

## 『日本語話し言葉コーパス』に現れる 「さ入れ言葉」に関する数量的分析

佐野真一郎

上智大学大学院生

**【要旨】**本研究では、日本語の使役形に現れる言語変異である「さ入れ言葉」(例、やらさせる、入らさせる)に関して、数量的分析を試みる。さ入れ言葉については、これまで言語内的要因に関しては詳細な分析が行われているが、言語外的要因との関係についてはあまり分析がない。しかしながら、言語変異において性差、スタイル差などの言語外的要因がその分布に大きな影響を与えていることが様々な研究において観察されている。さ入れ言葉の性質の解明のためには、言語外的要因に関して、更に各要因間の関係についても分析を行う必要がある。そこで、本研究では『日本語話し言葉コーパス』を用い、言語外的要因がさ入れ言葉の分布に与える影響、各言語内的・言語外的要因間の関係を明らかにすることを目的とする。分析の結果、各要因のさ入れ言葉の分布への影響、それらの影響の大きさ、各要因間の関係が新たに明らかになった。また、変異理論全体への示唆が得られた\*。

**キーワード:** さ入れ言葉, 日本語話し言葉コーパス, 言語変化, 言語内的要因, 言語外的要因

### 1. はじめに

本研究では、日本語の使役形に現れる共時的言語変異である「さ入れ言葉」(例、やらさせる、入らさせる)を対象として、『日本語話し言葉コーパス』(Corpus of Spontaneous Japanese 以下、CSJ とする)を用い、言語内的・言語外的要因とさ入れ言葉との関係に関して数量的分析を行う。これまで、さ入れ言葉に関しては国会会議録などを使った分析が行われている(陳 2002, 井上 2003, Okada 2003, 佐野 2008 他)。その結果、言語内的要因(接続する動詞の長さや種類、後続要素の種類など)に関しては、その特性が明らかになってきている。しかしながら、これまでスタイルに関して、さ入れ言葉があらたまった場面で使われる傾向があるということが経験的にはわかっているものの、そのことに対して形式的な説明を与えるまでには至っていない。国会におけるスタイル差を定式化することは難しく、また分析に堪えるほどの十分なスタイル差があるとは必ずしも言えないからである。

これに対して、自発音声コーパスである CSJ には、学会講演・模擬講演という

\* 本稿の執筆に際し、松田謙次郎氏、丸山岳彦氏、萩生田伸子氏、日比谷潤子氏、廣瀬英子氏から熱心なご指導と多数の貴重なご助言をいただいた。ここに記して感謝の意を表したい。加えて、重要なお指摘を数多く下さった2名の査読者の方にお礼を申し上げる。本稿における不備は全て筆者に帰するものである。

講演種に関する第一次的な区別に加え、「講演者属性データ」、「印象評定データ」などスタイル差を分析するのに十分な情報が豊富に付与されている。本研究ではこれらの情報を言語外的要因として捉え、そのさ入れ言葉への影響を明らかにし、更に言語内的要因を含めた各要因間の関係を明らかにすることを目的とする。

上記の「あらたまった場面で現れる傾向がある」など、さ入れ言葉がスタイルの影響を受けていることを示唆する振る舞いが観察されていることから、スタイルを始めとする言語外的要因に関する分析を行い、そのような観察に対して説得力のある客観的な説明を与えることによって、さ入れ言葉の性質を解明することができると考えられる。また言語変異研究においては、性差が重要な役割を果たしている研究事例が数多く報告されており、その重要性が認識されているものの、さ入れ言葉では性差を考慮した研究はほとんど存在しない。それは、国会では議員は圧倒的に男性が多く、性差による影響を検証することはほぼ不可能であるためであるが、CSJの一部（「模擬講演」）は男女のバランスを考慮して設計されているため、性差の影響を検証することは十分可能である。これらの背景を踏まえ、さ入れ言葉を日本語の共時的言語変異として捉え、さ入れ言葉とその正用法<sup>1</sup>との比較という形で要因別に分析を進める。第2節では、さ入れ言葉の特徴、及び先行研究を紹介する。第3節では、本研究で使用したデータの収集方法や規範に関する詳細を述べる。第4節では、収集したデータに関する各要因別分析を行い、さ入れ言葉と各要因との関係を検証する。第5節では、多変量解析を行い、各要因の影響の大きさ、及び各要因間の関係を明らかにする。最後に第6節で結論を述べる。

## 2. 問題の所在

### 2.1 さ入れ言葉

さ入れ言葉とは日本語の使役形における言語変異現象であり、従来使われてきた正用法に対して、誤用と見做されている。以下では、五段動詞と一段動詞を例として、正用法とさ入れ言葉の活用について示す。

#### (1) 正用法の活用

	五段動詞（飲む）	一段動詞（食べる）
語幹	nom-	tabe-
終止形	nom-u	tabe-ru
使役形	nom-ase-	tabe-sase-

(1) が示すように、従来の正用法では五段動詞には使役形態素「ase」が、一段動詞では「sase」が接続する。次にさ入れ言葉の活用について見る。

<sup>1</sup> 以下では、学校文法、規範文法で正しいと見做されている、五段動詞に使役形態素「せ」が接続した使役形を正用法とする。しかしながら、このことはさ入れ言葉が誤用であるという立場を支持するものではない。

## (2) さ入れ言葉の活用

	五段動詞 (飲む)	一段動詞 (食べる)
語幹	nom-	tabe-
終止形	nom-u	tabe-ru
使役形	nom- <u>asase</u> -	(*tabe-sasase-)

(2) が示すように、さ入れ言葉とは、正用法における使役形態素「ase」と比べ、「sa」が余分に挿入されている。このために文字通りさ入れ言葉と呼ばれるわけである。また、さ入れ言葉は五段動詞にのみ現れ、一段動詞では現れず、「\*食べさせ」のようにはならない。このことから井上 (2003) では、使役形態素を「sase」に統一しようとする類推的水平化 (analogical leveling, Matsuda 1993 他) による説明が提案されているが、意見の一致を見ていない。

## 2.2 先行研究

さ入れ言葉に関する先行研究としては、国会会議録などを使った研究が挙げられる (陳 2002, 井上 2003, Okada 2003, 佐野 2008 他)<sup>2</sup>。そこで明らかにされているさ入れ言葉の性質をまとめると、以下ようになる。さ入れ言葉は、1) 言語変化であり、近年増加の一途を辿っている、2) 言語変化の初期段階にあり、今後 S 字曲線的な変化を辿ると予測される。言語外的要因としては、3) スタイル上あらたまった発話で使用される傾向がある、言語内的要因に関しては、4) イディオム、名詞形では現れない、5) 正用法とは異なる意味、機能を持っている、6) 受益形式 (-ていただく、-てもらう型) への文法化の途上にある、7) 自制可能動詞 (self-controllable verb) にのみ現れる、8) 「ささ」の連続を含むものはない (二重「さ」制約: Double-sa constraint)、9) 5 モーラ以上の動詞ではほぼ現れない。このように、徐々にその性質が明らかになってはいるものの、性差、スタイルなどの言語外的要因がさ入れ言葉の分布に与える影響や要因間の関係などについて、詳細な分析は行われていない。

## 3. データ

## 3.1 CSJ

本研究では分析対象として CSJ を用いる。CSJ は、約 661 時間、752 万語の日本語音声からなる自発音声コーパスであり、1,417 人の話者による 3,302 の「講演」から構成されている。その大半は学会講演と模擬講演とに大別される自発的な独話である。学会講演とは複数の学会での発表講演を実況録音したものであり、模擬講演とは日常的なテーマ (「今までの人生を振り返って楽しかった出来事」, 「あなたの住んでいる町や地域について」など) に関する 10 分程度のスピーチを録音したものである。一般的に学会講演ではあらたまったスタイルの発話が多く、模擬

<sup>2</sup> 国会会議録の言語分析への応用については松田 (2004) が詳しい。

講演では比較的リラックスした発話が多い。全ての講演は「転記テキスト」によってフィラーや語断片、不明瞭な発音、咳などを含む全ての言語・非言語情報が忠実に書き起こされている。また、コアと呼ばれる約50万語分のデータに対しては、転記テキストの他に講演者に関する情報である「講演者属性データ」、人手で付与された二種類の「形態論情報」（長単位、短単位）、「節単位情報」、「係り受け構造情報」、「談話構造情報」、「イントネーションラベル」などの研究用情報が付与されている。更に、各講演に対してその講演がどのようなスタイル的特徴を持っているかに関する「印象評定データ」が付与されている（前川2004他）。印象評定データは、各講演が聞き手にどのような印象を与える講演であるかを捉えるのに有用であり、各評定項目は言語外的な要因の度合いを示す指標として用いることができる。

