

共起語情報を用いた新聞見出しに現れる略語の考察

赤間啓之 (東京工業大学)、清水由美子 (武蔵工業大学・東京工業大学)

1 はじめに

1.1 目的

ある語が多用され、社会に十分浸透すると、しばしばその語の略語が使用され始める。略語は単にもとの語を省略したというだけでなく、その用いられ方に独自の特徴を持つ。朝日新聞の見出しに使用される「インターネット」「ネット」を例に、語とその略語との関係を考える。また、社会におけるインターネットの位置の変化を、見出しがどのように反映させているのかについても考察する。さらに、こうした考察において、構文論的観点だけでなく、共起語情報から得られる知見をプラスすることの有効性を立証することを主な目的とする。

1.2 調査対象

調査には朝日新聞社の Web サイト、asahi.com のデータベースを利用した。対象期間は 1994 年 1 月～2000 年 8 月である。「インターネット」を含む見出しはこれ以前にも見られるが、「ネット」を含む見出しも共に現れる 1994 年からの分析の対象にする。「ネット」を含む見出しは対象期間中出现する 6225 件のうち、「ネット」が「インターネット」を意味すると認められる 1856 件とした。

表1 「インターネット」を含む見出しと「ネット」を含む見出し、年別件数

| | インターネット | ネット |
|--------------|---------|------|
| 1994 | 36 | 4 |
| 1995 | 285 | 6 |
| 1996 | 646 | 34 |
| 1997 | 592 | 146 |
| 1998 | 380 | 253 |
| 1999 | 404 | 634 |
| 2000 (8月末現在) | 335 | 779 |
| 計 | 2679 | 1856 |

1997年までは「インターネット」を用いた見出しが圧倒的に多いが、1998年から「ネット」を用いた見出しの割合が増加し、翌1999年には両者の関係が逆転する。

2 「インターネット」を含む見出し、「ネット」を含む見出し

2.1 キーワードの見出し冒頭からの位置

各キーワードの全体を 100 とした時の、それぞれのキーワードの見出し冒頭からの位置の占める割合、キーワード毎の「年別」見出し冒頭からの位置の割合を下記に示す。

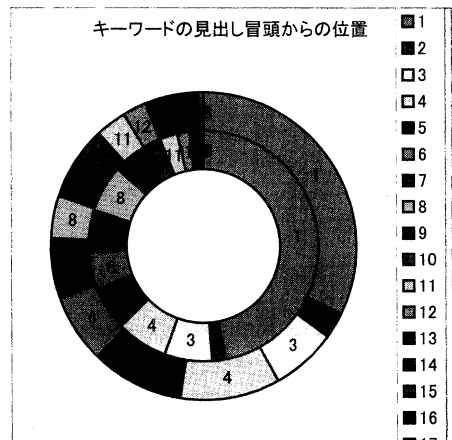


図1 各キーワードの全体を 100 とした時の、各見出し冒頭からの位置の占める割合 (内側「インターネット」、外側「ネット」)

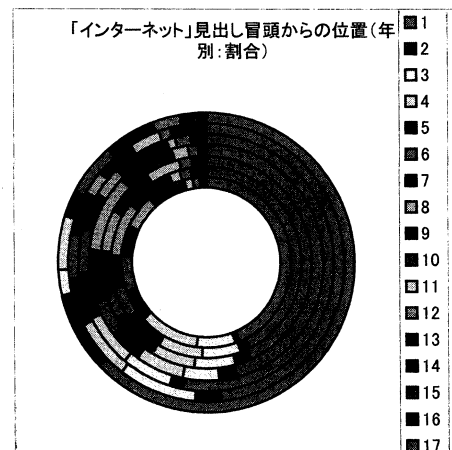


図2 「インターネット」の見出し冒頭からの位置(年別割合: 外側から 1994 年～2000 年)

