run: #51

-----  
Stepping down...

----- Level # 10 -----

Run # 56, 741 cells:  
No Convergence at Iteration 20  
Input 0.091  
Group # 1 -- X: 0.090, o: 0.765, i: 0.462  
Group # 2 -- T: 0.587, W: 0.391, M: 0.422  
Group # 3 -- O: 0.449, t: 0.996  
Group # 4 -- F: 0.544, M: 0.392, I: 0.366  
Group # 5 -- O: 0.495, J: 0.557  
Group # 6 -- O: 0.510, P: 0.119  
Group # 7 -- 2: 0.537, 4: 0.717, 3: 0.572, 1: 0.253, +: 0.672, 5: 0.821  
Group # 8 -- A: 0.537, B: 0.828, D: 0.260, E: 0.498, F: 0.439, G: 0.306,  
H: 0.397, J: 0.470, K: 0.730, C: 0.782  
Group # 9 -- o: 0.482, n: 0.519  
Group #10 -- a: 0.609, b: 0.385  
Log likelihood = -835.561

----- Level # 9 -----

Run # 57, 571 cells:  
Convergence at Iteration 11  
Input 0.115  
Group # 2 -- T: 0.734, W: 0.627, M: 0.274  
Group # 3 -- O: 0.448, t: 0.997  
Group # 4 -- F: 0.540, M: 0.409, I: 0.374  
Group # 5 -- O: 0.496, J: 0.552  
Group # 6 -- O: 0.506, P: 0.228  
Group # 7 -- 2: 0.547, 4: 0.652, 3: 0.592, 1: 0.276, +: 0.317, 5: 0.342  
Group # 8 -- A: 0.532, B: 0.824, D: 0.284, E: 0.492, F: 0.474, G: 0.336,  
H: 0.367, J: 0.469, K: 0.740, C: 0.679  
Group # 9 -- o: 0.473, n: 0.527  
Group #10 -- a: 0.612, b: 0.381  
Log likelihood = -907.228 Significance = 0.000

Run # 58, 680 cells:  
Convergence at Iteration 11  
Input 0.093  
Group # 1 -- X: 0.095, o: 0.813, i: 0.383  
Group # 3 -- o: 0.449, t: 0.996  
Group # 4 -- F: 0.543, M: 0.398, I: 0.366  
Group # 5 -- o: 0.497, J: 0.542  
Group # 6 -- o: 0.510, P: 0.115  
Group # 7 -- 2: 0.543, 4: 0.683, 3: 0.566, 1: 0.247, +: 0.690, 5: 0.843  
Group # 8 -- A: 0.530, B: 0.828, D: 0.261, E: 0.498, F: 0.435, G: 0.308,  
H: 0.405, J: 0.471, K: 0.732, C: 0.785  
Group # 9 -- o: 0.480, n: 0.520  
Group #10 -- a: 0.611, b: 0.382  
Log likelihood = -838.815 Significance = 0.042

Run # 59, 719 cells:  
No Convergence at Iteration 20  
Input 0.092  
Group # 1 -- X: 0.077, o: 0.747, i: 0.506  
Group # 2 -- T: 0.601, W: 0.394, M: 0.409  
Group # 4 -- F: 0.534, M: 0.456, I: 0.386  
Group # 5 -- o: 0.451, J: 0.921  
Group # 6 -- o: 0.509, P: 0.136  
Group # 7 -- 2: 0.539, 4: 0.709, 3: 0.570, 1: 0.249, +: 0.614, 5: 0.849  
Group # 8 -- A: 0.576, B: 0.825, D: 0.262, E: 0.510, F: 0.427, G: 0.293,  
H: 0.420, J: 0.457, K: 0.642, C: 0.762  
Group # 9 -- o: 0.486, n: 0.514  
Group #10 -- a: 0.589, b: 0.406  
Log likelihood = -921.296 Significance = 0.000