### 3.2 データ抽出

本研究では、CSJのコア、ノンコアを含めたあらゆる講演に現れたさ入れ言葉、正用法を対象とする<sup>3</sup>。つまり、生年、出身地などの制限は設けないということである。データの抽出に関しては、まずCSJから本研究で必要となる用例を形態論情報などを参照し行った<sup>4</sup>。次に抽出したデータの中から、本研究の対象とはならない用例、つまりさ入れ言葉、正用法とは見做されないものを除外した。以下にその規範<sup>5</sup>に関する詳細を示す。1) さ入れ言葉は（現時点では）五段動詞にのみ現れ、その他の動詞はさ入れ言葉にはなり得ないため、五段動詞以外の用例は除外する、2) 「お待たせ」、「お聞かせ」、「嫌がらせ」など慣用的に独立して現れる要素は名詞相当と見做し除外する、3) 「説明させ」、「拝見させ」、「発表させ」などの名詞+本動詞の「させ」を含むものは除外する、4) 「晴らせる」、「表せる」などのように使役形態素と音韻的に同じであるものの、可能形態素であるものは除外する、5) 「あわせる」、「知らせる」、「聞かせる」などの他動詞は除外する<sup>6</sup>、6) さ入れ言葉と正用法を厳密に区別する。

以下では6について補足しておく。使役形の中には「出させ」、「飛ばさせ」などのようにさ入れ言葉とも正用法とも分析できるものがある。つまり、これらはさ入れ言葉としては「d-as-ase-」、「tob-as-ase-」と、正用法としては「das-ase-」、「tobas-ase-」と分析できるのである。これらを区別するために以下の規準を設ける。

- I. 自動詞形がない（「\*d-as-」のように更に細かい形態素に分割できない）
- II. 明らかな目的語を持つ

<sup>3</sup> さ入れ言葉の全く現れなかった朗読（Rで始まる講演）、対話（Dで始まる講演）は分析対象から除外している。

<sup>4</sup> コアの形態論情報は、自動形態素解析の結果を人手修正しており、精度は99.9%となっている。一方、コア以外の部分は人手修正がほとんど入っていないため、精度は97～98%にとどまる。

<sup>5</sup> 本研究で用いる規範は佐野（2008）での規範に一部修正を加えたものである。

<sup>6</sup> 具体的には広辞苑第5版（1998）を用い、辞書項目として記載されているものは他動詞であると判断する。

もしこれらの規準を満たすならば、その動詞は他動詞であり、それを含む句は「das-ase-」, 「tobas-ase-」などと分析される他動詞と使役形態素からなる正用法である。もし満たさなければさ入れ言葉と見做す。例えば、「方法は発声者に提示音と同じ高さの音を出させて」(A05F0424)<sup>7</sup>は以下のように分析できる。

- I. 自動詞形がない \*d-u
- II. 明らかな目的語を持つ 同じ高さの音を

よって「das」が他動詞であることが確認でき、「出させる」という句は正用法であると断定できる。

このような動詞の多くは、上記のように他動詞かさ入れ言葉のどちらかに分かれるが、例外的に「飛ばさせ」は他動詞である場合とさ入れ言葉である場合の両方が存在する。後者の場合、例えば「ちょっと時間の関係で飛ばさしていただきまして」(A06F0210)は以下のように分析できる。

- I. 自動詞形がある tob-u
- II. 明らかな目的語を持たない

よって、「tob-」が自動詞であることが確認でき、「飛ばさせ-」という句がさ入れ言葉であると断定できる。自動詞形があっても明らかな目的語を持つ場合は他動詞とする。

### 3.3 要因

分析に移る前に、CSJに付与されている研究用情報の中から、分析の指標として用いる言語外的要因を選択しなければならない。その際に言語学的に説得力のある議論を展開できなければ、その選択は妥当であるとは言いがたい。例えば「聞き取りやすさ」, 「難解な専門用語の有無」, 「方言の多少」などはさ入れ言葉の出現に関して影響しているとは考えにくい。一方、さ入れ言葉があらたまった場面で使用される傾向があるという観察に基づき、「あらたまり度」に関係しているような要因を選択することは妥当であると言える。そこで、まずCSJに付与されている講演者属性データ、印象評定データの中から、言語学的に説得力のある議論が展開できるようなもの、更に事前に検討を行い、さ入れ言葉の分布と関係があると思われる要因の中から幾つかを抽出した。結果として、言語外的要因としては「講演種」(学会講演／模擬講演)、「性別」, 「生年」, 「居住年数」<sup>8</sup>, 「自発性」, 「発話スタイル」, 「得

<sup>7</sup> 例の末尾に付された英数字 (A05F0424) は「講演 ID」と呼ばれる各講演に与えられる指標である。各講演 ID におけるアルファベットのうち、先頭の「A」は学会講演を、「S」は模擬講演を、「M」は「その他」の講演であることをそれぞれ示す。また、中央の「M」は当該の講演の話者が男性であることを、「F」は話者が女性であることをそれぞれ示す。従って、例えば「A01M0001」は男性話者による学会講演、「S02F0002」は女性話者による模擬講演となる。この表記法は、付録における講演 ID でも同様である。

<sup>8</sup> 居住年数に関してはそのままの形では分析に使用できないため、整形を行う (4.2.4 節)。

手・不得手」,「講演経験」の8項目を,また,言語内的要因としては「動詞の長さ」,「後続要素」の2項目を,それぞれ分析に使用することとした。

#### 4. 分析

本節では,各要因に従い,具体的なさ入れ言葉の分析を行う。収集した用例中,さ入れ言葉の総トークン数は42,正用法の総トークン数は1,498であった。さ入れ言葉の全体に対する比率(さ入れ率:さ入れ言葉/(さ入れ言葉+正用法)×100)は2.73%である。表1にこの結果を示す。

表1 全使役形におけるさ入れ言葉の割合<sup>9</sup>

	トークン数
さ入れ言葉	42
正用法	1,498
さ入れ率 (%)	2.73

以下,4.1節で言語内的要因,4.2節で言語外的要因に関してそれぞれ生起率を用いて精査する。

#### 4.1 言語内的要因

本小節では,国会会議録に現れたさ入れ言葉を分析した佐野(2008)における主張のうち,動詞の長さ,後続要素という二つの言語内的要因について,CSJを用いて検証する。

##### 4.1.1 動詞の長さ

言語変化(形態・統語論的变化)において,その変化は短い動詞から長い動詞へと進行するという傾向が確認されている(例,ら抜き言葉, Matsuda 1993)。このことから,言語変化の初期においては長い動詞より短い動詞でさ入れ率が高くなると予測される。

佐野(2008)は国会会議録を用い,東京出身の議員81人の全発話(第1回国会(1947年8月5日)から第162回国会(2005年4月8日)まで)を対象として分析を行っている(総トークン数は,さ入れ言葉:317;正用法4,708,さ入れ率は6.31%である)。そして,佐野(2008)は上記の予測を検証するために,各動詞の長さ(モーラ単位)別にさ入れ言葉の生起率をトークン,タイプごとに分析している。以下に動詞の長さ別のさ入れ言葉の割合を示す。

<sup>9</sup> なお,全ての講演に対して全ての講演者属性データ,及び集合評定データが付与されているわけではなく,一部の項目に関しては「nr」などのように記録がない場合がある。本研究ではそのような場合,欠損値として扱い,対象から外し分析を進めている。そのため各要因別分析においてさ入れ言葉,正用法の総トークン数が必ずしも42,1,498になるわけではない。

表2 動詞の長さ（モーラ）別のさ入れ言葉の割合（国会会議録）

	トークン			タイプ		
	さ入れ言葉	正用法	さ入れ率(%)	さ入れ言葉	正用法	さ入れ率(%)
2モーラ	90	1,411	6.00	14	55	20.29
3モーラ	157	2,220	6.61	17	152	10.06
4モーラ	66	878	6.99	15	109	12.10
5モーラ	2	131	1.50	2	55	3.51
6モーラ	2	68	2.86	2	8	20.00
合計	317	4,708	6.31	50	379	11.66

$$\chi^2=7.952, \text{d.f.}=4, \text{n.s.}$$

$$\chi^2=9.788, \text{d.f.}=4, p<0.05^{10} \text{ (佐野 2008)}$$

表2が示すように、2, 3, 4モーラにおけるさ入れ言葉の全トークンはそれぞれ、90, 157, 66である。それに対し、5, 6モーラではそれぞれ2である。さ入れ率に関しても、2, 3, 4モーラでは6%以上を示しているのに対し、5, 6モーラでは1～3%である。さ入れ言葉は2, 3, 4モーラでは生起率が高く、5, 6モーラでは非常に低いという傾向がうかがえる。更にこの傾向はタイプに関しても同様に観察される。また正用法と比べ、さ入れ言葉は2, 3, 4モーラの動詞では頻繁に現れるのに対し、5モーラ以上の動詞ではほとんど現れない。このことから佐野(2008)では、言語変化の初期にある現段階ではさ入れ言葉は短い動詞に限られ、さ入れ言葉に関してもその変化は短い動詞から長い動詞へと進行すると主張した。

本研究でも同様に、言語変化の初期においては長い動詞より短い動詞でさ入れ率が高くなるという仮説を検証する。そこで、CSJから収集したさ入れ言葉、正用法を動詞の長さ（モーラ単位）に従い分類し、それぞれの分布を求めた。以下に結果を示す。

