

定量的分析に基づく「が/の」交替再考

南部智史

(神戸松蔭女子学院大学大学院)

キーワード：「が/の」交替，定量的分析，言語変化，言語変異

1. はじめに*

本稿では「が/の」交替 (GA/NO Conversion, 以降 GNC) について, コーパスを用いた定量的観点から分析を行う. GNC とは, (1) のように名詞修飾節等の従属節内で主格を標示する格助詞「が」と「の」が交替可能な変異現象のことである.

- (1) a. [太郎 が/の 書いた] 本
- b. [窓 が/の 小さい] 家
- c. [交替 が/の 可能な] 環境
- d. [花子 が/の 学生だった] 頃

GNC の統語的研究は, Harada (1971) を初めとして数多くなされており (Shibatani 1975, 井上 1976, Harada 1976, Nakai 1980, Miyagawa 1993, Ura 1993, Watanabe 1996, 西岡 1998, Ochi 2001, 菊田 2002, Hiraiwa 2002), その中で GNC に影響を与える要因, つまり「が」か「の」のどちらが現れるかに影響する要因として隣接性や他動性制約などが提案されてきた. また, 数は少ないが定量的観点に基づいた研究も行われている. 例えば, 小説をコーパスとして用いた Horie and Kang (2000) や金 (2002) では GNC と状態性の関係が指摘されている. さらに, 小熊 (2004) による認知言語学的観点からの研究や, Nakajima (2003) の言語獲得についての研究

もあり，GNC に関する研究が幅広く行われていることがわかる．

しかし，先行研究では扱われていない事柄がいくつかある．まず，Harada (1971) による言語変化仮説の検証である．先に挙げた Harada (1971) は統語的分析の他に，GNC には現在進行中の言語変化が存在するという仮説を立てていた．後で詳述するが，その仮説は名詞修飾節内の主格標示において「の」の使用から「が」の使用へ移行する変化が起きているというものであったが，これまでこの仮説が検証されることはなかった．また，先行研究では GNC に影響を与える要因が複数提案されていたが，それらは包括的に扱われてこなかった．このことは，1 つの要因の影響のみ考慮した分析だと，たとえその要因の GNC に対する影響がデータから示されたとしても，他の要因の影響を排除していない限り，その要因自体の影響の強さが示されたことにならないという問題を抱えている．

以上の GNC における言語変化と GNC に影響を与える要因という 2 点について，本稿ではコーパスを用いた定量的分析によって明らかにしたい．言語変化については，Harada (1971) の仮説にある変化が実際に起きているかどうかを，経年的データによって検証する．2 章で詳述するが，本稿では長期的発話データを持つコーパスとして国会会議録を用いることとした．GNC に影響を与える要因は先行研究での議論を参考にして想定し，それらの影響について定量的に検証を行う．先行研究で提案された要因の検証は，これまでも Maki et al. (2004) によって GNC の容認度を尋ねるアンケート調査という形で行われている．一方，本稿ではコーパス中のデータに基づいて検証を行うわけだが，どの要因がどの程度 GNC に影響を及ぼすのかを提示する方法には，クロス表とロジスティック回帰を用いることとした．

クロス表分析では，コーパス中の GNC に相当する「が」と「の」の生起率を用いることで，要因の GNC に対する影響について考察する．ロジスティック回帰は，ある要因が GNC に与える影響度を経験的に数値として特定することができる．さらに，他の要因の影響を排除した個々の要因独自の影響度を算出できるため，その値から要

因間で GNC に対する影響の強さを比較することも可能である。

ロジスティック回帰を用いた研究は、言語変異分析のためにロジスティック回帰を含む統計ソフトとして開発された VARBRUL が存在することからわかる通り数多くあり、近年では Bresnan et al. (forthcoming) の英語の与格交替の分析にもロジスティック回帰が用いられている。Bresnan et al. (forthcoming) は、コーパスに基づく定量的分析によって、与格交替に影響を与える複数の要因を特定できると主張している。本稿でも同様に、定量的観点から GNC に影響を与える要因の特定を行う。

本稿の構成は以下の通りである。2 章では本稿で用いるコーパスと分析方法について述べる。3 章ではまず言語変化の仮説の検証を行い、その後で想定した要因ごとにクロス表を用いて GNC に影響を与えるかどうかを考察する。4 章ではロジスティック回帰分析を行い、要因ごとの具体的な影響度の算出と要因間で影響度の比較を行う。

2. 方法論

2.1 国会会議録

本稿ではコーパスとして国会会議録を用いる。国会会議録は国会での議員らの発言を記録した会議録であり、ウェブサイト「国会会議録検索システム」(<http://kokkai.ndl.go.jp>)としてインターネット上に無料で公開されている。国会会議録は言語資料として近年注目され始めており、語彙(荒尾 2005)、統語(Chen 2002, Sano 2005)、談話(薄井 2005)などの多岐にわたって国会会議録を用いた研究が行われている。また、松田ほか(2005)では整文の解明という観点から国会会議録の特性を明らかにする試みがなされており、言語資料として国会会議録が持つ特徴が今後さらに明らかになるとと思われる。

本稿で国会会議録を採用した最も大きな理由は、その発話収録期間の長さである。国会会議録は、1947 年第一回国会以降およそ 60 年にわたって議員らの発話を記録し続けているという、コーパスとして非常に貴重な特徴を持っている(松田 2004)。

このことは、異なる世代の発話を比較できるという点で Harada (1971) の言語変化仮説を検証するためには欠かせない特徴である。

2.2 分析対象

本稿では分析対象として、国会会議録に収録されている東京方言話者である国会議員の発話を選択した。その理由は、本稿で検証を試みる Harada (1971) の言語変化仮説が東京方言話者のデータに基づいて立てられた仮説だったからである。また、これは GNC における方言間の差異を考慮する点においても重要である¹⁾。

東京方言を話す国会議員の定義は、東京出身の国会議員とし、東京出身の国会議員を選択するための参考資料として「貴族院 / 参議院議員名鑑 (衆議院・参議院 1990a)」、「衆議院議員名鑑」(衆議院・参議院 1990b)、「政治家人名辞典」(日外アソシエーツ 2003)、「現代政治家人名辞典」(日外アソシエーツ 1999)、「国会便覧」(日本政経新聞社 1998, 1999, 2000, 2001, 2002, 2003, 2004)を用いた。これら資料から東京出身の国会議員 180 名を選択し、その 180 名から彼らの生年ごとに 1 名ずつ抽出し、76 名を得た²⁾。生年を用いた国会議員の選択は、通時的に均質なデータに基づいて言語変化を考察するために行った。最終的に得られた 76 名は、彼らの生年で考えると 1876 年～1970 年とおよそ 100 年に及ぶこととなり、Harada(1971) の仮説を検証するための十分なデータを得ることができると考えられる。

以上の方法で得られた 76 名の発話データは、国会会議録の公式サイトから議員ごとにテキストファイル形式でダウンロードした。その際、それぞれの議員の発話で最も古い年代のものから順にダウンロードを行った。こうして得られた発話データから、GNC に関係する環境を発話年の古い順に 1 人 100 件、合計 7,600 件を抽出し、これを本稿で扱うデータとした。この 7,600 件のうち、主語が「が」で標示されていたのは 6,662 件 (87.7%) であり、「の」であったのは 938 件 (12.3%) であった。

2.3 抽出対象となる交替環境

前節で GNC に関係する環境を本稿のデータとして抽出したと述べたが、本節ではどの環境に現れた「が」と「の」を GNC に関係する環境としたのかという、その環境の定義について述べる³⁾。Matsuda (1995) で論じられているように、データ抽出の際には分析対象である変異現象が生起する環境を適切に制限する必要がある。本稿では先行研究での議論に基づいて、抽出する環境を決定した。

本稿で抽出対象とした環境として、まず、多くの先行研究で議論されている(2)のような名詞修飾節が挙げられる。

- (2) a. [風 が / の 強い] 日
- b. [太郎 が / の 買った] 文房具

名詞修飾節以外では、Watanabe (1996)、Hiraiwa (2001)、菊田 (2002) において、(3)にある「まで」節や「より」節のような従属節でも GNC が可能であると指摘されている。

- (3) a. [バス が / の 来るまで] 座っていらっしゃいますか。
- b. [客 が / の 来るより] 早く荷物が着いた。

菊田 (2002)

菊田 (2002) は (3) から、GNC には節によって修飾される主要部名詞の存在が必要条件ではないと述べており、本稿ではこの主張を検証するために「まで / より」節を GNC に関わる環境として抽出した。

また、井上 (1976) で扱われていた、「という / との」が述部と主要部名詞の間に介在する場合も抽出することとした。井上 (1976) では (4) のような節内では「の」

が容認されないと主張している。

(4) [彼ら が/*の 無事だった] という/との 知らせ

井上 (1976)

しかし Ura (1993) では, (5a) のように主要部名詞が動詞由来名詞でない場合には, (5b) のように動詞由来名詞である場合と異なって, 「という」節でも「の」が容認されると主張している。

(5) a. 僕が [[ローマにジョン が/の いた] という 噂/報告/疑惑] を否定した。

b. 僕が [[ローマにジョン が/*の いた] という 知らせ/話/疑い] を否定した。

(Ura 1993)

本稿では「という/との」節をデータ抽出環境に含み, 井上 (1976) の主張と Ura (1993) の主張を検討することとした。

それから (6) のように, 「て」形で並列する述部の主語が同一である場合についても GNC が見られるかどうかを確認するため抽出した。

(6) a. [太郎が考えて出した] 答え

b. [花子が引っ越して, 住み始めた] 町

3. 要因ごとの分析

ここでは, GNC に影響すると想定した要因を言語外的要因 (3.1 節) と言語内的要

因(3.2節)に分類して,要因ごとの分析を行う⁴⁾.言語変化に関わる要因は言語外的要因に含まれ,3.1.1節で散布図と回帰曲線を用いてHarada(1971)の仮説を検証する.その他の要因についてはクロス表を用いた分析を行う.また,それぞれの要因の影響度については,データ中における「が」と「の」の生起率,特にここでは「の」の生起率を用いて議論を進めることとした.