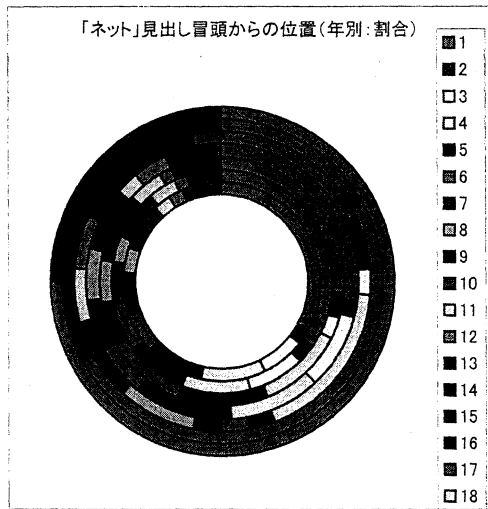


図3 「ネット」見出し冒頭からの位置(年別割合) : 外側から1994年～2000年)

キーワード双方ともに、見出し冒頭に置かれる割合が他の位置に置かれる割合に比べ最も高いが、「インターネット」が冒頭に位置する割合の方が「ネット」のそれよりも高い。「ネット」では冒頭から4語目、5語目に位置する割合の高さも冒頭に続いて目立つ。

各キーワードの年別割合を見ると、「インターネット」が冒頭に位置する割合が年を追って徐々に減っていること、逆に「ネット」の冒頭に位置する割合が段々が増えていくことが分かる。

2.2 共起頻度の高い語

年度毎にタグ付けした各キーワードを変数に、またそれらのキーワードと見出し内で共起する語をサンプルにし、年度ごとの共起語-サンプルの出現頻度(年毎の延べ語数に対する割合)をデータに取った行列を因子分析にかけた。因子分析には主成分分解を用い、固有値1以上という基準で因子を3個抽出し、この結果にバリマックス回転を施した。因子分析の結果抽出された3個の因子の寄与率はそれぞれ48.9%、10.2%、7.2%であり、また3番目までの因子の累積寄与率は66.4%であった。因子負荷行列を見ると、第1因子は変数のうちインターネット1997年、1998年、1999年、2000年の因子負荷が大きいことから、「インターネット中・後期」に関わる因子であると考えられる(「インターネット中・後期因子」と命名する)。第2因子はネット1996年、1997

年、1998年、1999年、2000年の因子負荷が大きいことから、「ネット全体」に関わる因子だと見られる(「ネット因子」とする)。また、第3因子はインターネット1994年、1995年、1996年の因子負荷が大きく、これは「インターネット前期」に関わる因子であると考えられる(この因子を「インターネット前期因子」と命名する)。この結果から「ネット」と「インターネット」、また「インターネット」1994年～1996年と「インターネット」1997年～2000年との間に用法の違いがあるのではないかと推測される。

また各因子の因子得点が高い共起語の上位100語(自立語のみ)は表2のようになった(ゴシック体はマイナスイメージを伴う語)。これらの共起語はそれぞれの因子を大きく含んでいると言える。

各上位100語中、三因子に共通する共起語は「サービス・会社・開設・活用・企業・実験・紹介・情報・接続・通信・電子・発信・利用」の13語であり、「インターネット中・後期因子」と「ネット因子」、「インターネット中・後期因子」と「インターネット前期因子」、「ネット因子」と「インターネット前期因子」とのそれぞれの共通共起語数は31語、28語、25語であった。

また、各上位100語中に含まれるマイナスイメージを伴う語に着目した。「インターネット中・後期因子」の高因子得点上位100語では“逮捕”“容疑”“詐欺”など、11語がマイナスイメージを伴うと認められる。「ネット因子」でも同じく11語がマイナスイメージを伴う語として挙げられるが、「インターネット前期因子」ではこれが“ポルノ”1語である点の特徴的である。

これらから、インターネット前期とインターネット中・後期とは、かなり異なる文脈の中でキーワードが用いられていることが伺える。また、それに比べ、インターネット中・後期とネットの用法の差はさほど大きくないのではないかと推測される。

2.3 キーワードと共起語の前後関係

共起語を・前(キーワードより前に来る)、・後(キーワードの後ろに来る)にわけ、・前・後共に共起頻度5以上の語について分析する。・前・後共に頻度5以上の共起語は58語あった。