表3 動詞の長さ（モーラ）別さ入れ言葉、正用法（トークン）の分布（CSJ）

	さ入れ言葉	正用法	さ入れ率(%)
2モーラ	8	515	1.53
3モーラ	31	736	4.04
4モーラ	1	192	0.52
5モーラ	2	51	3.77
6モーラ	0	4	0.00
合計	42	1,498	2.73

$$\chi^2=11.704, \text{d.f.}=4, p<0.02$$

<sup>10</sup> 以下では、帰無仮説（さ入れ言葉、及び正用法の分布（生起頻度）は各要因の影響を受けていない（独立である））を検証するためにカイ二乗検定を行う。検定結果が有意であれば帰無仮説は棄却され、さ入れ言葉、正用法の分布は各要因の影響を受けているということになる。カイ二乗検定において有意差が見られた場合はその有意確率を「 $p<$ 」という形で、有意差が見られない場合はn.s. (not significant) と示す。また、カイ二乗統計量を「 $\chi^2=$ 」、自由度を「d.f.=」という形で示す。

表3が示すように、長い動詞よりも短い動詞でさ入れ率が高くなっている。2モーラにおけるさ入れ率は高いとは言えないが（佐野（2008）においても、短い動詞の中でも3モーラ、4モーラでさ入れ率が高く、2モーラでは若干低い値を示している。）、トークンで見れば2モーラ、3モーラでは比較的多くのさ入れ言葉が現れるのに対し、4モーラ以上では極端に減少する<sup>11</sup>。従って、仮説の通りさ入れ言葉は短い動詞を選好し、それらの動詞から変化が進行すると言える。また、この結果は佐野（2008）のさ入れ言葉は短い動詞に限られるという主張を支持するものである。

#### 4.1.2 後続要素

次に、さ入れ言葉に後続する要素がさ入れ言葉の分布にどのような影響を与えているのか、つまりさ入れ言葉が特定の後続要素を選好する性質があるかということを検証する。

佐野（2008）では、後続要素に関しても分析をしている。以下に後続要素別のさ入れ言葉の割合を示す。

表4 後続要素別のさ入れ言葉の割合（国会会議録）

type	さ入れ言葉	正用法	さ入れ率(%)
-ていただく	307	1,944	13.64
-て	3	573	0.52
-てもら	2	61	3.17
-る	2	1,128	0.18
-た	1	172	0.58
-ない	1	257	0.39
-たい	1	63	1.56
-よう	0	99	0.00
-ます	0	281	0.00
-られ	0	41	0.00
-れば	0	21	0.00
φ	0	44	0.00
-ろ	0	12	0.00
-かねない	0	2	0.00
-しめる	0	8	0.00
-得る	0	1	0.00
-切れない	0	1	0.00

（佐野 2008）

表4は後続要素別にさ入れ言葉の生起率を割り出したものであるが、この表が示すように、さ入れ言葉の後続要素の中では「させていただく」という形が支配的

<sup>11</sup> 5モーラで高いさ入れ率を示しているのは、今回のデータにおけるさ入れ言葉の総トークン数が少ないために、1トークンが結果に大きく影響を与えているためであると考えられる。



ある。

更に、具体的に文法化の進行を把握するために、さ入れ言葉における「させていただく」の10年ごとの数量的変化も分析している。

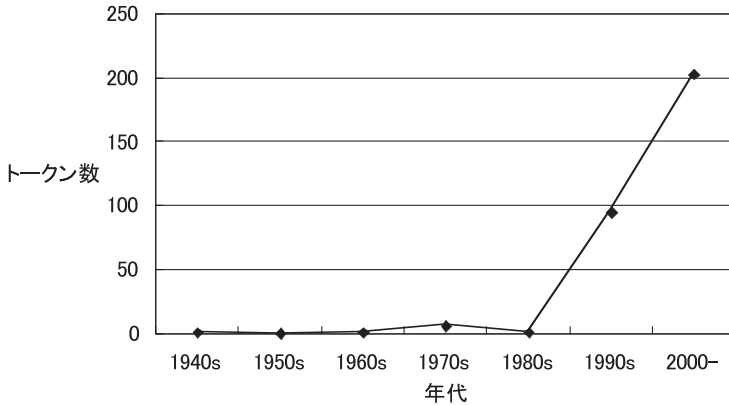


図1 さ入れ言葉における「させていただく」の年代的变化 (佐野 2008)

図1が示すように、1980年代までは数例が観察されるのみであるが、1990年代に95例、2000年に至っては203例と、爆発的に増加している。また、この変化の仕方はさ入れ言葉自体の変化とほぼ同じ傾向を示している。このことから佐野(2008)では、さ入れ言葉が「-させていただく、-させてもらう」という句へと急速に文法化 (grammaticalization, Meillet 1912, Hopper and Traugott 1993 他) しつつあると結論した。

本研究でも、受益形式 (-ていただく、-てもらう型) で圧倒的にさ入れ率が高く、トークンでも多くのさ入れ言葉が現れるという主張、そして文法化が急速に進行しているという仮説を検証するために、さ入れ言葉、正用法を受益形式とその他に分類し、それぞれの分布を見る。結果を以下に示す。

表5 後続要素別さ入れ言葉、正用法の分布 (CSJ)

	さ入れ言葉	正用法	さ入れ率 (%)
受益形式 (-ていただく、-てもらう型)	37	543	6.38
その他	5	955	0.52
合計	42	1,498	2.73

$$\chi^2=46.777, d.f.=1, p<0.001$$

表5が示すように、CSJにおいても受益形式で圧倒的に高いさ入れ率を示し、トークンでも多くのさ入れ言葉が現れていることから<sup>12</sup>、文法化の主張を支持する結果

<sup>12</sup> 受益形式以外では、「-て」(2例)、「-る」、「-てくれ」、そして接続する形態素がないもの(そ

となっている。このことは、さ入れ言葉があらたまった場面で使用される傾向があるということと関係があると考えられる。つまり、受益形式は丁寧な形式であるため、あらたまった場面で使用されるということが推測される。文法化によってその受益形式とまとまりとなって現れるさ入れ言葉は、やはりあらたまった場面で使用されるということである。

本小節では、さ入れ言葉の分布に影響を与える言語内的要因に関して分析を行った。そしてさ入れ言葉は1) ほぼ短い動詞に限られる、2) 「-させていただく、-させてもらう」という句へと文法化しつつあるという先行研究の主張を支持する結果が得られた<sup>13</sup>。

## 4.2 言語外的要因

本小節では、言語外的要因について分析する。扱う要因は講演種、性差、生年、地域差、講演の自発性、発話スタイル、得手・不得手、講演経験の8項目である。

### 4.2.1 講演種

上記のように、さ入れ言葉はあらたまった場面で使用されるということが経験的にわかっている。しかしながら、国会会議録を始めとする従来のデータでは、スタイルを分析するための指標が存在しないため、さ入れ言葉とスタイルとの関係は明らかではなかった。これに対して、印象評定データが付されているCSJでは、スタイル差がさ入れ言葉に与える影響を検証することができる。

CSJにおいて、講演種はスタイルに関連する最も主要な要因である。つまり、学会講演、模擬講演は、ともに複数の聴衆に向けた独話である点では同じであるものの、学会講演では内容も論理的で、比較的自発性の低い、あらたまった発話が顕著であるのに対し、模擬講演は自発性の高い、リラックスした、くだけた発話によって特徴付けられる(前川 2004)など、発話に関連するあらゆる点における違いがこの講演種に集約されると考えることができるのである。従って、ここではまず講演種という要因についての分析を行う。以上のような学会講演、模擬講演のスタイルの違いに基づき仮説を立てるならば、次のようになる。もしさ入れ言葉があらたまった場面で使用されるという観察が正しければ、模擬講演よりも学会講演でさ入れ率が高くなるはずである。この仮説の検証のために、講演種を学会講演、模擬講演の2種類に分類し<sup>14</sup>、さ入れ言葉、正用法それぞれの分布を見る。以下に結果を

れぞれ1例)が観察された。

<sup>13</sup> 佐野(2008)のさ入れ言葉の先行要素に関するその他の主張(イディオム、名詞形には現れない(例、\*口を滑らさせる、\*目を光らせる、\*お知らせ、\*嫌がらせ)、「ささ」という連続を含むさ入れ言葉はない(二重「さ」制約: Double-sa constraint)(例、\*出ささせていただきます、\*正ささせていただきます)、さ入れ言葉は自制可能動詞にのみ現れる(self-controllability restriction)(例、\*悲しませていただきます、\*冷え込ませていただきます)など)に関しても、それらを支持する結果となった。

<sup>14</sup> 本研究では「学会講演」、「模擬講演」、「その他」の3種類の講演を対象としているが、「そ

示す。

表6 講演種別さ入れ言葉, 正用法の分布

	さ入れ言葉	正用法	さ入れ率(%)
学会講演	14	545	2.50
模擬講演	18	913	1.93
合計	32	1,458	2.15

$$\chi^2=0.542, \text{d.f.}=1, \text{n.s.}$$

表6が示すように、模擬講演よりも学会講演でさ入れ率が高くなっている。つまり、仮説の通りさ入れ言葉はスタイル上くだけた発話が多い模擬講演よりもあらたまった学会講演を 선호すると考えることができる。ここから、さ入れ言葉はあらたまった場面で使用されるということが客観的な形で示された。ただし、表6の結果が有意でなかったことは、講演種という要因自体は有意ではなく、講演種とその他の要因との交互作用が講演種の分布に影響しているという可能性を示している。以下ではこの可能性を考慮し、分析を進める。