Run # 60, 495 cells:  
No Convergence at Iteration 20  
Input 0.093  
Group # 1 -- X: 0.090, o: 0.763, i: 0.466  
Group # 2 -- T: 0.589, W: 0.399, M: 0.420  
Group # 3 -- o: 0.450, t: 0.996  
Group # 5 -- o: 0.495, J: 0.559  
Group # 6 -- o: 0.510, P: 0.113  
Group # 7 -- 2: 0.535, 4: 0.724, 3: 0.569, 1: 0.259, +: 0.655, 5: 0.832  
Group # 8 -- A: 0.544, B: 0.828, D: 0.262, E: 0.511, F: 0.421, G: 0.296,  
H: 0.408, J: 0.469, K: 0.719, C: 0.803  
Group # 9 -- o: 0.474, n: 0.527  
Group #10 -- a: 0.598, b: 0.396  
Log likelihood = -846.214 Significance = 0.000

Run # 61, 693 cells:  
No Convergence at Iteration 20  
Input 0.091  
Group # 1 -- X: 0.089, o: 0.764, i: 0.464  
Group # 2 -- T: 0.583, W: 0.386, M: 0.426  
Group # 3 -- o: 0.447, t: 0.997  
Group # 4 -- F: 0.544, M: 0.393, I: 0.366  
Group # 6 -- o: 0.510, P: 0.117  
Group # 7 -- 2: 0.540, 4: 0.719, 3: 0.575, 1: 0.245, +: 0.676, 5: 0.820  
Group # 8 -- A: 0.538, B: 0.828, D: 0.258, E: 0.497, F: 0.439, G: 0.307,  
H: 0.396, J: 0.467, K: 0.734, C: 0.784  
Group # 9 -- o: 0.481, n: 0.519  
Group #10 -- a: 0.609, b: 0.384  
Log likelihood = -835.753 Significance = 0.546

Run # 62, 732 cells:  
No Convergence at Iteration 20  
Input 0.094  
Group # 1 -- X: 0.103, o: 0.758, i: 0.454  
Group # 2 -- T: 0.602, W: 0.418, M: 0.407  
Group # 3 -- O: 0.449, t: 0.996  
Group # 4 -- F: 0.544, M: 0.394, I: 0.364  
Group # 5 -- O: 0.494, J: 0.576  
Group # 7 -- 2: 0.540, 4: 0.687, 3: 0.561, 1: 0.276, +: 0.479, 5: 0.745  
Group # 8 -- A: 0.534, B: 0.820, D: 0.268, E: 0.498, F: 0.440, G: 0.307,  
H: 0.402, J: 0.479, K: 0.725, C: 0.783  
Group # 9 -- o: 0.480, n: 0.521  
Group #10 -- a: 0.608, b: 0.386  
Log likelihood = -840.100 Significance = 0.005

Run # 63, 476 cells:  
No Convergence at Iteration 20  
Input 0.097  
Group # 1 -- X: 0.124, o: 0.765, i: 0.421  
Group # 2 -- T: 0.640, W: 0.485, M: 0.369  
Group # 3 -- O: 0.449, t: 0.997  
Group # 4 -- F: 0.544, M: 0.385, I: 0.367  
Group # 5 -- O: 0.487, J: 0.658  
Group # 6 -- O: 0.506, P: 0.220  
Group # 8 -- A: 0.514, B: 0.805, D: 0.285, E: 0.511, F: 0.448, G: 0.314,  
H: 0.412, J: 0.494, K: 0.706, C: 0.802  
Group # 9 -- o: 0.483, n: 0.518  
Group #10 -- a: 0.607, b: 0.387  
Log likelihood = -847.786 Significance = 0.000

Run # 64, 257 cells:  
No Convergence at Iteration 20  
Input 0.104  
Group # 1 -- X: 0.115, o: 0.753, i: 0.448  
Group # 2 -- T: 0.616, W: 0.455, M: 0.392  
Group # 3 -- O: 0.450, t: 0.996  
Group # 4 -- F: 0.549, M: 0.430, I: 0.339  
Group # 5 -- O: 0.492, J: 0.598  
Group # 6 -- O: 0.508, P: 0.168  
Group # 7 -- 2: 0.545, 4: 0.665, 3: 0.465, 1: 0.358, +: 0.588, 5: 0.805  
Group # 9 -- o: 0.478, n: 0.523  
Group #10 -- a: 0.618, b: 0.375  
Log likelihood = -902.420 Significance = 0.000