キーワードに対する共起語出現数の・前・後比の違いにより、この58語は下記の4種類に分類される。

パターン1: 共起語の出現数が「インターネット」「ネ

表2 各因子に対する因子得点の高い語（上位100語）

| | |
|--------------------------------|---|
| <p>因子1 (インターネット中・後期因子)</p> | <p>情報・逮捕・サービス・接続・利用・公開・販売・容疑・NTT・詐欺・中継・電話・ホームページ・使う・調査・県警・会社・提供・開設・速報・放送・設立・無料・取引・テレビ・事業・開始・中国・女性・紹介・活用・学校・電子・通信・CATV・来月・教育・検索・配信・県内・定額・画像・授業・消費・実験・ケーブル・図書館・普及・男・東京・研究・体験・地裁・急増・発信・トラブル・購入・通販・アクセス・相談・活動・携帯・起訴・業者・講演・京都・障害・商議・番組・被告・回線・銀行・若者・個人・TV・全国・不正・教室・悪用・わいせつ・県立・参入・HP・規制・支店・対象・会議・地域・生活・児童・就職・男性・技術・衛星・議事・PR・ネット・知事・きょう・企業</p> |
| <p>因子2 (ネット因子)</p> | <p>容疑・情報・逮捕・販売・会社・パソコン・社会・開設・公開・県警・取引・議論・交流・企業・事業・証券・時代・わいせつ・利用・犯罪・画像・提携・HP・広告・配信・ソフト・設立・若い・詐欺・発注・発信・ルール・使う・サービス・聴取・必要・オンライン・視野・著作・募集・ビジョン・接続・国際・成功・印刷・市場・紹介・新聞・生む・ソニー・業者・年賀状・銀行・事件・倫理・シリコンバレー・マルチメディア・日本・大阪・マイクロソフト・対策・米国・音楽・社長・警察庁・活用・金融・電子・メディア・ソフトバンク・送検・男性・仮想・対応・流す・売る・システム・ポルノ・通信・大手・探す・市民・福岡・業界・広島・摘発・メール・支援・ビジネス・商店・海外・OK・実験・街・電脳・ハイテク・個人・家電・開局・口座</p> |
| <p>因子3 (インターネット前期因子)</p> | <p>情報・世界・提供・開始・参加・接続・発信・ホームページ・記念・官邸・瞬時・乗せる・宣言・人間・式典・関係・熱気・夫婦・仲間・開設・電子・パソコン・富士通・実験・ソフト・メール・会議・通信・日本・研修・大学・企業・国境・利用・紹介・募集・発足・ネット・通じる・越える・福岡・協会・観光・国際・活用・体験・料金・計画・商用・会社・来月・続々・参入・KDD・学校・コンピューター・入門・交換・サービス・教育・結ぶ・就職・開発・検索・採用・教室・次々・PR・機関・パスワード・リクルー ト・ポルノ・反響・中継・TI・プロジェクト・経営・登場・市民・大阪・仮想・発売・記事・メディア・大統領・日本語・規模・三重大・通産省・禁止・ホワイトハウス・学研・文部省・エイズ・新車・特許庁・費用・地球・施設・広報</p> |

ット」共にキーワードの「後」より「前」の方が多し(データ・県内・全・電子・電脳 計5語)。

パターン2: 共起語の出現数が「インターネット」ではキーワードの「前」が多く、「ネット」では「後」が多い(メール・海外・中・投票・年 計5語)。

パターン3: 共起語の出現数が「インターネット」では「後」が多く、「ネット」では「前」が多い(テレビ・パソコン・ポルノ・わいせつ・画像・情報・電話 計7語)。

パターン4: 共起語の出現数が「インターネット」「ネット」共にキーワードの「前」より「後」の方が多し(1・2・3・NTT・サービス・する・ソフトバンク・円・音楽・化・

会・会社・学校・企業・銀行・研究・県・公開・向け・使う・市・社・者・女性・新人・世界・逮捕・通信・提携・日・日本・発信・販売・米・米国・募集・本・容疑・利用 計41語)。

58語すべてについて χ^2 乗検定を行い、 χ^2 乗値3以上の語(自立語のみ)に着目した。該当したのは次の7語である[パソコン・ポルノ・メール・わいせつ・画像・情報・電話]。これらのうち「メール」のみ、パターン2に分類される。これは「インターネット」と共起する「メール」の用例24件中10件がシリーズ名の一部(例:電子メール:森井教授のインターネット講座第64回)であるためであり、その他の χ^2 乗値3以上の語はすべてパターン3に属する。