#### 4.2.2 性差

言語変異研究において性差の重要性は認識されてはいるものの、データの性質等の問題もあり、これまでさ入れ言葉に関してその影響が検証されることはなかった。従って、性差のさ入れ言葉への影響を検証することは本研究のもう一つの主眼に挙げられる。言語変化(形態・統語論的变化)では、女性がそれをリードするという傾向があるとされる(Labov 1990, 2001, Milroy and Gordon 2003 他)。従って、さ入れ言葉においても女性が変化をリードしているならば、変化の初期段階である本データにおいて女性がより高いさ入れ率を示すと予測できる。この仮説を検証する。

表7 性別ごとのさ入れ言葉, 正用法の分布

	さ入れ言葉	正用法	さ入れ率(%)
男性	34	890	3.68
女性	8	608	1.30
合計	42	1,498	2.73

$$\chi^2=7.898, \text{d.f.}=1, p<0.005$$

表7が示しているように、女性よりも男性の方が高いさ入れ率を示している。こ

---

の他」に関しては、その性質は厳密に定義されておらず、また定義することは難しいと思われる。従って、たとえ得られた結果が有意であったとしても、それに対して言語学的に説得力のある議論を展開することが不可能であると思われるため、講演種の分析に関しては「その他」は除外する。

れは上記の仮説に反する。しかしながら、CSJでは男女の比率が学会講演と模擬講演で大きく異なるため、その講演種ごとの男女の構成比がこの結果に影響を与えているということが考えられる。そこで、講演種ごとの性差に関するさ入れ言葉、正用法の分布を見る。以下に結果を示す。

表8 講演種及び性別ごとのさ入れ言葉、正用法の分布

学会講演	さ入れ言葉	正用法	さ入れ率 (%)	模擬講演	さ入れ言葉	正用法	さ入れ率 (%)
男性	9	450	1.96	男性	15	416	3.48
女性	5	95	5.00	女性	3	497	0.60
合計	14	545	2.50	合計	18	913	1.93

$$\chi^2=198.650, \text{d.f.}=3, p=0.00$$

表8が示すように、学会講演では男性よりも女性の方が高いさ入れ率を示している。一方、模擬講演では女性よりも男性の方がさ入れ率が高くなっている。また、さ入れ言葉は2種類の講演で全く異なる分布を示していることから、要因としての性差には講演種との交互作用が反映していると考えられる。更に男性、女性それぞれの学会講演と模擬講演でのさ入れ率の差を見ると、男性は約1.5%の差であるのに対し、女性は約4.5%もの差がある。では、なぜ学会講演では、女性が男性よりさ入れ言葉を選好し、模擬講演では男性の方がさ入れ言葉を選好しているのか、またなぜ男性、女性で学会講演と模擬講演におけるさ入れ率の差が異なるのであろうか。このことに関して、女性の方が男性よりも場面によるスタイル差が大きく、また女性がさ入れ言葉に対してあらたまった場面で使うことに積極的であるのに対し、男性の方は規範意識が強いのか、それを抑える傾向にあるためであるという説明が考えられる。

#### 4.2.3 生年

ここでは、生年がさ入れ言葉の分布と関係があるかということに関して検討する。もし、さ入れ言葉が言語変化であるならば、生年との相関が見られるはずである。すなわち、生年を見かけ時間 (apparent time) と捉えるならば、その時間の流れとさ入れ言葉の分布に何らかの関係があるということである。この可能性を検証するために、まず講演者の生年を10年ごとに区分し、それに従ってさ入れ言葉、正用法を分類した。以下に結果を示す。

表9 生年代別さ入れ言葉, 正用法の分布

	さ入れ言葉	正用法	さ入れ率(%)
1920s	0	5	0.00
1930s	2	144	1.37
1940s	2	217	0.91
1950s	6	284	2.07
1960s	9	365	2.41
1970s	15	464	3.13
合計	34	1,479	2.25

$\chi^2=4.191, d.f.=5, n.s.$

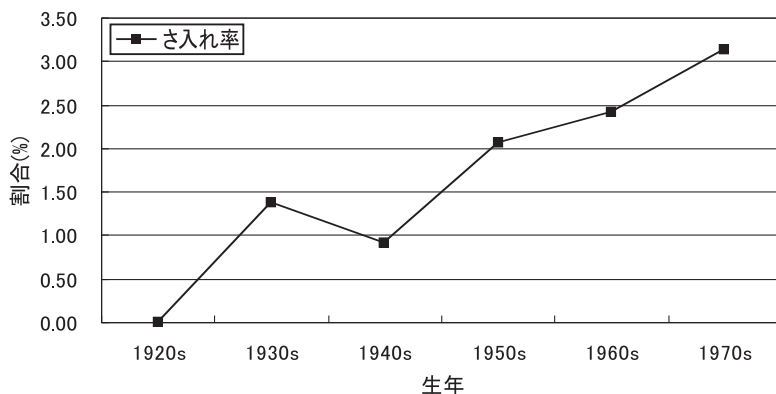


図2 生年代別さ入れ言葉の割合の変化

表9, 図2において, トークンで見ると1920sではさ入れ言葉は観察されず, それ以降増加している。さ入れ率では1920sから, 1940sで停滞してはいるものの, それ以降徐々に高くなっている。

佐野(2008)ではさ入れ言葉と国会議員の生年との相関に関する分析, 及び仮説の検証を行なっている。以下に表と議論を示す。

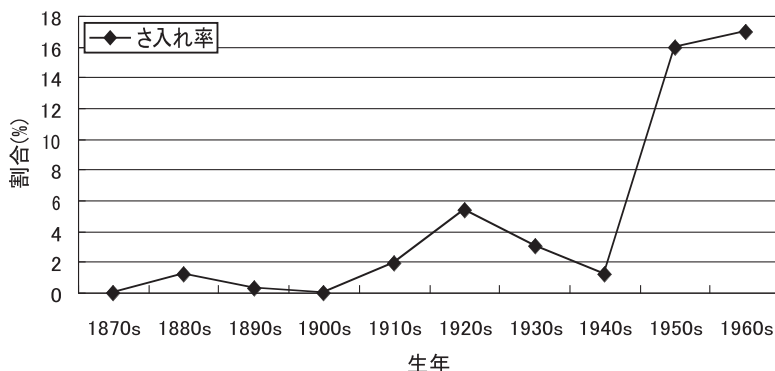


図3 生年別のさ入れ言葉の割合 (佐野 2008)

図3が示すように、生年の若い議員ほどさ入れ言葉を頻繁に使っている。この傾向はとりわけ1940年以降に生まれた議員に顕著である。よって、議員の生年はさ入れ言葉の生起率に対して影響を与えていると言える。この結果から、さ入れ言葉が言語変化であり、更に生年が若くなればなるほどさ入れ率が増加しているということから、さ入れ言葉の変化は進行中であると結論している。

本研究における上記の分布は、佐野(2008)におけるものと酷似している。この結果から、さ入れ率は生年が若くなるに従って高くなっていると言える、つまり生年との相関が見られるのである。このことはさ入れ言葉が進行中の言語変化であることを示している。またその分布を見ると、その変化は初期段階にあり、今後増加を続けるということが予測できる。佐野(2008)は、さ入れ言葉が進行中の言語変化であると主張しているが、その主張とも一致する。

#### 4.2.4 地域差

先行研究の中でも、これまでさ入れ言葉の地域差に関する研究は行われていない。従って、さ入れ言葉の地域的な分布やどの地域で出現したのか、どのように伝播しているのかということは明らかになっていない。よって、ここではさ入れ言葉の地域差について分析する。

CSJには講演者属性データとして、地域差に関係する「居住年数」、「居住年数(言語形成期)」などが付与されている。ここでは居住年数、居住年数(言語形成期)を分析に使用する。まず居住地を北海道、東北地方、関東地方、中部地方、近畿地方、中国地方、四国地方、九州地方、沖縄、外国の10の地域に分割し、居住年数に関しては居住歴中最近10年間で最も長く(最低3年以上)住んだ地域(以下、居住地とする)という規準に従って、さ入れ言葉、正用法のデータを各地域に振り分けた。また居住年数(言語形成期)に関しては、言語形成期中で最も長く住んだ地域(以下、居住地(言語形成期)とする)という規準で分類している。以下がその結果である。

表10 居住地、及び居住地（言語形成期）別さ入れ言葉、正用法の分布

居住地 (言語形成期)	さ入れ 言葉	正用 法	さ入れ 率(%)	居住地	さ入れ 言葉	正用 法	さ入れ 率(%)
北海道 (23人)	8	40	16.67	北海道 (6人)	0	8	0.00
東北地方 (27人)	0	84	0.00	東北地方 (16人)	2	45	4.26
関東地方 (340人)	23	801	2.79	関東地方 (493人)	38	1,160	3.17
中部地方 (77人)	6	175	3.31	中部地方 (33人)	2	60	3.23
近畿地方 (87人)	2	169	1.17	近畿地方 (74人)	0	128	0.00
中国地方 (39人)	0	73	0.00	中国地方 (22人)	0	43	0.00
四国地方 (22人)	0	37	0.00	四国地方 (3人)	0	3	0.00
九州地方 (52人)	2	103	1.90	九州地方 (13人)	0	19	0.00
沖縄 (1人)	0	1	0.00	沖縄 (1人)	0	1	0.00
外国 (6人)	1	11	8.33	外国 (14人)	0	29	0.00
合計	42	1,494	2.73	合計	42	1,496	2.73