Run # 65, 542 cells:  
No Convergence at Iteration 20  
Input 0.091  
Group # 1 -- X: 0.089, o: 0.765, i: 0.463  
Group # 2 -- T: 0.585, W: 0.380, M: 0.424  
Group # 3 -- O: 0.449, t: 0.996  
Group # 4 -- F: 0.545, M: 0.387, I: 0.363  
Group # 5 -- O: 0.495, J: 0.557  
Group # 6 -- O: 0.510, P: 0.116  
Group # 7 -- 2: 0.536, 4: 0.720, 3: 0.575, 1: 0.253, +: 0.666, 5: 0.816  
Group # 8 -- A: 0.536, B: 0.831, D: 0.263, E: 0.494, F: 0.442, G: 0.309,  
H: 0.394, J: 0.470, K: 0.732, C: 0.776  
Group #10 -- a: 0.608, b: 0.385  
Log likelihood = -836.227 Significance = 0.254

Run # 66, 545 cells:  
 No Convergence at Iteration 20  
 Input 0.096  
 Group # 1 -- X: 0.086, o: 0.761, i: 0.473  
 Group # 2 -- T: 0.600, W: 0.359, M: 0.411  
 Group # 3 -- O: 0.452, t: 0.995  
 Group # 4 -- F: 0.535, M: 0.396, I: 0.398  
 Group # 5 -- O: 0.492, J: 0.604  
 Group # 6 -- O: 0.509, P: 0.129  
 Group # 7 -- 2: 0.530, 4: 0.703, 3: 0.572, 1: 0.269, +: 0.625, 5: 0.838  
 Group # 8 -- A: 0.552, B: 0.830, D: 0.276, E: 0.487, F: 0.411, G: 0.301,  
 H: 0.406, J: 0.462, K: 0.718, C: 0.816  
 Group # 9 -- o: 0.483, n: 0.517  
 Log likelihood = -860.144 Significance = 0.000

Cut Group # 5 with factors OJ

----- Level # 8 -----

Run # 67, 528 cells:  
 Convergence at Iteration 10  
 Input 0.115  
 Group # 2 -- T: 0.729, W: 0.621, M: 0.279  
 Group # 3 -- O: 0.446, t: 0.997  
 Group # 4 -- F: 0.540, M: 0.409, I: 0.374  
 Group # 6 -- O: 0.506, P: 0.224  
 Group # 7 -- 2: 0.549, 4: 0.654, 3: 0.595, 1: 0.268, +: 0.320, 5: 0.343  
 Group # 8 -- A: 0.532, B: 0.824, D: 0.282, E: 0.492, F: 0.474, G: 0.336,  
 H: 0.366, J: 0.467, K: 0.743, C: 0.679  
 Group # 9 -- o: 0.473, n: 0.527  
 Group #10 -- a: 0.613, b: 0.381  
 Log likelihood = -907.404 Significance = 0.000

Run # 68, 627 cells:  
 Convergence at Iteration 11  
 Input 0.093  
 Group # 1 -- X: 0.095, o: 0.810, i: 0.387  
 Group # 3 -- O: 0.448, t: 0.997  
 Group # 4 -- F: 0.543, M: 0.398, I: 0.366  
 Group # 6 -- O: 0.510, P: 0.114  
 Group # 7 -- 2: 0.544, 4: 0.685, 3: 0.569, 1: 0.241, +: 0.692, 5: 0.842  
 Group # 8 -- A: 0.531, B: 0.828, D: 0.259, E: 0.497, F: 0.435, G: 0.308,  
 H: 0.404, J: 0.469, K: 0.735, C: 0.786  
 Group # 9 -- o: 0.480, n: 0.520  
 Group #10 -- a: 0.612, b: 0.382  
 Log likelihood = -838.923 Significance = 0.044