さらにこれら7語の中でも特に高い χ^2 乗値を出し、

共起頻度も高い「情報」について、詳細に検討した。文脈を描えるために、1994年当初からよく見られるパターンである「インターネットで～情報」「ネットで～情報」「～情報インターネットで」「～情報ネットで」に該当するものについて、それらの比率の変化を見た。

表3 文脈を描えた対比：数字はそれぞれのパターンに該当する用例件数（括弧内はキーワードが見出し冒頭に位置しない用例件数）

| | インター ネットで ～情報 | ～情報イ ンターネ ットで | ネット で ～情報 | ～情報 ネットで |
|------|---------------------|---------------------|-----------------|-------------|
| 1994 | 3 | 0 | 0 | 0 |
| 1995 | 27(13) | 6 | 0 | 0 |
| 1996 | 38(8) | 13 | 1(1) | 0 |
| 1997 | 25(7) | 11 | 0 | 5 |
| 1998 | 13(3) | 5 | 4(2) | 15 |
| 1999 | 10 | 9 | 8(4) | 16 |
| 2000 | 4(1) | 6 | 8(5) | 17 |

年を経るに従って、「インターネットで～情報」対「～情報インターネットで」の比率が変化していく様が見て取れる。2000年には「身近な生活情報をインターネットで配信」のような、「情報」の“中身”に焦点を置いた見出しの方が、冒頭に「インターネット」を置くものより多くなっている。「～情報ネットで」のパターンの増加とも併せ、話題の焦点が「インターネット」という“媒体”からその媒体が運ぶ“中身”に徐々に移行していると言えよう。

3 おわりに

朝日新聞の記事では「インターネット」という語を、1996年初め頃には社会にかなり浸透したものとして扱っている[清水]。ここでは、それと共に出現の度合いを強めた「ネット」と「インターネット」との関係を探った。

分析の観点は以下の3点であった。

- 1) キーワードの見出し冒頭からの位置
 - 2) 共起頻度の高い語
 - 3) キーワードと共起語の前後関係
- 「ネット」に比べ「インターネット」はより目立つ位置

に置かれる（話題の焦点として提出されることが多い）ことが多かったが、その割合は年々減少している。「ネット」の見出し冒頭の位置を占める割合が年毎に増加しつつあることから、両者の違いが縮まっていると言えそうである。

“共起頻度の高い語”の調査では、「インターネット」の初期の用法と後期の用法とでかなりの違いがあることが分かった。また、「インターネット」の後期の用法と「ネット」の用法との違いがさほど大きくないのではないかと予測した。これが「インターネット」の言葉としての「格」が下落し、略語である「ネット」の用法に近づいていく過程なのかどうか（もしそうであるならば、非常に短期間における「言葉の格」変化であり、その背景には想像以上に速かった「インターネット」の普及状況が反映されていることになろう）等、今後の詳細な検討による。

「パソコン、ポルノ、メール、わいせつ、画像、情報、電話」といった語は、「インターネット」と「ネット」とで、その共起の仕方を変えることで両キーワードのふるまいの違いを際立たせていた。ただし、これも詳細に検討すると、年を追うごとに両者のふるまいの差は小さくなる傾向にある。

キーワードの見出し冒頭からの位置のような構文的特徴から見えてくるものもあるが、見え方が大まかであり、共起語情報をプラスすることでより詳細な分析が可能になる。今後、共起語情報を用いた更に多様な観点からの分析を進めたい。また、複合語を作れることは略語から一般語移行への一つの指標と認められる[上野]。「ネット」が作る複合語と「インターネット」が作る複合語の違いなどの観点からの分析も今後の課題である。

参考文献

- 1 清水由美子、2000.2、「インターネット」ということば、武蔵工業大学環境情報学部紀要創刊号
- 2 上野 力、1995.11、ワードプロセッサからワープロへー言葉の変化試論一、常葉国文（常葉学園短期大学）20