表10はそれぞれ言語形成期の居住地と最近10年間のうち最低3年間の居住地(以下、現在の居住地とする)に基づくさ入れ言葉の分布を示すものである。各地方名の右側に、今回のデータにおける、その地方に属する話者の人数を示す(例、北海道(23人))。言語形成期の居住地では、さ入れ言葉は広範な地域に分布している。それに対し現在の居住地では、関東地方を中心に東北地方、中部地方にのみ分布している。もし言語変化は波状に伝播するという立場に立つのであれば、言語形成期の居住地に基づく分布より、現在の居住地に基づく分布の方が自然なものとなる。つまり、さ入れ言葉は成人になってから獲得されるということになる。しかしながら、さ入れ言葉は言語形成期に獲得され、それが成人になっても保たれているのか、それとも成人になってから獲得されたのかという疑問が生じる。もし後者であるならば、形態・統語的な範疇に属するさ入れ言葉が成人になってから獲得されることになり、語彙的な要素以外(音韻、形態、統語)は子供のときに獲得され、その後の変化は見られないという言語獲得に関する主張(Labov 1994, Weinreich 1968)に対する反例となる。そこでより詳細な分析を行うために、以下にさ入れ言葉と正用法それぞれの言語形成期の居住地と現在の居住地に基づくクロス表を示す。

表 11 さ入れ言葉に関する居住地、及び居住地（言語形成期）のクロス表

言語形成期 \ 現在	北海道	東北	関東	中部	近畿	中国	四国	九州	沖縄	外国
北海道			8							
東北										
関東		1	22							
中部		1	3	2						
近畿			2							
中国										
四国										
九州			2							
沖縄										
外国			1							

表 12 正用法に関する居住地、及び居住地（言語形成期）のクロス表

言語形成期 \ 現在	北海道	東北	関東	中部	近畿	中国	四国	九州	沖縄	外国
北海道	1		34	2	2			1		
東北	1	18	56	6	1			1		1
関東	4	17	710	10	36	6	1	4		13
中部	1	2	128	31	8			1		4
近畿	1	1	97	1	60	6		2		1
中国		5	32	3	4	22	1	1		5
四国			14	2	10	3	1	1	1	5
九州		2	78	3	6	6		8		
沖縄			1							
外国			8	2	1					

縦軸が言語形成期の居住地、横軸が現在の居住地を表している。表 11 が示すように、さ入れ言葉を使用する話者の中で言語形成期を関東以外で過ごした者のほとんどが、現在では関東に集中している。このことは、そのような話者は、関東に移住してからさ入れ言葉を獲得したということを示唆する。更に、そのことからさ入れ言葉が関東地方で出現し、周辺地域へと伝播したという推測ができる。また近畿地方以西や北海道、外国では現在の分布ではさ入れ言葉が観察されていない。よって、そのような地域で言語形成期（過去）に、現在の言語変化であるさ入れ言葉を言語獲得のインプットとして聞いたとは考えにくい。それゆえ、大人になってから



獲得したと考えるのが自然である。対照的に表 12 が示すように、正用法は現在のデータで多くの地域での観察が認められており、言語形成期（過去）でも正用法がその地域で使われていた可能性は保障される。よって、正用法を言語形成期に獲得し、それを保ったまま移住をし、現在の分布となっている（つまり、成人での獲得ではない）ということは十分考えられる。

更に、網掛けのしてあるセルは言語形成期と現在の居住地が一致するデータを示している。つまり、言語形成期から現在まで関東地方、中部地方に一貫して住んでいる話者がさ入れ言葉を使用するということである。そのような話者がさ入れ言葉のインプットを得るためには、この地域でさ入れ言葉が使用されていなければならない。このことは関東あるいは中部地方でさ入れ言葉が出現し、周辺地域へと伝播したという可能性を支持するものである。

また上記の、形態・統語的な範疇に属するさ入れ言葉が成人になってから獲得されることになり、語彙的な要素以外（音韻、形態、統語）は子供のときに獲得され、その後の変化は見られないという言語獲得に関する主張に対する反例となるという議論に関しては、もし文法化の仮説が正しければ、さ入れ言葉は語彙的要素ということになり、成人におけるその獲得は上記の主張に対する反例とはならない。しかしながら、文法化は進行中であることが示されており、語彙的な要素であるとは断言できない。従って本研究における結果が、言語獲得に関する主張に対する反例となる可能性は十分にある。

#### 4.2.5 講演の自発性

ここでは、講演の自発性がさ入れ言葉の分布に与える影響を検証する。以下で見える発話スタイル、講演経験とともに、講演の自発性は上記の講演種（学会講演／模擬講演）の下位範疇<sup>15</sup>であると考えられる。つまり、講演種とは発話に関する様々な要因を総合したものであり、その中に講演の自発性、発話スタイルが含まれているということである。しかしながら、講演種別の自発性の高低は講演種間で定めた属性であり、講演種内での属性ではない。つまり、比較的自発性が低いとされる学会講演の中でもそれが高い講演があり、自発性が高いとされる模擬講演の中でもそれが低い講演があるということは容易に想像できる。よって、講演の自発性を講演種に帰することによる要因の検証では不十分であり、講演の自発性という個別の要因の形で検証がなされなければならない。そこで、以下に自発性の高低に従いさ入れ言葉、正用法それぞれを分類した表を示す。

<sup>15</sup> 本研究において、「下位範疇」とは各要因が言語現象に影響を与える際のそれらの関係、構造を指すものであり、CSJ の設計において各要因間に構造を設けているということを意味するものではない。

表 13 講演の自発性別さ入れ言葉, 正用法の分布

	さ入れ言葉	正用法	さ入れ率(%)
自発性が低い	2	139	1.42
自発性がやや低い	2	125	1.57
普通	6	272	2.16
自発性がやや高い	8	407	1.93
自発性が高い	24	554	4.15
合計	42	1,497	2.73

$$\chi^2=7.305, \text{d.f.}=4, \text{n.s.}$$

表 14 性別及び講演の自発性ごとのさ入れ言葉, 正用法の分布

男性	さ入れ言葉	正用法	さ入れ率 (%)	女性	さ入れ言葉	正用法	さ入れ率 (%)
自発性が低い	0	87	0.00	自発性が低い	2	52	3.70
自発性がやや低い	2	58	3.33	自発性がやや低い	0	67	0.00
普通	5	155	3.13	普通	1	117	0.85
自発性がやや高い	6	253	2.32	自発性がやや高い	2	154	1.28
自発性が高い	21	336	5.88	自発性が高い	3	218	1.36
合計	34	889	3.68	合計	8	608	1.30

$$\chi^2=24.927, \text{d.f.}=12, p<0.02$$

表 13 では、自発性が高いほどさ入れ率が高くなる傾向が見て取れる。つまり、さ入れ言葉が自発性の高い講演を選好しているということである。これは講演種の結果からの予測と正反対の結果である。また表 14 から、女性では講演の自発性との目立った関係は見受けられず、この傾向は男性によるものであるということがわかる。このことから、講演の自発性という要因自体では高いほどさ入れ率がより高くなり、トークンでも多くのさ入れ言葉が現れる。講演種においても自発性が高い模擬講演を選好している。しかしながら、その他の発話スタイルなどの様々な要因の影響により自発性の影響が相殺され、講演種という要因としては、自発性の低い学会講演においてさ入れ率が高くなっているという結果となっている。そして、性別との相関があるということが推測される。

ここで、自発性が高いほどさ入れ率が高くなるということから、さ入れ言葉の変化は書き言葉に現れるほどには進行していないという点を指摘しておきたい。陳(2002)が(変化の初期段階では)書き言葉にはさ入れ言葉はほぼ現れないと主張しているように、言語変化は話し言葉から始まり、書き言葉へと進行してゆくと考えることができる。従って、自発性の高い発話が「話し言葉的」であるのに対し、自発性の低い発話は朗読調、つまり「書き言葉的」であるために、さ入れ言葉が前者を選好しているという推論ができるのである。

#### 4.2.6 発話スタイル

次に講演種のもう一つの下位区分と考えられる発話スタイルについての検証を行う。発話スタイルについても、講演の自発性と同様の議論により、それ自体での検証が必要であると考えられる。さ入れ言葉があらたまった場面を選好するということから、ここでは発話スタイルという要因に関して、さ入れ言葉はあらたまった発話において、よりさ入れ率が高くなるという仮説を立てる。以下に、発話スタイルに従い、さ入れ言葉、正用法の分布をまとめた表を示す。