3.1 言語外的要因

3.1.1 議員の生年と言語変化

ここでは発話者76名の生年を用いてHarada(1971)の言語変化仮説の検証を行う.Harada(1971)の仮説は,東京方言において名詞修飾節内の主格標示としての「の」の使用が減少し「が」の使用が増加してきているという仮説と解釈される.もしこの仮説が正しければ,ここで行う検証では,生年が早い時期の議員の発話ほど「の」の生起率が高く,生年が遅い時期の議員の発話ほど「の」の生起率が低くなるという結果が予測される.

図1はGNCにおける「の」の生起率と議員の生年を軸とした散布図である.図1にプロットされている点は,76名それぞれの発話データ100件における「の」の生起率に相当する.また,図1には議員の生年と「の」の生起率を用いたロジスティック回帰によって算出した回帰曲線が描かれている⁵⁾.

図1から,議員の生年が上がるにつれて「の」の生起率が次第に減少していることがわかる.これはGNCにおいて「の」ではなくて「が」が次第に用いられるようになってきていることを表しており,Harada(1971)の言語変化仮説を支持する結果である.Harada(1971)の仮説はこれまで検証されてこなかったが,ここでは長期的発話記録である国会会議録を用いることでその仮説が正しいことを実証できた.

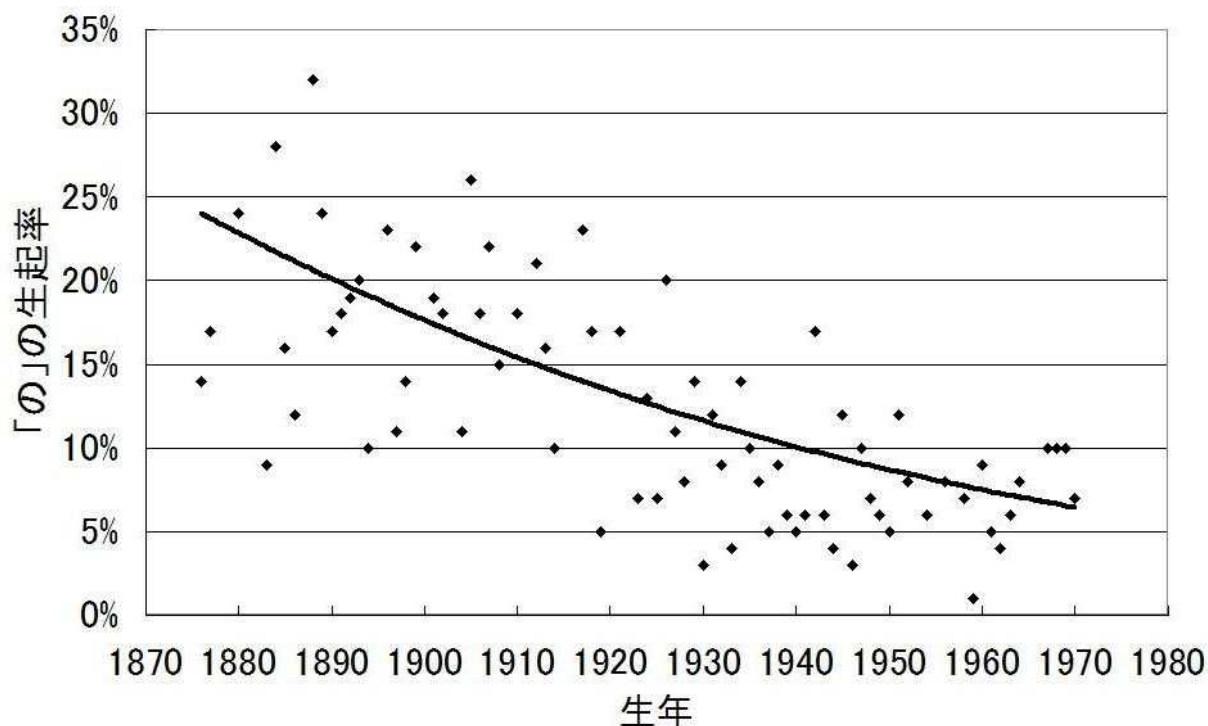


図 1：議員の生年と言語変化

3.1.2 衆議院 / 参議院差，会議の種類

中川（1987）は，どういった影響があるのか特定はしていないが，スタイル差が GNC に影響を与えている可能性を示唆している．ここで GNC のスタイル差を考察する前に問題となるのは，本稿で用いた国会会議録自体のスタイルの捉え方である．国会会議録に収録されている発話のスタイル差について，ここでは国会での発話状況の違いから考察してみる．まず本会議や委員会といった会議の種類が考えられる．国会では多様な会議が開かれているが，本会議では国会議員全員出席の下，テレビ中継も行われる中で会議が開催されるという，他の会議とは異なる特徴を持っている．そのため，本会議とその他委員会という分類ができる．さらに，会議が衆議院・参議院のどちらで行われているか（あるいは両院合同か）といった分類もあげられる⁶⁾．しかし，それぞれの類型における発話状況の特徴や，その状況が発話に与える影響については，今のところ明らかにされていない．この問題については，国会での発話状況と国会会議録の研究によって明らかにされるべき課題と言えるだろう．

以上を踏まえて、本稿では探索的研究として衆議院・参議院の違いと本会議・その他委員会の2つの分類によって国会でのスタイル差が捉えられると仮定して、衆・参の違いと会議の種類2つの要因がGNCに与える影響を考察することとした。

まず、衆・参の違いがGNCに影響を与えているかどうか検討するために、衆・参それぞれにおける「が」と「の」の生起率を表1に示した。

	衆議院	参議院	合計
「が」	88.1% (4,609/5,230)	86.6% (2,053/2,370)	87.7% (6,662/7,600)
「の」	11.9% (621/5,230)	13.4% (317/2,370)	12.3% (938/7,600)
合計	68.8% (5,230/7,600)	31.2% (2,370/7,600)	100% (7,600/7,600)

$$\chi^2 = 3.40, \text{ d.f.} = 1, p < 0.10$$

表1： 衆議院・参議院の違いと「が/の」の生起率

表1からわかるように、衆議院と参議院には「の」の生起率にほとんど差は見られない。また、衆議院・参議院の差は表右下に示されたカイ2乗検定において統計的に有意とならないことが示されている。このため、衆議院・参議院の違いがGNCに影響を与えている可能性は否定される。

次に、会議の種類について検討する。表2に本会議とその他委員会における「が」と「の」の生起率を示した。

表2から、本会議の方が委員会より「の」の生起率が高いことがわかる。さらにその差は、カイ2乗検定において統計的に有意であることが示されている。

それでは、なぜ本会議と委員会で差が現れたのだろうか。その原因を追究するため、ここでは会議の種類と議員の生年との関係に注目し、表2に見られた本会議と委員会の差に対して、実は議員の生年、つまり言語変化が影響しているのではないかという可能性を考えてみる。言語変化の影響について検討するため、本会議と委員会それぞれ

	本会議	その他委員会	合計
「が」	81.7% (437/535)	88.1% (6,225/7,065)	87.7% (6,662/7,600)
「の」	18.3% (98/535)	11.9% (840/7,065)	12.3% (938/7,600)
合計	7.0% (535/7,600)	93.0% (7,065/7,600)	100% (7,600/7,600)

$$\chi^2 = 19.00, \text{ d.f.} = 1, p < 0.001$$

表 2： 会議の種類と「が／の」の生起率

れについて生年で 20 年ごとにデータを区分したところ、委員会のデータは 20 年ごとの区分にほぼ均等に存在していたのに対して、本会議のデータの 49.8% (267/536) が 1876～1895 年と生年が最も初期の 20 年に集中していることがわかった。これは 3.1.1 節で見た、生年が昔であるほど「の」の生起率が高くなるという言語変化の影響を、本会議の方が委員会より受けやすくなることを意味している。そのことで、本会議の「の」の生起率の高さについて一応の説明がつくと考えられるが、GNC に対するスタイル差の影響についてのより詳細な議論のためには、本会議と委員会に存在する発話状況の違いを明らかにする必要がある。