Run # 69, 653 cells:  
 Convergence at Iteration 12  
 Input 0.105  
 Group # 1 -- X: 0.079, o: 0.690, i: 0.576  
 Group # 2 -- T: 0.525, W: 0.312, M: 0.484  
 Group # 4 -- F: 0.532, M: 0.495, I: 0.384  
 Group # 6 -- O: 0.510, P: 0.103  
 Group # 7 -- 2: 0.592, 4: 0.747, 3: 0.638, 1: 0.111, +: 0.711, 5: 0.839  
 Group # 8 -- A: 0.615, B: 0.832, D: 0.200, E: 0.508, F: 0.413, G: 0.303,  
 H: 0.399, J: 0.380, K: 0.688, C: 0.755  
 Group # 9 -- o: 0.492, n: 0.508  
 Group #10 -- a: 0.597, b: 0.397  
 Log likelihood = -987.571 Significance = 0.000

Run # 70, 454 cells:  
No Convergence at Iteration 20  
Input 0.093  
Group # 1 -- X: 0.089, o: 0.761, i: 0.468  
Group # 2 -- T: 0.585, W: 0.394, M: 0.424  
Group # 3 -- O: 0.448, t: 0.997  
Group # 6 -- O: 0.510, P: 0.112  
Group # 7 -- 2: 0.537, 4: 0.726, 3: 0.573, 1: 0.250, +: 0.659, 5: 0.832  
Group # 8 -- A: 0.545, B: 0.828, D: 0.260, E: 0.511, F: 0.420, G: 0.296,  
H: 0.406, J: 0.466, K: 0.722, C: 0.804  
Group # 9 -- o: 0.474, n: 0.527  
Group #10 -- a: 0.599, b: 0.395  
Log likelihood = -846.423 Significance = 0.000

Run # 71, 683 cells:  
No Convergence at Iteration 20  
Input 0.094  
Group # 1 -- X: 0.103, o: 0.757, i: 0.457  
Group # 2 -- T: 0.596, W: 0.412, M: 0.412  
Group # 3 -- O: 0.447, t: 0.997  
Group # 4 -- F: 0.544, M: 0.394, I: 0.364  
Group # 7 -- 2: 0.543, 4: 0.689, 3: 0.565, 1: 0.265, +: 0.486, 5: 0.744  
Group # 8 -- A: 0.535, B: 0.820, D: 0.265, E: 0.497, F: 0.439, G: 0.308,  
H: 0.400, J: 0.476, K: 0.729, C: 0.785  
Group # 9 -- o: 0.479, n: 0.521  
Group #10 -- a: 0.608, b: 0.385  
Log likelihood = -840.448 Significance = 0.004

Run # 72, 421 cells:  
No Convergence at Iteration 20  
Input 0.098  
Group # 1 -- X: 0.124, o: 0.765, i: 0.421  
Group # 2 -- T: 0.631, W: 0.475, M: 0.378  
Group # 3 -- O: 0.443, t: 0.998  
Group # 4 -- F: 0.544, M: 0.384, I: 0.367  
Group # 6 -- O: 0.506, P: 0.217  
Group # 8 -- A: 0.514, B: 0.803, D: 0.281, E: 0.510, F: 0.448, G: 0.316,  
H: 0.410, J: 0.491, K: 0.711, C: 0.804  
Group # 9 -- o: 0.483, n: 0.518  
Group #10 -- a: 0.608, b: 0.385  
Log likelihood = -849.422 Significance = 0.000

Run # 73, 226 cells:  
No Convergence at Iteration 20  
Input 0.105  
Group # 1 -- X: 0.115, o: 0.751, i: 0.451  
Group # 2 -- T: 0.610, W: 0.448, M: 0.398  
Group # 3 -- O: 0.447, t: 0.997  
Group # 4 -- F: 0.549, M: 0.431, I: 0.339  
Group # 6 -- O: 0.508, P: 0.165  
Group # 7 -- 2: 0.549, 4: 0.667, 3: 0.470, 1: 0.344, +: 0.592, 5: 0.804  
Group # 9 -- o: 0.478, n: 0.523  
Group #10 -- a: 0.619, b: 0.374  
Log likelihood = -903.029 Significance = 0.000