表15 発話スタイル別さ入れ言葉、正用法の分布<sup>16</sup>

	さ入れ言葉	正用法	さ入れ率(%)
くだけた	8	346	2.26
普通	20	678	2.87
あらたまった	14	422	3.21
合計	42	1,446	2.82

$$\chi^2=0.653, \text{d.f.}=2, \text{n.s.}$$

表16 性別及び発話スタイルごとのさ入れ言葉、正用法の分布

男性	さ入れ言葉	正用法	さ入れ率(%)	女性	さ入れ言葉	正用法	さ入れ率(%)
くだけた	6	216	2.7	くだけた	2	130	1.52
普通	17	377	4.31	普通	3	301	0.99
あらたまった	11	264	4.00	あらたまった	3	158	1.86
合計	34	857	3.82	合計	8	589	1.34

$$\chi^2=8.167, \text{d.f.}=6, \text{n.s.}$$

表15が示すように、あらたまった発話でさ入れ率が高くなっている。また表16からも、性別によりある程度の差はあるものの、やはりさ入れ言葉はくだけた発話よりもあらたまった発話を選好していると言える。よって、仮説の通りさ入れ言葉

<sup>16</sup> 発話スタイルに関しては、CSJでは本来1～5の5段階の尺度（「1」から「5」に向かって、「あらたまり度」が高くなる）での評定結果が各発話に付与されているが、そのままの尺度では、項目によっては、トークン数が極端に少なくなってしまうことなどにより、統計的に意味のある結果を導き出すことが不可能となってしまうため、このような分析に耐え得るような分布を得る（項目による極端なばらつきを無くす）目的で、「1」と「2」を結合し「くだけた」へと、「4」と「5」を結合し「あらたまった」へと変換し、それぞれの分布を分析する（「3」に関しては、その必要がないため変換を行っていない）。以下、同様の理由により、得手・不得手、講演経験それぞれの尺度の変換を行っている。得手・不得手に関しては、本来「得意」、「やや得意」、「やや不得手」、「不得手」の4段階での評定となっているが、「得意」、「やや得意」を結合し「得手」へと、「やや不得手」と「不得手」を結合し「不得手」へとそれぞれ変換している。また、講演経験に関しても、本来「はじめて」、「5（回以上）」、「10」、「15」、「20」、「21」の6段階での評定となっているが、「はじめて」と「5」を結合し「5回以下」へと、「10」、「15」、「20」、「21」を結合し「6回以上」へとそれぞれ変換している。

はあらたまった発話を選好するということが支持された。しかしながら、カイ二乗検定の結果が有意でないことから、他の要因との交互作用の存在が示唆されるため、以下で議論する。また講演種の低位範疇としての発話スタイルという要因に関して、あらたまった発話ほどよりさ入れ率が高くなり、講演種でも同じくあらたまった学会講演が選好されている。つまり、さ入れ言葉の分布に対する影響に関して、発話スタイルは講演の自発性とは異なり、講演種という上位の範疇を構成する主要な要因であると言える<sup>17</sup>。

#### 4.2.7 得手・不得手

ここでは、得手・不得手がさ入れ言葉の分布に与える影響に関して検証する。得手・不得手とは講演の得意・不得意に関する尺度である。これは第三者による評定ではなく、講演者自身が事前のアンケートの際に講演が得意か不得意かを回答したものである。従って、上記の第三者の評定である発話スタイルや講演の自発性などとは異なるものであり、独立した要因であると考えられる。ここでは、講演が得意であると答えた話者を「得手」、不得意であると答えた話者を「不得手」としている。この要因の検証のために、得手・不得手という尺度に従ってさ入れ言葉、正用法を分類し、それぞれの分布を求めた。以下に結果を示す。

表 17 得手・不得手別さ入れ言葉、正用法の分布

	さ入れ言葉	正用法	さ入れ率(%)
得手	19	401	4.52
不得手	23	1,055	2.13
合計	42	1,456	2.80

$$\chi^2=6.336, d.f.=1, p<0.02$$

表 17 から明らかであるように、「得手」の方がさ入れ率が高く、トークンでもより多くのさ入れ言葉が現れている。つまり、講演が得意と意識している話者ほどさ入れ言葉を選好するということである。このことは話者の講演の得意・不得意に関する意識がさ入れ言葉の使用に影響を与えているということを示している。

#### 4.2.8 講演経験

最後に講演経験がさ入れ言葉の分布に与える影響を検証する。この要因に関しても、学会講演をするような話者は模擬講演でスピーチをする話者よりも講演経験が豊富であると予測されるが、講演内での講演経験の差を明らかにしなければならない。この検証のために、講演経験を 6 回以上、5 回以下に分け、それに従いさ入れ言葉、正用法それぞれを分類した。以下に結果を示す。

<sup>17</sup> これらの要因の相関に関しては第 5 節で詳細な議論をする。

表 18 講演経験別さ入れ言葉、正用法の分布

	さ入れ言葉	正用法	さ入れ率(%)
6回以上	20	454	4.22
5回以下	22	1,018	2.12
合計	42	1,472	2.77

$$\chi^2=5.344, \text{d.f.}=1, p<0.025$$

表 18 では、講演経験が 6 回以上の話者が 5 回以下の話者よりも高いさ入れ率を示している。つまり、講演経験が豊富なほどさ入れ言葉を選好しているということである。また、講演種の下位範疇としての講演経験という要因に関して、講演経験が豊富な話者ほどよりさ入れ率が高く、講演種でも同じく講演経験が豊富な話者が多いであろう学会講演が選好されている。つまり、さ入れ言葉の分布に対する影響に関して、講演経験も発話スタイルと並び、講演種という上位の範疇を構成する主要な要因であると予測できる。

更に、上記の得手・不得手とさ入れ言葉の関係に関する議論の結果、4.2.7 節では講演が得意と意識している話者ほどさ入れ言葉を選好するということが明らかになり、更にここで講演経験が豊富な話者ほどさ入れ言葉を選好しているということが明らかになった。この二つの要因に対して、さ入れ言葉は同様の振る舞い(「講演」に対して肯定的な項目を選好する)を示している。ここで、講演経験が豊富な話者ほど自身の講演に自信を持っており、結果として講演経験が豊富な話者は講演が得意だと意識しているという可能性が考えられる。その結果、さ入れ言葉がこの二つの要因に対して同様の振る舞いを示すということである。このように、この二つの要因は相互に関係している可能性がある。これについては、次節の多変量解析でその関係の検証を行う。

本小節では、各言語外的要因に関する分析を行った。その結果、さ入れ言葉が 1) 模擬講演よりも学会講演で、2) 学会講演では女性によって、模擬講演では男性によって、3) 生年が若い話者によって、4) 関東地方で、5) 自発性が高い講演で、6) スタイルの高い発話で、7) 講演が得意と意識している話者によって、8) 講演経験が豊富な話者によって選好されるということが明らかになった。このことは、さ入れ言葉が現在初期段階にある進行中の言語変化であり、あらたまった場面で使用される傾向があるということ強く支持する。また、性差がさ入れ言葉の変化において重要な役割を果たしていることを客観的な形で示した。更に、関東地方または中部地方で出現し、成人になって獲得された要素であり、言語獲得における仮説の反例となり得るという可能性を指摘した。

## 5. 多変量解析

前節では、言語内的要因に加え言語外的要因がさ入れ言葉の分布に影響を与えているということを明らかにした。しかしながら、言語変異において各要因がそれぞ

れ完全に独立して影響を及ぼすということはなく、相互に複雑に作用し合いながらその言語変異に対して影響を与えているはずである。そして、それらの相互作用の結果さ入れ言葉の分布が規定されるのである。例えば、各要因別分析において講演種がさ入れ言葉の分布に影響を与えているという結果が出たとしても、実はそれらは発話スタイルという下位範疇の影響が講演種という形で現れたものに過ぎないなどという可能性もある。また、影響のある要因の中でも全てが一樣に作用しているということは考え難く、例えば講演種が最も大きく、発話スタイルが2番目に影響を与えているなどというように、影響の大きさにも差がある可能性が高い。従って、さ入れ言葉の解明にはそれら要因間の関係を考慮に入れた更なる詳細な分析が必要となる。本節では、前節の要因別分析の結果を踏まえ、各要因がどの程度さ入れ言葉の分布に影響を与えているのか、更に各要因がどのような関係を持ち得るのかということ进行を明らかにする。具体的には、さ入れ言葉／正用法を従属変数とし、言語内的要因である動詞の長さ、後続要素、言語外的要因である講演種、性別、生年、居住地、自発性、発話スタイル、得手・不得手、講演経験を独立変数として、二項ロジスティック回帰による分析を行う<sup>18</sup>。分析には統計ソフト SPSS を用いた。導き出された結果の中で、各要因の持つ係数がさ入れ言葉に対する影響の大きさ（重み：weight）を表していると考えることができる。結果のうち、まず各変数間の相関行列を以下に示す。

表 19 従属変数（さ入れ言葉／正用法）、及び独立変数（各言語内的要因、言語外的要因）間の相関行列

	Constant	講演種	動詞長さ	後続要素	性別	生年	居住地	自発性	スタイル	得手不得手	講演経験
Constant	1.000	.111	-.315	-.208	-.196	-.447	-.482	-.488	-.444	.063	-.068
講演種	.111	1.000	-.022	-.117	-.310	-.216	-.087	.318	-.297	.125	-.539
動詞長さ	-.315	-.022	1.000	.108	-.012	-.022	.000	-.019	-.036	-.010	.049
後続要素	-.208	-.117	.108	1.000	.022	.003	-.076	-.054	-.043	-.066	.067
性別	-.196	-.310	-.012	.022	1.000	.060	-.041	-.012	.171	-.143	.054
生年	-.447	-.216	-.022	.003	.060	1.000	-.038	.065	.028	.004	.239
居住地	-.482	-.087	.000	-.076	-.041	-.038	1.000	-.081	.048	-.005	.047
自発性	-.488	.318	-.019	-.054	-.012	.065	-.081	1.000	.277	-.043	-.239
スタイル	-.444	-.297	-.036	-.043	.171	.028	.048	.277	1.000	-.191	-.012
得手不得手	.063	.125	-.010	-.066	-.143	.004	-.005	-.043	-.191	1.000	-.250
講演経験	-.068	-.539	.049	.067	.054	.239	.047	-.239	-.012	-.250	1.000