3.2 言語内的要因

3.2.1 隣接性

Harada (1971) は、(7) のように名詞修飾節内の主語とそれに対応する述部の隣接性が GNC に大きく影響することを指摘している。

(7) a. [お金 が／の ある] 人

b. [子供たち が／*の 皆で勢いよく駆け上った] 階段

Harada (1971)

(7a)では主語と述語が隣接しているが、(7b)では主語と述語の間に「皆で勢いよく」という介在要素が存在するため、主語と述語が隣接していない。Harada(1971)は(7b)において「の」の使用が容認されないのは、隣接性の影響によると主張している。その影響とは、(7b)のような非隣接環境では「の」の容認度が(7a)の隣接環境よりも低下するというものである。

隣接性の影響を本稿のデータを用いて検証した場合、もし隣接性がGNCに影響を与えているならば、非隣接環境における「の」の生起率は低くなると予測される。表3に隣接環境と非隣接環境における「が」と「の」の生起率をあげた。

	隣接	非隣接	合計
「が」	82.9% (4,371/5,274)	98.5% (2,291/2,326)	87.7% (6,662/7,600)
「の」	17.1% (903/5,274)	1.5% (35/2,326)	12.3% (938/7,600)
合計	69.4% (5,274/7,600)	30.6% (2,326/7,600)	100% (7,600/7,600)

$$\chi^2 = 363.87, \text{ d.f.} = 1, p < 0.001$$

表3：隣接性と「が/の」の生起率

表3から、主語と述部が隣接である場合の「の」の生起率が17.1%、非隣接である場合が1.5%と、非隣接環境下で「の」の生起率が著しく低下することがわかる。また、隣接環境と非隣接環境の差は、カイ2乗検定によって統計的に有意となっている。これは、隣接性がGNCに強い影響を与えていることを示しており、Harada(1971)の主張を支持する結果である。以下に本稿のデータ中で、非隣接環境下で「の」が使用された例をあげておく⁷⁾。

- (8) a. [歴史的価値 の 非常に高い] ものが [小澤潔 80/05/08 衆 地方行政]
 b. [私 の 一番懸念する] 問題は [山本桑吉 55/06/25 衆 法務]

また Harada (1976) では、非隣接環境下で現れる介在要素の種類も GNC に影響すると主張している。その影響とは、介在要素の名詞性が高いほど「の」の容認度が低下するというものであった。そこで、本稿で観察された非隣接環境下での「の」の使用例 35 件を見てみると、介在要素が名詞句であったのは 4 件であり、それらは全て「～になる」という述部であった。そのうちの 2 例を (9) にあげておく。

- (9) a. 只今 [河井さん の お話になつた] ことと [徳川頼貞 47/08/20 参 皇室]
b. [増田先生 の 御指摘になりました] 通り [越智通雄 61/08/09 衆 災害]

この「～になる」は他動詞と比較しても、その介在要素的卓立は低いと考えられ、また、この 4 件以外はすべて副詞句の介在であったことは、介在要素の名詞性の高さが「の」の容認度を低くするという Harada(1976)の主張は妥当であると考えられる。

さらに、主語・述語間に存在する介在要素について Shibatani (1975) は、「の」の多義性の問題について言及している。

- (10) a. [太郎 が / の 昨日 買った] 本
b. [太郎 が / * の 友達と買った] 本
c. 誰も [太郎 が / の アメリカへ来た] ことを知らない

Shibatani (1975)

Shibatani (1975) は、(10a) のように介在要素が副詞句である場合には問題にならないが、(10b) のように名詞句が後続する場合には、「の」に所有の意味が現れる可能性を指摘している。Shibatani(1975)は(10b)のような環境を「NP-no NP」環境とし、この環境で所有の意味が現れる場合には「の」の使用が認められないと主張している。よって(10c)のように、主語である「太郎」を所有者として解釈する

ことが不可能な場合には、「の」を使用することが可能となるのである。

本稿のデータを用いて「の」の多義性の影響について見てみると、主格標示の「の」で所有の意味を持つことができる「の」であった例は1件も見られなかった。このことは、Shibatani (1975) の主張を支持する結果であると言える。

3.2.2 他動性制約

Watanabe (1996) では、GNC に対して他動性制約が働いていると主張している。他動性制約とは、(11) のように節内に直接目的語が存在する場合は、「の」の使用が容認されないという制約である。

(11) a. *[John の LGB を 貸した] 人

b. *[LGB を John の 貸した] 人

Watanabe (1996)

(11a) のみでは隣接性との違いが明確でないが、(11b) では隣接性が保たれているにもかかわらず「の」が容認されないの、直接目的語の存在が GNC に影響していることになる。これが正しければ、本稿のデータ中でも直接目的語が存在する節において「の」の使用が観察されないはずである。表 4 に述部が動詞である場合において、節内に直接目的語がある場合とない場合の「が」と「の」の生起率を示した。

	直接目的語あり	直接目的語なし
「が」	100% (641/641)	87.9% (5,070/5,768)
「の」	0% (0/641)	12.1% (698/5,768)
合計	8.4% (641/7,600)	75.9% (5,768/7,600)

表 4： 他動性制約と「が／の」の生起率

表 4 から、直接目的語が節内に存在する場合、つまり他動性制約が働く環境の場合において、「の」が全く使用されていないことがわかる。これは、Watanabe (1996) の主張する通り、他動性制約の GNC に対する影響を示す証拠である。以下に、データ中で直接目的語が節内に存在する場合に「が」が使用された例をあげておく。

(12) a. [この事業の申請を 運輸省さん が 受けられた] 際には

[赤羽一嘉 94/06/07 衆 予算]

b. [多くの人々 が 戸惑いを感じている] のが [米津等史 97/03/25 衆 厚生]

3.2.3 状態性

Horie and Kang (2000) は、状態性が GNC に影響すると主張している。Horie and Kang (2000) は、コーパス中の述部の種類と「の」の生起数との関係として、「の」の生起数が動詞 > 存在述部、形容詞 > 連結詞の順に高くなることから、Givón (1979) で議論されていた状態性の概念において、状態性が高いほど「の」の容認度は下がると主張している。つまり動詞のように状態性が低い場合は、「の」の容認度は高いということになる。しかし、Horie and Kang (2000) はコーパス中の「の」の生起数だけを数えているという点に問題がある。上にあげた 4 種類の述部ごとの「の」の生起数は、それぞれの述部自体がコーパス中で占める割合に影響を受ける危険性があるため、述部ごとに「が」と「の」の生起率を出して議論したほうがよいと考えられる。

本稿のデータから状態性と GNC の関係について考察するため、Horie and Kang (2000) に従って、述部が動詞、形容詞、存在述部、連結詞である場合の「の」の生起率を比較する。表 5 に 4 種類の述部ごとの「が」と「の」の生起率をあげた。また、ここでは Horie and Kang (2000) にある 4 種類以外に形容動詞とて形で接続された述部もあげている。

	形容詞	存在述部	動詞	連結詞	形容動詞	て形で接続された述部
「が」	72.2% (575/796)	87.6% (1,107/1,264)	89.4% (4,549/5,090)	99.2% (249/251)	88.2% (127/144)	100% (55/55)
「の」	27.8% (221/796)	12.4% (157/1,264)	10.6% (541/5,090)	0.8% (2/251)	11.8% (17/144)	0% (0/55)
合計	10.5% (796/7,600)	16.6% (1,264/7,600)	67.0% (5,090/7,600)	3.3% (251/7,600)	1.9% (144/7,600)	0.7% (55/7,600)

$$\chi^2 = 227.51, \text{ d.f.} = 5, p < 0.001$$

表5：品詞の種類と「が/の」の生起率

表5から、「の」の生起率の高さは形容動詞とて形で接続された述部を除くと、形容詞 > 存在述部 > 動詞 > 連結詞の順に高いことがわかる⁸⁾。これと Horie and Kang (2000)の階層と比較すると、本稿の階層では形容詞と動詞の順序が入れ替わっていて、状態性の点で一貫した結果とはなっていない。このように、本稿の分析結果からは Horie and Kang (2000)の主張するような状態性の影響は見られなかった⁹⁾。

ここで、状態性と GNC の関係をさらに検討するため、状態性の影響を述部の種類から見ることに加えて、動詞の種類による状態性の違いが GNC に与える影響を検証してみることとした。動詞の種類については、状態動詞とそれ以外の動詞の2つに分類した。状態動詞は金田一(1976)による分類に従い、「ている」が後続できない動詞、「できる」や可能形の動詞とし、それ以外の動詞と比較して GNC に対する状態性の影響を考察することとした。もし、Horie and Kang (2000)の主張どおりに状態性の影響があるならば、状態性が低い方が「の」の生起率が高くなるはずである。

本稿のデータ中で、動詞が状態動詞であった場合の「の」の生起率は 14.2% (224/1,803) で、それ以外の動詞の場合は 10.4% (474/4,551) となり、カイ2乗検定では $\chi^2 = 5.33, \text{ d.f.} = 1, p < 0.05$ と 5%水準で統計的有意差が認められた。この結