Run # 74, 495 cells:  
No Convergence at Iteration 20  
Input 0.092  
Group # 1 -- X: 0.089, o: 0.764, i: 0.464  
Group # 2 -- T: 0.581, W: 0.376, M: 0.429  
Group # 3 -- O: 0.447, t: 0.997  
Group # 4 -- F: 0.545, M: 0.387, I: 0.363  
Group # 6 -- O: 0.510, P: 0.115  
Group # 7 -- 2: 0.539, 4: 0.722, 3: 0.578, 1: 0.245, +: 0.670, 5: 0.816  
Group # 8 -- A: 0.536, B: 0.831, D: 0.261, E: 0.493, F: 0.441, G: 0.309,  
H: 0.393, J: 0.467, K: 0.736, C: 0.778  
Group #10 -- a: 0.609, b: 0.385  
Log likelihood = -836.425 Significance = 0.252

Run # 75, 498 cells:  
No Convergence at Iteration 20  
Input 0.097  
Group # 1 -- X: 0.085, o: 0.759, i: 0.477  
Group # 2 -- T: 0.593, W: 0.350, M: 0.418  
Group # 3 -- O: 0.449, t: 0.997  
Group # 4 -- F: 0.535, M: 0.396, I: 0.398  
Group # 6 -- O: 0.509, P: 0.126  
Group # 7 -- 2: 0.534, 4: 0.706, 3: 0.577, 1: 0.255, +: 0.633, 5: 0.837  
Group # 8 -- A: 0.553, B: 0.830, D: 0.273, E: 0.486, F: 0.410, G: 0.302,  
H: 0.404, J: 0.457, K: 0.724, C: 0.818  
Group # 9 -- o: 0.483, n: 0.517  
Log likelihood = -860.808 Significance = 0.000

Cut Group # 9 with factors on

----- Level # 7 -----

Run # 76, 362 cells:  
Convergence at Iteration 10  
Input 0.115  
Group # 2 -- T: 0.726, W: 0.608, M: 0.282  
Group # 3 -- O: 0.446, t: 0.997  
Group # 4 -- F: 0.542, M: 0.402, I: 0.370  
Group # 6 -- O: 0.506, P: 0.215  
Group # 7 -- 2: 0.548, 4: 0.654, 3: 0.598, 1: 0.269, +: 0.308, 5: 0.332  
Group # 8 -- A: 0.529, B: 0.829, D: 0.284, E: 0.487, F: 0.477, G: 0.342,  
H: 0.360, J: 0.466, K: 0.746, C: 0.669  
Group #10 -- a: 0.613, b: 0.381  
Log likelihood = -908.920 Significance = 0.000

Run # 77, 436 cells:  
Convergence at Iteration 11  
Input 0.093  
Group # 1 -- X: 0.095, o: 0.809, i: 0.390  
Group # 3 -- O: 0.448, t: 0.997  
Group # 4 -- F: 0.545, M: 0.392, I: 0.363  
Group # 6 -- O: 0.510, P: 0.112  
Group # 7 -- 2: 0.543, 4: 0.686, 3: 0.571, 1: 0.242, +: 0.685, 5: 0.837  
Group # 8 -- A: 0.529, B: 0.831, D: 0.262, E: 0.493, F: 0.438, G: 0.311,  
H: 0.400, J: 0.469, K: 0.737, C: 0.779  
Group #10 -- a: 0.612, b: 0.382  
Log likelihood = -839.723 Significance = 0.040

Run # 78, 462 cells:  
Convergence at Iteration 11  
Input 0.105  
Group # 1 -- X: 0.078, o: 0.690, i: 0.577  
Group # 2 -- T: 0.525, W: 0.309, M: 0.484  
Group # 4 -- F: 0.532, M: 0.492, I: 0.383  
Group # 6 -- O: 0.510, P: 0.102  
Group # 7 -- 2: 0.592, 4: 0.747, 3: 0.639, 1: 0.111, +: 0.708, 5: 0.838  
Group # 8 -- A: 0.614, B: 0.833, D: 0.201, E: 0.507, F: 0.414, G: 0.304,  
H: 0.397, J: 0.381, K: 0.689, C: 0.752  
Group #10 -- a: 0.597, b: 0.397  
Log likelihood = -987.725 Significance = 0.000