まず、表 19 の各要因間の相関に関して考察をする。言語変異においては言語外的要因同士では交互作用が見られる一方で、言語内的要因同士では大抵交互作用は

<sup>18</sup> 今回のデータは正用法に対して、さ入れ言葉の割合が非常に低いいわゆるレアケースであるため、ロジスティック回帰により完全に正確な結果を得ることは難しいが、ある程度予測を立てる上では意味のある結果が得られると考える。

なく互いに独立である。更に、言語外的要因と言語内的要因間に関しても独立であるという傾向がある (Labov 1982)。この主張が正しければ、言語内的要因 (動詞の長さ、後続要素) は互いに相関がなく、各言語外的要因とも相関がないはずである。実際、表 19 が示すように動詞の長さ、後続要素ともに相関はなく、言語外的要因との相関もない。よって、各言語内的要因が独立した形でさ入れ言葉の分布に作用しているということであり、上記の主張は支持される。また、前節で議論をした講演種とその下位範疇と考えられる言語外的要因に関してであるが、講演の自発性、発話スタイル、講演経験ともに講演種との相関がある程度見られる。更に、講演の自発性と発話スタイルも僅かに相関を示している。このことから、下位範疇として講演の自発性、発話スタイルがまず相互作用をし、それに講演経験が加わり、これらが講演種という上位の範疇に作用をし、さ入れ言葉の分布に影響を与えていると考えることができる。また性別も講演種との相関を示しているが、これは学会講演では男性が支配的であるという事実が反映されたものであると推測される。更に、講演経験は得手・不得手とも若干の相関を示しており、4.2.8 節における仮説の通り、両要因は相互に関係している。これは講演経験が豊富な話者は講演が得意だと意識しているということが反映されたものであると言える。その結果としてさ入れ言葉がこの二つの要因に対して同様の振る舞いを示しているという説明ができる。これらをモデル化すると以下のようなになる。

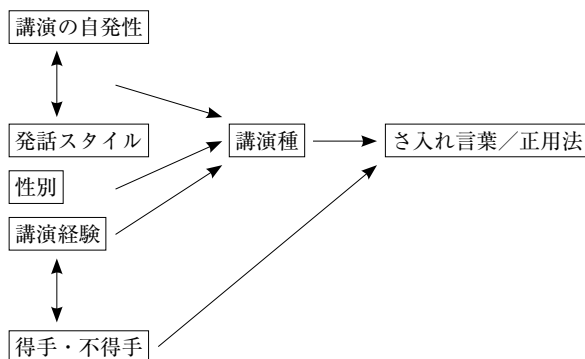


図4 講演の自発性、発話スタイル、性別、講演経験、得手・不得手、及び講演種間の関係に関するモデル

次に各要因がどの程度さ入れ言葉の分布に影響を与えているのかということに関して議論する。以下に方程式中の変数を示す。

表 20 ロジスティック回帰の結果

	方程式中の変数					
	B	標準誤差	Wald	自由度	有意確率	Exp(B)
講演種	4.843	2.603	3.461	1	.063	126.829
動詞の長さ	.185	.321	.331	1	.565	1.203
後続要素	2.608	.556	21.964	1	.000	13.571
性別	1.932	.672	8.258	1	.004	6.906
生年	.292	.170	2.957	1	.086	1.339
居住地	-.747	.329	5.144	1	.023	.474
講演の自発性	.643	.676	.904	1	.342	1.902
発話スタイル	.870	1.348	.417	1	.519	2.388
得手・不得手	-.417	.591	.497	1	.481	.659
講演経験	.511	.803	.405	1	.525	1.666
講演種 by 性別	-3.307	.942	12.319	1	.000	.037
講演経験 by 講演種	-1.544	.920	2.818	1	.093	.213
講演経験 by 得手不得手	1.339	.940	2.028	1	.154	3.816
発話スタイル by 自発性	-.208	.289	.515	1	.473	.812
発話スタイル by 講演種	-.809	.742	1.187	1	.276	.446
講演種 by 自発性	-.066	.448	.022	1	.883	.936
定数	-8.889	3.672	5.859	1	.016	.000

Nagelkerke R<sup>2</sup>=.251<sup>19</sup>

まず、表 20 における用語の簡単な説明をする。「B」は回帰係数であり、ロジスティック回帰方程式での各要因が持つ係数を表している。プラスであればさ入れ言葉の現れる確率が高く、マイナスであればさ入れ言葉の現れる確率が低いということである。この値が大きいほどその要因は大きな影響を与えているということになる。「標準誤差」とは（実際のデータではなく）推定値における標準偏差のことである。この値が大きいとバラつきが大きいということになる。「Wald」とは Wald 統計量である。これは Wald 検定によって回帰係数が有意か有意でないかの判断をする際の検定統計量である、その有意確率が 5 列目に示されている。「Exp(B)」とはオッズ比であり、例えば後続要素の Exp(B) は 13.571 となっているが、後続要素が受益形式である場合は、それ以外の場合よりも 13.571 倍さ入れ言葉が現れる確率が高くなることを示している。この値が 1 に近いとその要因の影響は小さいとい

<sup>19</sup> Nagelkerke R<sup>2</sup> はモデルのフィットを表し、この値が 1 に近いほどデータをよく説明しているモデルとなる。



うことになる。また「講演種 by 性別」などの変数は、交互作用が認められる変数を組み合わせた「交互作用項」である。この交互作用項を投入することにより、交互作用を考慮に入れた分析を行うことができる。

表 20 が示すように、まず有意確率から講演種、後続要素、性別、居住地、講演種 by 性別、に関してはかなりの確率で有意である。また、生年、講演経験 by 講演種に関しても 5% 域には達しないが、有意差への傾向が見られる。つまり、これらの要因はさ入れ言葉の分布に影響を与えていると言える。一方、動詞の長さ、講演の自発性、発話スタイル、得手・不得手、講演経験、講演経験 by 得手不得手、発話スタイル by 自発性、発話スタイル by 講演種、講演種 by 自発性に関しては有意ではない。つまり、影響を与えているとは言えない。ここで注目したいのは、例えば講演経験、得手不得手はそれ自体では有意ではないが、講演経験 by 得手不得手という交互作用項とした場合、有意差への傾向が高まる。従って、これらの要因は単独ではなく、互いに作用し合いながらさ入れ言葉の分布に働きかけていると言える。次に各要因の持つ係数（表 20 の B）は、影響を与えている講演種、後続要素、性別、居住地、講演種 by 性別、生年、講演経験 by 講演種の中でも講演種が最も大きく（重みがあり）、以下講演種 by 性別、後続要素、性別、講演経験 by 講演種、居住地、生年の序列となっている。つまり、講演種が最も大きく、以下講演種 by 性別、後続要素、性別、講演経験 by 講演種、居住地、生年の順でさ入れ言葉の分布に影響を与えていると考えることができる。

また、有意である言語外的要因の中で、交互作用項が高い係数を示していることから、図 4 に示したように各要因が、単独ではなく互いに作用し合い、最終的に講演種という形でさ入れ言葉に影響を与えているということが読み取れる。つまり、講演種が下位範疇である各要因とどのような関係にあるかという構造が明らかになったと言える。更に、講演種の極めて高い係数は、その下位範疇である様々な言語外的要因がこの要因に反映され、最終的に講演種という形でさ入れ言葉に影響を与えているという上記の議論を支持するものである<sup>20</sup>。CSJ において、講演種は第一次的な要因であるが、今回の分析においてその特徴が顕著に現れた。このように、要因別分析のみでは明らかにならなかった各要因の与える影響に関する詳細な知見を得ることができた。

## 6. 結論

本研究では CSJ を用い、さ入れ言葉の分布に関する言語内的要因、言語外的要因の影響に関する分析を行った。その結果、要因別分析では 1) 言語内的要因に関して、さ入れ言葉はほぼ短い動詞に限られ、「-させていただく、-させてもらう」という句へと文法化しつつあるという先行研究の主張を支持する結果が得られた。