果は、Horie and Kang (2000) の主張に反して、状態動詞の方がむしろ「の」の生起率が高くなることを示している。

この結果を受けて、Horie and Kang (2000) の議論とは対照的に、状態性が高いほど「の」の生起率が高くなるという仮説を立ててみる。表5で見た品詞の種類の間層は、形容詞 > 存在述部 > 動詞 > 連結詞となっており、形容詞、存在述部、動詞についてはその仮説は妥当であるが、連結詞は状態性が高いにもかかわらず「の」の生起率が低くなっているため、その仮説に反してしまう。ここで、この連結詞の例外性を説明する試みとして、述部が連結詞である場合と3.2.1節で見たShibatani (1975) の「NP-no NP」環境との関係について考えてみる。

Shibatani (1975) の「NP-no NP」環境とは、「の」に所有の意味が現れる場合には主格として「の」を使用することができないというものであった。そして、(13)のように述語が連結詞である場合に主格の「が」の代わりに「の」を使用すると、「の」に所有の意味が現れて「NP-no NP」環境を形成することがわかる。

(13) a. [先生 が 学生だった] 頃

b. [太郎 が 子供な] わけ

3.2.1節で述べたように、本稿のデータ中では所有の意味にとれる環境で主格として「の」が使用された例はなかった。そのため、連結詞の場合でも(13)の「NP-no NP」環境については同様の理由で「の」が抑制されている可能性が考えられる。本稿のデータ中で実際に(13)のように、述部が連結詞である場合に「の」を使用すると曖昧さが生じる環境がどの程度存在するかを確認してみると、その環境は47.9% (56/117) あり、無論すべて「が」が使用されていた。この結果から、「NP-no NP」環境での「の」の曖昧さが、連結詞の「の」の生起率を低下させる原因の1つとして考えられるが、連結詞の場合に「の」が2件しか現れていないことを考えると、「NP-no NP」

環境を除いた残りの約半数においてなぜ「の」が現れにくいのかは説明されなければならない。

以上より、述部の種類と動詞の種類の2項目で状態性のGNCに対する影響について考察してきたが、Horie and Kang(2000)の主張どおりの結果は得られなかった。また、状態性がGNCに一貫して影響を与えている可能性を考えるにしても、連結詞の例外性の解明が必要である。

3.2.4 有生性

Croft(1990)は多言語間に及ぶ文法現象に対する有生性の影響から(14)の階層を提唱している。

(14) 有生性の階層(Croft 1990)

first, second-person pronouns < third-person pronoun < proper names < human common noun < nonhuman animate common noun < inanimate common noun

Croft(1990)はこの階層において、有生性が高いほど多様な文法現象と共起すると主張している。本稿ではGNCに対する有生性の影響を、主語名詞の有生性を用いて考察することとした。Croft(1990)の有生性の議論から、もしGNCが有生性の影響を受けているならば、主語名詞が有生性の階層で上位にあるほど、「が」と「の」の両方を主格標示として持つことができる動機が強くなるはずである。表6にCroft(1990)の有生性の階層に従った主語名詞の分類におけるそれぞれの「が」と「の」の生起率を示した。

表6において、データ数が少ない3人称代名詞と非人間有生普通名詞を除くと「の」の生起率では人間普通名詞と非有生名詞においてCroft(1990)の有生性の階層に反する結果を示している。ただ、その他のカテゴリーでは有生性の階層と一貫した順序

を示しているため，人間普通名詞と非有生名詞の例外性について，その他の要因が影響している可能性を検討しなければならない。

	1,2 人称 代名詞	3 人称 代名詞	固有名詞	人間普通 名詞	非人間有生 普通名詞	非有生名詞
「が」	65.8% (263/400)	75.0% (3/4)	76.2% (93/122)	89.9% (550/612)	100% (2/2)	87.6% (4,550/5,196)
「の」	34.3% (137/400)	25.0% (1/4)	23.8% (29/122)	10.1% (62/612)	0% (0/2)	12.4% (646/5,196)
合計	5.3% (400/7,600)	0.1% (4/7,600)	1.6% (122/7,600)	8.1% (612/7,600)	0.03% (2/7,600)	68.4% (5,196/7,600)

$$^2 = 166.55, \text{ d.f.} = 5, p < 0.001$$

表 6： 有生性と「が/の」の生起率

3.2.5 主語名詞句の形式差

松田(2000)は東京方言における格助詞「を」の使用と不使用(ゼロ形)について，「を」の付くことができる目的語名詞句の形式差がその変異に影響することを東京語話者の発話データから実証している。その影響として，4種類に分類された目的語名詞句と共起するゼロ形は，疑問詞>語彙名詞>代名詞>名詞節の順にゼロ形が現れる率が高くなっている。本稿で扱っている「が/の」も「を」と同様に格助詞であるので，それらの付く主語名詞句の形式差が影響するかどうかは検討すべき項目である。主語名詞句の形式は松田(2000)に従って，疑問詞，語彙名詞，代名詞，名詞節の4種類に分類し，それぞれの「が」と「の」の生起率を割り出したところ，表7のようになった。

	疑問詞	語彙名詞	代名詞	名詞節	合計
「が」	100% (28/28)	87.1% (4,974/5,710)	77.8% (487/626)	94.9% (1,173/1,236)	87.7% (6,662/7,600)
「の」	0% (0/28)	12.9% (736/5,710)	22.2% (139/626)	5.1% (63/1,236)	12.3% (938/7,600)
合計	0.4% (28/7,600)	75.1% (5,710/7,600)	8.2% (626/7,600)	16.3% (1,236/7,600)	100% (7,600/7,600)

$$^2 = 121.77, \text{ d.f.} = 3, p < 0.001$$

表 7：形式差と「が/の」の生起率

表 7 から、疑問詞において「の」の使用例が見られず、名詞節に関しても「の」の生起率が非常に低くなっていることがわかる。また、代名詞の場合は語彙名詞よりも「の」の生起率が高い。そして表 7 に示された名詞句の形式差は、表右下に示されたカイ 2 乗検定では有意であることが示されている。松田 (2000) の格助詞「を」の場合と比較すると、「の」の生起率は代名詞 > 語彙名詞 > 名詞節 > 疑問詞となっていて、松田 (2000) とは異なる階層を示しており、その原因についても吟味する必要があるが、今回の分析では主語名詞句の形式差が GNC に関わっていることが示された。

3.2.6 否定性

Givón (1979) は言語変化における否定性の影響について論じている。Givón (1979) は言語変化を、単純化に向かう変化 (simplificatory change) と多表現のための変化 (creative-elaborative change) に分類している。そして、多表現のための変化である場合に限り、否定性を含んだ環境はその言語変化に対して保守的であると主張している。GNC に関して言えば、3.1.1 節で観察された言語変化は「の」の使用が減少し「が」のみを使うようになる変化だったので、それは単純化に向かう変化ということになる。従って、Givón (1979) の主張と GNC は関連がないことになるが、否定性

が何らかの影響を与えている可能性を考えてその影響を検討することとした。否定環境の定義は、述語の接辞として「ない」がある場合、もしくは述語が形容詞「ない」である場合とし、「少ない」や「しなければならない」などの「ない」は否定の意味を持っていないので、否定環境には含まなかった。(15)は否定環境である場合の例である。否定環境とそれ以外の環境における「が」と「の」の生起率は表8に示した。

	否定環境	その他	合計
「が」	78.9% (690/875)	88.8% (5,972/6,725)	87.7% (6,662/7,600)
「の」	21.1% (185/875)	11.2% (753/6,725)	12.3% (938/7,600)
合計	11.5% (875/7,600)	88.5% (6,725/7,600)	100% (7,600/7,600)

$$^2 = 70.79, \text{ d.f.} = 1, p < 0.001$$

表8： 否定環境と「が/の」の生起率

(15) a. [太郎 が/の 行かなかった] 店

b. [理由 が/の ない] 喧嘩

表8から、否定環境の方が「の」の生起率が高く、否定環境とその他の環境との差は統計的に有意となっている。この結果は一見すると否定性の影響があると考えられるのだが、より詳細に観察すると、表8で示された「の」の生起率の差は否定性の影響によるものではないことが判明した。否定環境における「の」の生起率の高さは、否定環境に含まれる、述部が形容詞の「ない」である場合の影響によるものであった。否定環境のうち、述部が形容詞の「ない」である場合は「の」の生起率が36.7% (125/341)と非常に高く、形容詞の「ない」を除いた場合の否定環境では11.4% (62/542)となり、否定環境以外の場合とほとんど差がなくなることがわかった。形容詞の「ない」の場合に「の」の生起率が高くなることは、3.2.3節で見たように、

述部が形容詞の場合は「の」の生起率が高くなることから説明がつく。このことから、否定環境自体が GNC に影響を与えていないことが明らかになった。

3.2.7 「という/との」節

2.3 節で見たように、井上 (1976) が「という/との」節で「の」が容認されないと主張しているのに対して、Ura (1993) では「という」節に関しては主要部名詞が動詞由来名詞でない場合には、「の」が容認されるとしている。本稿のデータで彼らの主張を検討するため、表 9 に「という」節、「との」節、それ以外の環境下の「が」と「の」の生起率を比較した。