Run # 79, 309 cells:  
No Convergence at Iteration 20  
Input 0.094  
Group # 1 -- X: 0.089, o: 0.761, i: 0.469  
Group # 2 -- T: 0.581, W: 0.381, M: 0.428  
Group # 3 -- O: 0.448, t: 0.997  
Group # 6 -- O: 0.510, P: 0.108  
Group # 7 -- 2: 0.535, 4: 0.730, 3: 0.577, 1: 0.250, +: 0.649, 5: 0.826  
Group # 8 -- A: 0.543, B: 0.832, D: 0.261, E: 0.505, F: 0.424, G: 0.300,  
H: 0.401, J: 0.465, K: 0.725, C: 0.796  
Group #10 -- a: 0.598, b: 0.396  
Log likelihood = -847.765 Significance = 0.000

Run # 80, 481 cells:  
No Convergence at Iteration 20  
Input 0.094  
Group # 1 -- X: 0.102, o: 0.757, i: 0.457  
Group # 2 -- T: 0.594, W: 0.401, M: 0.415  
Group # 3 -- O: 0.447, t: 0.997  
Group # 4 -- F: 0.545, M: 0.388, I: 0.361  
Group # 7 -- 2: 0.542, 4: 0.690, 3: 0.568, 1: 0.265, +: 0.481, 5: 0.736  
Group # 8 -- A: 0.534, B: 0.824, D: 0.268, E: 0.493, F: 0.442, G: 0.311,  
H: 0.396, J: 0.476, K: 0.731, C: 0.778  
Group #10 -- a: 0.608, b: 0.386  
Log likelihood = -841.281 Significance = 0.003

Run # 81, 277 cells:  
No Convergence at Iteration 20  
Input 0.098  
Group # 1 -- X: 0.123, o: 0.765, i: 0.421  
Group # 2 -- T: 0.628, W: 0.467, M: 0.381  
Group # 3 -- O: 0.443, t: 0.998  
Group # 4 -- F: 0.545, M: 0.379, I: 0.365  
Group # 6 -- O: 0.506, P: 0.210  
Group # 8 -- A: 0.512, B: 0.805, D: 0.283, E: 0.505, F: 0.450, G: 0.319,  
H: 0.407, J: 0.491, K: 0.713, C: 0.799  
Group #10 -- a: 0.608, b: 0.386  
Log likelihood = -850.043 Significance = 0.000

Run # 82, 145 cells:  
No Convergence at Iteration 20  
Input 0.105  
Group # 1 -- X: 0.113, o: 0.751, i: 0.452  
Group # 2 -- T: 0.606, W: 0.433, M: 0.402  
Group # 3 -- 0: 0.447, t: 0.997  
Group # 4 -- F: 0.550, M: 0.427, I: 0.336  
Group # 6 -- 0: 0.508, P: 0.161  
Group # 7 -- 2: 0.548, 4: 0.669, 3: 0.474, 1: 0.343, +: 0.580, 5: 0.798  
Group #10 -- a: 0.618, b: 0.375  
Log likelihood = -904.151 Significance = 0.000

Run # 83, 349 cells:  
No Convergence at Iteration 20  
Input 0.097  
Group # 1 -- X: 0.085, o: 0.759, i: 0.477  
Group # 2 -- T: 0.591, W: 0.342, M: 0.421  
Group # 3 -- 0: 0.449, t: 0.997  
Group # 4 -- F: 0.536, M: 0.391, I: 0.395  
Group # 6 -- 0: 0.509, P: 0.123  
Group # 7 -- 2: 0.533, 4: 0.709, 3: 0.580, 1: 0.254, +: 0.629, 5: 0.834  
Group # 8 -- A: 0.552, B: 0.833, D: 0.275, E: 0.482, F: 0.413, G: 0.304,  
H: 0.400, J: 0.457, K: 0.726, C: 0.813  
Log likelihood = -861.373 Significance = 0.000

All remaining groups significant

Groups eliminated while stepping down: 5 9  
Best stepping up run: #51  
Best stepping down run: #74