	「という」節	「との」節	その他	合計
「が」	98.9% (1,308/1,322)	100% (8/8)	85.3% (5,346/6,270)	87.7% (6,662/7,600)
「の」	1.1% (14/1,322)	0% (0/8)	14.7% (924/6,270)	12.3% (938/7,600)
合計	17.4% (1,322/7,600)	0.1% (8/7,600)	82.5% (6,270/7,600)	100% (7,600/7,600)

表 9: 「という/との」節と「が/の」の生起率

表 9 から、「という」節内では「の」の生起率が非常に低くなるのがわかる。「との」節については事例が少ないので、この結果から議論することは難しい。

「という」節内で「の」が使用された例に関して、Ura (1993) で主張されていたように主要部名詞が動詞由来名詞でなかったかどうか確認してみると、Ura (1993) の主張に反して動詞由来名詞と考えられる例が 1 件観察された。(16) にその例をあげた。

(16) [[十分成算 の 立つ] という] 見込み がつくならば [福田篤泰 52/02/07 衆 経済]

動詞由来名詞でない方が「という」節内においても「の」がより多く生起したことを考えると、確かに Ura (1993) の主張を支持する結果であると言えるが、14 件という数の少なさと、(16) のような例外の存在を考慮すると、「という」節で「の」の使用が認可される条件については更なる検討が必要である。しかし、本稿のデータ中で「という」節内にほとんど「の」が現れなかったことを考えると、「という」節が GNC に大きな影響を与えていることは明らかである。

また、「という」節内で「の」が使用された例は全て、生年が 1917 年以前の国会議員の発話から観察されている。これは「の」から「が」へ移行する言語変化の観点から考えると、この環境はすでに GNC の起こらない環境となってしまったと考えられる。

3.2.8 主要部名詞

2.3 節で見たように、Watanabe (1996)、Hiraiwa (2001)、菊田 (2002) は、GNC には主要部名詞の存在が必要条件でないと主張している。本稿では菊田 (2002) で主要部名詞がない例として挙げられていた「より」節と「まで」節を用いて、主要部名詞の存在が GNC に影響するかどうか確認することとした。もし、主要部名詞の存在が必要条件ならば、「より/まで」節において GNC は見られないはずである。

また、主要部名詞の種類が GNC に影響するかどうかも検討する。Harada (1976) は主要部名詞の種類に基づいて、節の名詞性が高いほど「の」が容認されやすくなることを主張している。Harada (1976) のこの主張は、関係節と補文節における主要部名詞の違いから議論しているが、本稿では主要部名詞が形式名詞である場合とそうでない場合を比較することとした。益岡・田窪 (1992) によると、形式名詞は名詞であり

ながら意味的に希薄で、修飾要素なしでは使うことのできない語とされている。このことと Harada (1976) の主張を関連付けると、通常の名詞から逸脱している形式名詞が主要部名詞である場合には「の」の生起率が低くなるという結果が予測される。本稿では益岡・田窪(1992)を参考として、「こと/の/わけ/ところ/よう/もの」を形式名詞とした。

	「まで」節	「より」節	「よう」	「こと」	「の」
「が」	89.5% (17/19)	100% (2/2)	84.9% (706/832)	92.9% (1,054/1,134)	98.5% (2,138/2,171)
「の」	10.5% (2/19)	0% (0/2)	15.1% (126/832)	7.1% (80/1,134)	1.5% (33/2,171)
合計	0.3% (19/7,600)	0.03% (2/7,600)	10.9% (832/7,600)	14.9% (1,134/7,600)	28.6% (2,171/7,600)

「ところ」	「わけ」	「もの」	その他	合計
51.5% (88/171)	100% (645/645)	74.0% (162/219)	76.9% (1,850/2,407)	87.7% (6,662/7,600)
48.5% (83/171)	0% (0/645)	26.0% (57/219)	23.1% (557/2,407)	12.3% (948/7,600)
2.3% (171/7,600)	8.5% (645/7,600)	2.9% (219/7,600)	31.7% (2,407/7,600)	100% (7,600/7,600)

表 10： 主要部名詞と「が/の」の生起率

表 10 に「より/まで」節の場合と主要部名詞が形式名詞である場合の「が」と「の」の生起率を示した。菊田(2002)が指摘するように「まで」節については「の」の使用が見られたのだが、「より」節では「の」が用いられた例はなかった。以下に本稿のデータ中で「まで」節において「の」が使用された例をあげておく。

(17) a. [納得 の 行く] まで御説明を願いたいと考えております

[山口シヅエ 53/07/03 衆 通商]

b. [大臣 の 来る] まで若干の質問をしたいと思っております

[石川要三 78/05/11 衆 地方]

「まで/より」節自体のデータ数の少なさは考慮しなければならないが、GNC は主要部名詞の存在が必要条件ではないとする菊田(2002)の主張を支持する結果が本稿のデータから示された。

次に形式名詞についてだが、表 10 の「の」の生起率からわかるとおり、GNC に対して一貫した影響があるとは言えない。「ところ」と「もの」は高い「の」の生起率を示すのに対して、「わけ」や「の」はそれと反対の結果を示している。これは形式名詞の GNC に対する一貫した影響というよりも、むしろ個々の語彙による差異が関係すると考えられる。ここでは形式名詞を用いて名詞性の影響を分析した結果、GNC に一貫した影響を与えているわけではなかったが、主要部名詞の種類が「の」の生起率を左右することが示された。

3.2.9 「の」先行環境

本稿では(18)のように、「が/の」がつく主語名詞句が「の」で終わる場合を、「の」先行環境としてその影響を考察することとした。

(18) a. [大きいもの が/の ある] 店

b. [ピアノ が/の 聞こえる] 教室

(18)では主格標示として「の」が用いられると、「の - の」という同一音の連続が現れることになる。推測ではあるが、もしそのような連続した同一音が発話中で避け

られるとすれば，この環境では「の」の生起率が減少することになる．このような先行環境の影響がどの程度 GNC に影響を与えるのか検討する．表 11 に「の」先行環境とそれ以外の環境における「が」と「の」の生起率を比較した．

	「の」先行環境	その他	合計
「が」	98.3% (343/349)	87.1% (6,319/7,251)	87.7% (6,662/7,600)
「の」	1.7% (6/349)	12.9% (932/7,251)	12.3% (938/7,600)
合計	4.6% (349/7,600)	95.4% (7,251/7,600)	100% (7,600/7,600)

$$\chi^2 = 38.15, \text{ d.f.} = 1, p < 0.001$$

表 11： 「の」先行環境と「が/の」の生起率

表 11 から，「の」先行環境における「の」の生起率が 1.7% と非常に低く，統計的有意差もあることから，「の」が先行する場合には「の - の」という連続を避けるために「が」を用いるという本稿での仮説を支持する結果となった．

4. ロジスティック回帰分析

3 章では GNC に影響すると想定した要因ごとに，その影響度を「の」の生起率によって議論した．しかしながら，クロス表分析ではカイ 2 乗検定によってその要因の影響が統計的に示されたとしても，他の要因がそこに影響を与えている可能性を否定できない．そこで本章ではロジスティック回帰分析を行い，他の要因の影響を排除した上で，GNC に影響する要因について考察する．

Rietveld and van Hout (1993) で詳述されているように，ロジスティック回帰では求めたい従属変数の値を複数の独立変数から算出する回帰モデルを構築する．GNC の場合は，「が」であるか「の」であるかの 2 値を従属変数の値とし，その値を決定するための独立変数には 3 章で扱った隣接性などの要因を想定することになる．独立

変数の個々の影響の強さはデータに基づいて経験的に決定され、回帰モデル構築の過程で GNC に影響しないと判断された独立変数は、回帰モデルから排除することができる。このようにロジスティック回帰を用いることで、本稿で想定した要因の中でどの要因が GNC に影響するのか特定できる。また、ロジスティック回帰では他の要因の影響を排除した個々の要因の影響度も割り出せるので、要因ごとの GNC に対する単独の影響についても明確に提示できる。

4.1 分析手法

変異研究では、ロジスティック回帰を行うソフトとしてしばしば VARBRUL が用いられるが、VARBRUL は連続変数を独立変数として想定することができない¹⁰⁾。本稿では議員の生年という連続変数を扱いたいので、それを独立変数として想定可能な SPSS を用いることとした。

また、ロジスティック回帰を行う前に、言語変異分析で knockout factor と呼ばれる、従属変数の値として一方の形式のみが現れるカテゴリカルな影響を及ぼす要因を排除しなければならない (Paolillo 2002)。例えば、3.2.2 節で見た他動性制約の場合、直接目的語が節内に存在する環境では「が」しか現れない。このようにカテゴリカルな影響を及ぼす要因については、その影響を受ける環境、例えば、他動性制約の場合であれば直接目的語が節内に存在する環境 (641 件)、をデータから排除しなければならない。このようにカテゴリカルな影響を受ける環境をデータから排除して、最終的に 6,291 件をロジスティック回帰に用いた¹¹⁾。

独立変数の候補には、前章で議論した、衆議院・参議院の差、会議の種類、議員の生年、隣接性、品詞の種類、有生性、主語名詞句の種類、「という」節、主要部名詞、「の」先行環境、の 10 項目が考えられる¹²⁾。これらを用いて分析を行うのだが、その前に、ロジスティック回帰分析で問題となる要因間の従属性についてここで指摘しておかなければならない。要因間の従属性とは片方の要因の値が決まれば自動的にも

う片方の要因の値が決まるという場合を指し、従属性があるとロジスティック回帰では個々の要因の影響を誤って測定してしまうのである。

本稿の要因における従属性は、有生性と主語名詞句の形式差、主語名詞句の形式差と「の」先行環境の2つのペアにおいて観察される。まず、有生性において主語が人称代名詞と分類される場合は、必ず主語名詞句の種類も代名詞となる。それから、代名詞には「の」で終わる語はないため、主語が代名詞である場合は必ず「の」先行環境とはならない。主語名詞句の種類、有生性、「の」先行環境の3変数の従属性の問題を排除するために、ここではそれらを掛け合わせた「主語名詞句の種類×有生性×「の」先行環境(S×Y×N)」という要因を独立変数として想定し、主語名詞句をa(非有生普通名詞)、b(人称代名詞)、c(節)、d(人間普通名詞)、e(「の」で終わる非人間普通名詞)、f(「の」で終わる節)、g(固有名詞)の7種類に分類した¹³⁾。要するに、主語名詞句の種類の中で有生性の分類を行い、それに「の」先行環境による分類を追加するというわけである。その結果、最終的に独立変数として8項目を用いることになった。

また、SPSSによるロジスティック回帰の変数選択は、尤度比に基づくステップワイズ法で行った。ステップワイズ法を用いると、独立変数として想定した8項目のうち、GNCに影響しないと判断された独立変数は最終的に出される回帰モデルから除外される。ここでは、変数増加法と減少法の両方を行い、結果が一貫しているか確認を行った。

4.2 分析結果

表12にロジスティック回帰の分析結果を示した。表12は列にロジスティック回帰で得られる項目、行にロジスティック回帰で構築された回帰モデルに含まれる独立変数が表されている。表12には、独立変数として想定していた衆議院・参議院の差、

	B	標準 誤差	Wald	自由度	有意 確率	Exp(B)	Exp(B)の信頼 区間	
							下限	上限
隣接性	2.42	.19	167.27	1	.000	11.19	7.76	16.14
「という」節	3.11	.28	122.40	1	.000	22.32	12.88	38.69
議員の生年	-.02	.00	99.81	1	.000	.98	.98	.99
品詞の種類			179.55	4	.000			
(形容詞)	1.02	.33	9.49	1	.002	2.79	1.45	5.35
(動詞)	-.56	.32	3.03	1	.082	.57	.31	1.07
(存在述部)	.14	.33	.17	1	.680	1.15	.60	2.20
(連結詞)	-2.41	.79	9.23	1	.002	.09	.02	.43
S × Y × N			120.15	6	.000			
a	-1.06	.27	15.73	1	.000	.35	.20	.58
b	.31	.29	1.13	1	.288	1.37	.77	2.43
c	-1.59	.31	27.39	1	.000	.203	.11	.37
d	-.82	.30	7.24	1	.007	.44	.24	.80
e	-2.21	.56	15.57	1	.000	.11	.04	.33
f	-3.24	1.10	8.74	1	.003	.04	.005	.34
主辞名詞			401.87	5	.000			
「ところ」	.89	.26	11.57	1	.001	2.42	1.46	4.03
「よう」	-.79	.21	14.22	1	.000	.45	.30	.68
「の」	-3.42	.25	182.65	1	.000	.03	.02	.05
「こと」	-.74	.22	11.03	1	.001	.48	.31	.74
「その他」	.08	.19	.19	1	.661	1.09	.75	1.57
定数	27.70	3.29	70.95	1	.000	1.07E+012		

表 12 : SPSS によるロジスティック回帰の出力結果

会議の種類が回帰モデルに入っていないことに注目したい。これは前節で述べたように、ステップワイズ法による変数選択の結果、衆議院・参議院の差、会議の種類はその影響を考慮する必要がないと判断されたため、独立変数としてモデルに含まれなかったからである。よって、表 12 にある要因は GNC に影響を与えていることになる。

それでは、表 12 にあげられた数値について左から順に見ていく。1 列目の B は回

帰係数を表している。回帰係数は個々の独立変数の影響度を表しており、プラスであれば「の」の現れる確率が高く、マイナスであれば「の」の現れる確率が低いという解釈となる。回帰係数の値は Wald 検定によって有意かどうか判断される。3 列目にある Wald 検定値はカイ 2 乗分布に従い、4 列目の自由度が検定に使用される。そしてその有意確率が 5 列目に示されている。表 12 で回帰係数の値が有意とならないのは、品詞の種類における動詞と存在述部、 $S \times Y \times N$ の b (人称代名詞)、主要部名詞の「その他」の 4 つである。しかし、議員の生年のように他と単位が異なる場合、それが回帰係数の大きさに反映されてしまうため、その影響度を要因同士で比較する場合は $\text{Exp}(B)$ で見たほうがよいと思われる。

$\text{Exp}(B)$ は、ある事象の起こりやすさを表すオッズ比に相当する。オッズ比とは、ある結果が起こる確率 (p) とそれが起こらない確率 ($1 - p$) と、もう一方の結果が起こる確率 (q) と起こらない確率 ($1 - q$) の比のことである。 $\text{Exp}(B)$ の値は、その要因に対する他の要因の影響を考慮した値であるため、純粹にその要因のみの GNC への影響を見ることが出来る。また、この値は独立変数中のあるカテゴリーをレファレンスとして算出されている。例えば表 12 の隣接性では、隣接環境と非隣接環境という 2 つのカテゴリーのうち、非隣接環境をレファレンスとしている。そして隣接性の $\text{Exp}(B)$ の値は、隣接環境のほうがレファレンスである非隣接環境よりも 11.19 倍「の」が現れる確率が高くなることを意味する。そのため、 $\text{Exp}(B)$ の値は 1 から遠い程その要因の影響が大きいと解釈できる。

以降、要因ごとに $\text{Exp}(B)$ について順にみていく。「という」節の場合には、「という」節である環境がレファレンスとなっており、 $\text{Exp}(B)$ は「という」節以外の節のほうが 22.32 倍「の」の率が高くなることを示している。この結果は 3.2.7 節で見たクロス表分析の結果に従っている。また、 $\text{Exp}(B)$ は要因間での比較が可能であるため、「という」節と隣接性の 2 つの要因では「という」節のほうが GNC に与える影響がより大きいことがわかる。

議員の生年は連続変数であったため，Exp(B)の数值は1年ごとに「の」の率が0.98倍ずつ減少すると解釈される．このことは，3.1.1節で見た「の」の使用の減少という言語変化を適切に表わしている．

品詞の種類の場合は，形容動詞をレファレンスとしており，その Exp(B)から形容詞>存在述部>形容動詞>動詞>連結詞の順に「の」の率が高くなることが示されている．この結果も3.2.3節で見たクロス表分析と同じ結果である．

次に $S \times Y \times N$ では，人称代名詞>固有名詞>人間普通名詞>非有生普通名詞>節>「の」で終わる非有生普通名詞>「の」で終わる節，という順に「の」の率が高くなることが Exp(B)からわかる．この階層には，3.2節のクロス表分析によって示された主語名詞句の種類，有生性，「の」先行環境の3つの要因のそれぞれの影響が適切に反映されている．主語名詞句の種類に関しては，「の」の生起率が代名詞>語彙名詞>名詞節の順に高いという3.2.5節で観察された影響を反映している．有生性に関して3.2.4節では，有生性が高いほど「の」の生起率が高くなるという仮説を立てた場合に，クロス表において非有生名詞の方が人間普通名詞よりも「の」の生起率が高くなることが問題であったが，上記のロジスティック回帰分析の結果によると，その例外性は現れておらず人間普通名詞>非有生名詞の順で「の」が現れやすいという有生性の仮説に沿った階層となっている．この結果から，有生性がGNCに一貫した影響を与えていると結論付けられる．最後に，「の」先行環境では3.2.9節で観察されたように，主語名詞句が「の」で終わる場合はそうでない場合より「の」が現れにくいことが示された．

主要部名詞の Exp(B)は，「ところ」>「その他」>「もの」>「こと」>「よう」>「の」の順に「の」の率が高いことを表しており，3.2.8節のクロス表とは異なっている．この相違は，Exp(B)の値が他の要因からの影響を排除しているからである．

最後にモデルの適合度についてだが，ここでは Hosmer Lemeshow の検定から適合度についてみておく．Hosmer Lemeshow の検定は，カイ2乗分布を用いた検定方法

であり、それが有意でない結果となった場合に、そのモデルはよいモデルであると判断される。ここで構築された回帰モデルについて Hosmer Lemeshow の検定では、 $\chi^2 = 7.311$, d.f. = 8, $p = .503$ と有意でない数値が出されたので、このモデルはよいモデルだということになる。また、このモデルの Nagelkerke's R^2 値は 0.453 であり、誤判別を表す分類表によるこのモデルの GNC の予測率は 87.2% となった。

以上のようにロジスティック回帰を用いることで、衆参の違い、会議の種類といった要因が GNC には影響を与えていないことが確認できた。また、クロス表からは推測することが困難な、他の要因の影響を排除した要因の単独の影響度が $\text{Exp}(B)$ によって提示され、またその値から個々の要因の影響度を相対的に比較することができた。

5. 結論

本稿では、どの要因が GNC に影響を与えるかを解明するために、クロス表とロジスティック回帰を用いた分析を行った。これまでの GNC を扱う研究では、複数の要因を考慮しておらず、その要因の純粋な影響について議論できているのかは明らかでなかった。ここでは、先行研究で提案されていた要因を含めた複数の要因を同時に分析に含めることで、他の要因の影響を取り除いて個々の要因自体の影響を割り出すことができた。具体的には、他動性制約が働く環境では「の」が使用されないというカテゴリーカルな影響が確認され、議員の生年、隣接性、名詞句の形式差、「という」節、品詞の種類、有生性、主要部名詞、「の」先行環境が独立して GNC に影響すること、また、その影響の強さを明らかにした。

GNC における言語変化についても、Harada (1971) の仮説が提唱されてから 30 年以上も検証されずにいたが、本稿においてその仮説の正しさを実証することができ、GNC には現在「の」の使用から「が」の使用へと移行する変化が存在していることが明らかとなった。

このような言語変化の分析を可能にしたのは、国会会議録である。2.1 節で述べた

ように、長期的発話記録であるというコーパスとして稀有の特徴を持っている国会会議録を用いることによって、Harada (1971) の仮説を検証できたのである。また、発話者についても国会会議録に記載されている発話者名から東京方言話者に対照を絞ることが容易であったことも、国会会議録を利用する利点であった。

しかし、本稿の分析結果の信頼性をより高めるためにも、国会会議録だけでなく他のコーパスを用いた分析が必要である。言語変化の分析には長期的発話記録でなければならないという制限がついてしまうが、GNC に影響を与える要因の特定に関しては、青空文庫などの書き言葉のコーパスや、話し言葉を収録した日本語話し言葉コーパスを利用して、今回の分析で特定された要因と同様の要因が GNC に影響を与えているのかどうかを検証する必要がある¹⁴⁾。

言語変化についても、アンケート等による意識調査を用いて、東京方言で GNC に世代差が見られるかどうかを検討することも今後の課題であろう。本稿で確認されたような言語変化が実際に存在しているならば、上の世代の方が若い世代よりも「の」の容認度が高いという結果が得られるであろう。

また、国会会議録自体の特徴についても明らかにしていく必要がある。今回は言語外的要因として衆議院・参議院の差と本会議・委員会の差を想定したが、国会会議録の発話状況についてより詳細な特徴づけを行い、そこから GNC に影響を及ぼすと思われる要因について考察していかなければならない。

注

* 本稿をまとめるにあたり，松田謙次郎先生には熱心な御指導を頂いたことを，ここに記して感謝申し上げます．また，Joseph Emonds 先生，郡司隆男先生，西垣内泰介先生，松井理直先生，Philip Spaelti 先生，田窪行則先生，小熊猛先生，岸田眞樹氏から貴重なコメント，ご助言を頂いた．この場を借りて感謝の意を表したい．また本稿の内容の一部を，第 2 回神戸オックスフォード言語学セミナー（2004 年 9 月 28 日，於，神戸インスティテュート），第 30 回関西言語学会（2005 年 6 月 4 日，於，関西大学），第 130 回日本言語学会（2005 年 6 月 12 日，於，国際基督教大学），The 17th International Conference on Historical Linguistics（2005 年 8 月 2 日，於，University of Wisconsin），日本語学会春季大会（2006 年 5 月 14 日，於，東京学芸大学）において発表させていただいたが，その席上，有益なコメントをして頂いた方々にもお礼申し上げます．

1) 例えば藤原（2000）では GNC の方言間でのゆれに関して，通常は「の」の使用が容認されない主節においても，九州方言では「の」を主格標示として使用可能であることが指摘されており，GNC を扱う場合は方言間の差異に注意しなければならない．

2) この段階で，TV アナウンサーやタレントなどの職業で特別な言語教育を受けた可能性のある国会議員は分析対象から除外した．

3) ここであげている環境以外にも，「が」と「の」の交替が可能な環境がある．まず，Kuno（1973）で議論されている（i）にある Multiple Subject Construction（MSC）があげられる．

（i）[太郎 が / の 弟 が 病気である] こと

しかし，MSC の場合は「の」が所有の意味を持つことや，Tateishi（1994）で考えられている MSC の統語構造が Miyagawa（1993）で考えられている GNC の構造とは異なるため，本稿では MSC を GNC とは異なる現象として分析対象から排除した．その他に，「が」と「の」の交替が可能である環境として以下の例があげられる．

(ii) 太郎 が / の 試験中に, 事件が起こった .

この環境での交替は, Iida (1987), Hasegawa (1991), 影山 (1993) などで議論されているが, 本稿では(ii)の交替環境には時制が存在しないことに注目し, 時制が必ず存在する GNC とは別の交替現象であると仮定して分析対象から除外した .

4) 言語外的要因として男女差も想定できるかもしれないが, 本稿で発話データを扱った 76 名のうち女性は 2 名だけであり, 男女の比較を行うには偏りが見られるため, ここでは男女差についての考察は行わない .

5) 回帰曲線は LogisticAj_X0.04 を用いたロジスティック回帰によって求めた . 回帰曲線の回帰式は, X=議員の生年, Y=「の」の生起率とする $Y=1/(1+e^{-(29.239-0.0162X)})$ である . また, この回帰式における議員の生年の回帰係数は Wald 検定において, $\chi^2 = 143.499$, d.f. = 1, $p < .0001$ と有意であることが示されており, この回帰モデルの Pseudo R^2 の値は 0.4734 となっている .

6) 山本・安達 (2003) では発話状況として, (1) 質疑応答型, (2) 所信表明型, (3) 趣旨説明型, (4) 承認型, (5) 混合型の 5 つの分類が提案されており, この観点からスタイル差を捉えられる可能性も指摘しておきたい .

7) 国会会議録中の事例を用いる際は, 発言者名, (西暦)年月日, 衆/参, 会議名を記した .

8) 形容詞と動詞における「の」の生起率の差異が, 動詞が他動詞である場合の隣接性と他動性制約の影響を受けている可能性を否定するため, 主語と述部が隣接であり, 動詞については他動性制約のかからない環境における形容詞と動詞の「の」の生起率を比較すると, 形容詞が 31.9% (215/674) 動詞が 15.9% (686/4,309) となり, やはり形容詞と動詞の間には大きな差が認められる .

9) ここで, Horie and Kang (2000) のように「の」の生起数から本稿のデータを見てみると, 「の」の生起数は動詞 > 存在述部 > 形容詞 > 連結詞の順で多くなり, Horie and Kang (2000) のデータと同じ階層が得られた .

- 10) VARBRUL においても連続変数を名義変数に変更することで分析が可能であるが、その際に恣意的な区分を用いなければならなくなる。
- 11) 他動性制約の他にも、主要部名詞が「より」（2件）、「わけ」（645件）である場合、述部がて形で接続されている場合（55件）、主語名詞句が疑問詞である場合（28件）、名詞修飾節と主要部名詞の間に「との」が介在する場合（8件）、それから後述する「主語名詞句の種類 × 有生性 × 「の」先行環境」という混合要因における「の」で終わる人間普通名詞（2件）の場合は、「が」しか現れないためロジスティック回帰の際はそれらを排除して分析を行った。また、主要部名詞のない「まで」節は19件と他と比べるとデータ数が少ないため、これも排除した。
- 12) 否定性については、3.2.6 節で論じたように否定性ありの場合には形容詞「ない」の影響が働いてしまっているため、ここでは候補から除外した。
- 13) 3 人称代名詞は4件と少なかったため、1/2 人称代名詞と合わせて「人称代名詞」と一括りにした。主要部名詞に関しても、件数が少ない「まで」節を「その他」に含むこととした。また、非人称代名詞は knockout factor であったため、分析から除外した。
- 14) 青空文庫はウェブ上に著作権の切れた作品がテキスト形式で無料公開されており (<http://www.aozora.gr.jp/>)、日本語話し言葉コーパスは、国立国語研究所・情報通信研究機構（旧通信総合研究所）・東京工業大学の共同開発により、2004 年から配布が開始されている (<http://www2.kokken.go.jp/csj/public/releaseinfo/index.htm>)。

参照文献

荒尾禎秀 (2005) 「形容動詞化する漢語：「具体」と「具体的」の場合」 『東京学芸大学紀要』 56: 1-14.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina, and Harald Baayen (forthcoming) Predicting the Dative Alternation. In: Gerlof Bouma, Irene Krämer, and Joost Zwarts (eds.) *Cognitive Foundations of Interpretation*. Amsterdam: Royal Netherlands Academy of Science.

Chen, Ch'iu-yen (2002) Changes in Japanese honorifics: a study of `sa-intrusion': An analysis based on electronic data. Unpublished master's thesis, Tokyo University of Foreign Studies.

Croft, William (1990) *Typology and universals*. Cambridge: Cambridge University Press.

藤原与一 (2000) 『日本語方言文法』 東京: 武蔵野書院.

Givón, Talmy (1979) *On understanding grammar*. Orland: Academic Press.

Harada, Shin-ichi (1971) Ga-No conversion and idiolectal variations in Japanese. *Gengo Kenkyu* 60: 25-38.

Harada, Shin-ichi (1976) Ga-No conversion revisited: A reply to Shibatani. *Gengo Kenkyu* 70: 23-38.

Hasegawa, Nobuko (1991) On head movement in Japanese: The case of verbal nouns. 『上智大学言語学会会報』 6: 8-32.

Hiraiwa, Ken (2001) On Nominative-Genitive Conversion. In: Elena Guerzoni and Ora Matushansky (eds.) *MIT Working Papers in Linguistics* 39: 66-125.

Hiraiwa, Ken (2002) Nominative-genitive conversion revisited. In: Noriko Akatsuka, Susan Strauss, and Bernard Comrie (eds.) *Japanese/Korean Linguistics* 10: 546-559.

- Horie, Kaoru and Bongshik Kang (2000) Action/state continuum and nominative-genitive conversion in Japanese and Korean. In: Ritsuko Kikusawa and Kan Sasaki (eds.) *Modern Approach to Transitivity*, 93-114. Tokyo: Kuroshio Publishers.
- Iida, Masayo (1987) Case-assignment by nominals in Japanese. In: Masayo Iida, Stephan Wechsler, and Draga Zec (eds.) *Working Papers in Grammatical Theory and Discourse Structure*, 93-138. Stanford: CSLI Publications.
- 井上和子 (1976) 『変形文法と日本語』東京: 大修館書店.
- 影山太郎 (1993) 『文法と語形成』東京: ひつじ書房.
- 神尾昭雄 (1983) 「名詞句の構造」井上和子 (編) 『日本語の基本構造』77-126. 東京: 三省堂.
- 菊田千春 (2002) 「が・の交替現象の非派生的分析: 述語連体形の名詞性」『同志社大学英語英文学研究』74: 93-136.
- 金銀珠 (2002) 「「の」の論理について: 連体修飾節における性質を中心に」田島毓堂, 釘貫亨 (編) 『名古屋大学日本語学研究室 過去・現在・未来』21-34. 名古屋大学大学院文学研究科.
- 金田一春彦 (1976) 『日本語動詞のアスペクト』東京: むぎ書房.
- 小熊猛 (2004) 「いわゆる「が・の」交替の認知的再考」『日本認知言語学会論文集』4: 381-391.
- 此島正年 (1970) 「「が」の意味・用法」『月刊文法』2: 10-17.
- Maki, Hideki, Tamami Morishima, and Jessica Dunton (2004) A Statistical Analysis of the Nominative/Genitive Alternation in Japanese: A Preliminary Study. *Bulletin of the Faculty of Regional Studies* 14: 87-119.
- 益岡隆志・田窪行則 (1992) 『基礎日本語文法』東京: くろしお出版.
- Matsuda, Kenjiro (1995) *Variable Zero-Marking of (o) in Tokyo Japanese*. UMI

Dissertation Services.

松田謙次郎 (2000) 「東京方言格助詞「を」の使用に関わる言語的諸要因の数量的検証」『国語学』51(1): 61-76.

松田謙次郎 (2004) 「言語研究としての国会会議録検索システム」*Theoretical and Applied Linguistics at Kobe Shoin* 7: 55-82.

松田謙次郎・薄井良子・岡田裕子・南部智史 (2005) 「国会会議録はどれほど発言に忠実か：整文化の実態を探る」『社会言語科学会 第16回発表論文集』267-275.

Miyagawa, Shigeru (1993) Case-checking and minimal link condition. In: Colin Phillips (ed.) *MIT Working Papers in Linguistics* 19: 213-254.

中川良雄 (1987) 「格助詞「が」/「の」の交替について」『研究論叢』28: 309-318.

Nakai, Satoru (1980) A reconsideration of ga-no conversion in Japanese. *Papers in Linguistics*, 13: 279-320.

Nakajima, Motoki (2003) The acquisition of nominative-genitive conversion in Japanese. In: Yukio Otsu (ed.) *The Proceedings of the Fourth Tokyo Conference on Psycholinguistics*, 185-196. Hitsuji Shobo.

日外アソシエーツ (編) (1999) 『現代政治家人名辞典』東京: 日外アソシエーツ株式会社

日外アソシエーツ (2003) 『政治家人名辞典』東京: 日外アソシエーツ株式会社.

日本政経新聞社 (編) (1998) 『国会便覧』第99版. 東京: 日本政経新聞社.

日本政経新聞社 (編) (1999) 『国会便覧』第101版. 東京: 日本政経新聞社.

日本政経新聞社 (編) (2000) 『国会便覧』第103版. 東京: 日本政経新聞社.

日本政経新聞社 (編) (2001) 『国会便覧』第105版. 東京: 日本政経新聞社.

日本政経新聞社 (編) (2002) 『国会便覧』第108版. 東京: 日本政経新聞社.

日本政経新聞社 (編) (2003) 『国会便覧』第110版. 東京: 日本政経新聞社.

日本政経新聞社 (編) (2004) 『国会便覧』第113版. 東京: 日本政経新聞社.

西岡いずみ (1998) 「名詞修飾語標示のノ格助詞と「が・の交替」文のノ格助詞の位置の相違」『九州大学言語学研究室報告』19: 25-35.

Ochi, Masao (2001) Move F and Ga/No Conversion in Japanese. *Journal of East Asian Linguistics*, 10: 247-286.

Paolillo, John (2002) *Analyzing Linguistic Variation*. Stanford: CSLI.

Rietveld, Toni and Roeland van Hout (1993) *Statistical Techniques for the Study of Language and Language Behaviour*. Berlin : Mouton de Gruyter.

Sano, Shinichiro (2005) On the nature of *sa*-insertion: A formal analysis of synchronic variation. Unpublished master's thesis, Sophia University.

Shibatani, Masayoshi (1975) Perceptual strategies and the phenomena of particle conversion in Japanese. In: Robin E. Grossman, L. James San, and Timothy J. Vance (eds.) *Papers from the Parasession on Functionalism*, 469-480. Chicago Linguistic Society, University of Chicago.

衆議院・参議院 (編) (1990a) 『貴族院・参議院議員名鑑』東京: 大蔵省印刷局.

衆議院・参議院 (編) (1990b) 『衆議院議員名鑑』東京: 大蔵省印刷局.

Tateishi, Koichi (1994) *The Syntax of 'Subjects'*. Stanford: CSLI.

Ura, Hiroyuki (1993) L-Relatedness and Its Parametric Variation. In: Colin Phillips (ed.) *MIT Working Papers in Linguistics* 19: 377-399.

薄井良子 (2005) 「国会における反対意見表明の構造 : 国会会議録検索システムを使用して」『社会言語科学会 第 15 回発表論文集』202-205.

Watanabe, Akira (1996) Nominative-genitive conversion and agreement in Japanese: A cross-linguistic perspective. *Journal of East Asian Linguistics*, 5: 373-410.

山本和英・安達康昭 (2005) 「国会会議録を対象とする話し言葉要約」『自然言語処理学』12 (1): 51-78.

要旨

定量的分析に基づく「が/の」交替再考

南部智史

(神戸松蔭女子学院大学大学院生)

本論文では、「が/の」交替について Harada (1971) で提唱された言語変化仮説の検証と、「が/の」交替に影響を与える要因の特定を定量的観点から行う。本論文で扱う要因は先行研究による分析を参考として想定し、それらの要因がどの程度「が/の」交替に影響を与えているのかを経験的に特定する。本論文で扱うデータは、国会会議録に収録されている東京都出身の国会議員 76 名の発話データとし、分析にはクロス表とロジスティック回帰を採用した。分析の結果、Harada (1971) の仮説通り、「が/の」交替において「の」から「が」への移行という言語変化が観察された。また、本論文で想定した 12 要因のうち、生年、隣接性、他動性制約、主語名詞句の形式差、「という」節、述部の種類、有生性、主要部名詞、「の」先行環境、の 9 要因において「が/の」交替に対する影響が確認され、それらの影響度も特定された。

Abstract

Reconsideration of *Ga/No* Conversion Based on a Quantitative Analysis

Satoshi Nambu

Graduate student, Kobe Shoin Graduate School

The purpose of this article is first to verify the language change hypothesis of Harada (1971), but also to identify the factors which affect GNC. Taking advantage of previous research, I consider how potential factors affect GNC, and clarify the degree of their effects on GNC based on a quantitative analysis. The data used for this study comes from the Minutes of the Japanese Diet, and I use cross tabulations and logistic regression. The statistical analysis clarifies not only

which factors affect GNC, but also provides values allowing a comparison of the degree of each effect. The result shows that there is an ongoing change in GNC as hypothesized by Harada (1971) where the use of *no* is decreasing in favor of *ga*. Finally the result shows that of 12 factors considered 9 have a statistically significant effect on GNC: birth year of Diet member, adjacency, transitivity restriction, animacy, predicate type, subject NP type, head noun, *to yuu/to no* clause, and *no* precedence environment.

(受領日 2006 年 11 月 25 日 最終受理日 2007 年 1 月 30 日)