<sup>20</sup> 交互作用項を投入せずにロジスティック回帰分析を行った場合、講演種は非有意となることから、各要因間の交互作用の存在が確認できる。

また、2) さ入れ言葉が性別、生年、居住地、及びスタイルに関する言語外的要因の影響を受けており、それが現在初期段階にある進行中の言語変化であるということ、そしてあらたまった場面で使用される傾向があるという経験的知識を強く支持する結果となった。また、3) 関東地方または中部地方で出現し、大人になって獲得された要素であるという可能性を指摘した。この結果が言語獲得に関する主張の反例となる可能性は十分にある。更に、4) 講演種とその下位範疇である講演の自発性、発話スタイル、性別、講演経験との関係、及び講演経験と得手・不得手との関係に関するモデルを提案し、多変量解析を行い、各要因が与える影響の大きさを明らかにした。そして講演種が最も大きく、以下講演種 by 性別、後続要素、性別、講演経験 by 講演種、居住地、生年の順でさ入れ言葉の分布に影響を与えているということがわかった。また、言語変異においては言語外的要因同士では交互作用が見られる一方で、言語内的要因同士では互いに独立であり、言語外的要因と言語内的要因間に関しても独立であるという傾向があるという主張 (Labov 1982) を支持する結果が得られた。このように、これまでなされて来なかった言語外的要因の詳細な分析、及び多変量解析により、さ入れ言葉の性質に関する知見が新たに得られた。また、CSJにおける各要因はこれまで並列的に捉えられており、要因間の関係や構造に関する議論は管見の限りなされて来なかったが、本研究ではそれらの関係や構造が明らかになった。

さ入れ言葉の変化は、スタイルが高いところから始まる変化でありながら、社会的には矯正への圧力があり、「誤り」と見做されている(正用法が正しい形であるという概念がまだ通用している)。従って、さ入れ言葉の変化は「上からの変化」(change from above) であると考えられる。更に、上からの変化であるためにあらたまった場面で女性がより高いさ入れ率を示すという結果は、「上からの変化において、女性は男性よりも新たな威信のある形式を好む」という Principle Ia (Labov 1990) とも一致する。また、あらたまった場面でより「丁寧度」の高い形式を使用しようとした結果、誤って正用法ではなくさ入れ言葉を使用したという過剰訂正 (Hypercorrection, Labov 1966) がさ入れ言葉出現の原因であるという推測ができる。Labov (1966) の過剰訂正の例は社会階層とスタイルの交互作用を示したものであると考えられるが、今回の場合、社会階層に相当するのが性別であるとする、正に同じ現象が起きていると言える。

さ入れ言葉は、革新形がくだけた発話を選好する「下からの変化」(change from below) と異なり、上からの変化であるために、あらたまった場面で使われる傾向を示すのであるが、このことに関して、意味・機能的な説明を試みたい。Okada (2003) はさ入れ言葉は丁寧さを強調する機能を有する (honorific booster) と主張している。この主張が正しければ、さ入れ言葉が丁寧さを強調する機能を有するために、くだけた発話よりもより丁寧な表現が要求されるであろうあらたまった発話を選好するということになり、自然な説明を与えることができる。次に 4.1.2 節で述べたように、受益形式 (-ていただく、-てもらう型) は丁寧な形式であるため、あらたまっ

た場面で使用されるということが推測される。文法化によってその受益形式とまともとなり現れるさ入れ言葉は、やはりあらたまった場面で使用されるということである。

今回のさ入れ言葉の数量的分析は言語変化の初期を捉えている。これまでの言語変化の研究では、多くがある程度変化が進行した段階での分析であり（例、ら抜き言葉, Matsuda 1993; ガ行鼻濁音化, Hibiya 1995）、言語変化の初期が扱われたことはなかった。さ入れ言葉に関する今後の継続的な分析により新たな知見が得られる、また、そのことが言語変化の全貌の解明に資する可能性は高い。

最後に、CSJが持つ言語コーパスとしての可能性について触れておきたい。上記のように、CSJには話者に関する属性や発話場面に関する情報など、研究用情報が豊富に付与されている。従来、これらの情報は調査の現場で取得しなければならず、その入手は多大な負荷を伴うものであった。この点、CSJに付与された多くの研究用情報は、調査における負担を大幅に軽減するだけでなく、今回の言語外的要因に関する詳細な分析など、これまで極めて実現しにくかった研究を容易に可能にするものであり、これによって言語研究の可能性が広がることは確実であると言える。

## 参 照 文 献

- Bayley, Robert (2002) The quantitative paradigm. In: J.K. Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds.) *The handbook of language variation and change*, 117–141. Oxford: Basil Blackwell.
- 陳 秋燕 (2002) 「日本語の敬語の変化「さ入れ言葉」についての考察—電子資料に基づく分析—」修士論文, 東京外国語大学.
- Eckert, Penelope (2004) The meaning of style. In: Wai Fong Chiang, Elaine Chun, Laura Mahalingappa and Siri Mehus (eds.) *Proceedings of the eleventh annual symposium about language and society—Austin, Texas linguistic forum 47*, 41–53.
- Hibiya, Junko (1995) The velar nasal in Tokyo Japanese: A case of diffusion from above. *Language variation and change* 7(2): 139–152.
- Hopper, Paul J. and Elizabeth Closs Traugott (1993) *Grammaticalization*. Cambridge: Cambridge University Press.
- 井上史雄 (2003) 『日本語は年速一キロで動く』東京：講談社現代新書.
- 籠宮隆之 (2004) 『記録票データ・対話記録票データ・講演者属性データ・対話参加講演者の講演一覧の解説』『日本語話し言葉コーパス』附属電子文書.
- 籠宮隆之・山住賢司・楨洋一 (2004) 『印象評定データの概要』『日本語話し言葉コーパス』附属電子文書.
- Labov, William (1966) Hypercorrection by the lower middle class as a factor in linguistic change. In: William Bright (ed.) *Sociolinguistics: Proceedings of the UCLA sociolinguistics conference, 1964*, 84–113. The Hague: Mouton.
- Labov, William (1972) *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, William (1982) Building on empirical foundations. In: Winfred P. Lehmann and Yakov Malkiel (eds.) *Perspectives on historical linguistics*, 17–92. Amsterdam: John Benjamins.
- Labov, William (1990) The intersection of sex and social class in the course of linguistic change. *Language variation and change* 2: 205–254.
- Labov, William (1994) *Principles of linguistic change: Internal factors*. Oxford: Basil Blackwell.
- Labov, William (2001) The anatomy of style-shifting. In: Penelope Eckert and John R. Rickford (eds.) *Style and sociolinguistic variation*, 85–108. Cambridge: Cambridge University Press.
- 前川喜久雄 (2004) 『『日本語話し言葉コーパス』の概要』『日本語科学』15: 111–133.
- Matsuda, Kenjiro (1993) Dissecting analogical leveling quantitatively: The case of the innovative

- potential suffix in Tokyo Japanese. *Language variation and change* 5: 1–34.
- 松田謙次郎 (2004) 「言語資料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.
- Meillet, Antoine (1912) L'évolution des formes grammaticales. *Scientia (Rivista di scienza)* 12(26). Reprinted in Antoine Meillet, ed. (1958) *Linguistique historique et linguistique generale*, 130–148. Paris: Champion.
- Milroy, Lesley and Matthew Gordon. (2003) *Sociolinguistics: Method and interpretation*. Oxford: Basil Blackwell.
- Okada, Judy. (2003) Recent trends in Japanese causatives: The *sa*-insertion phenomenon. *Japanese/Korean linguistics* 12: 28–39.
- 佐野真一郎 (2008) 「さ入れ言葉—国会会議録を使った数量的分析—」松田謙次郎 (編) 『国会会議録を使った日本語研究』. 東京: ひつじ書房. (印刷中)
- 新村 出 (編) (1998) 『広辞苑』, 第5版. 東京: 岩波書店.
- Weinreich, Uriel (1968) *Languages in contact: Findings and problems*. The Hague: Mouton.

著者連絡先: [受領日 2007年10月11日  
102-8554 東京都千代田区紀尾井町7-1 最終原稿受理日 2007年12月17日]  
上智大学大学院外国語学研究科  
shinichirosano@gmail.com

## 付録：さ入れ言葉の全 42 例

講演 ID	さ入れ言葉
A01M0115	呼ばさしていただきます
A01M0838	呼ばさしていただきますが
A01M0889	呼ばさしていただきます
A02F0800	入らさしていただきます
A02M0854	読まさせていただけますと
A03M0016	終わらさしていただきます
A03M0583	動かさして
A04M0229	終わらさせていただけます
A05M0858	省かさしていただきました
A06F0210	飛ばさしていただきまして
A07F0366	やらさせていだいたんですけれども
A07F0366	配らさせていだいたんですけれども
A08F0323	配らさして
A10M0580	振り返らさしていただきます
M01M0007	組み合わせ
M01M0011	終わらさせていたきたいと思いますけれども
S00F0595	終わらさせていただきます
S01F1522	かわいがらさしていただいています
S01M0460	終わらさしていただきます
S01M0464	終わらさせていたきたいと思いますけれども
S01M0851	終わらさしていただきます
S02M0478	終わらさしていただきます
S02M0772	終わらさしていただきます
S02M1331	終わらさせていただきます
S02M1372	終わらさしていただきます
S03M0081	終わらさしていただきます
S03M0976	終わらさしていただきます
S04M0379	終わらさしていただいでよろしいでしょうか
S04M0880	就かさしていただき
S04M1251	終わらさしていただきます
S04M1451	飲まさしていただきました
S05M0626	終わらさしていただきます
S06M0784	終わらさしていただきます
S06M0872	終わらさしていただきます
S06M1125	回らさしていただきました
S07F1336	作らさせる幼稚園だったんですけれど
S07M0714	終わらさしていただきます
S07M1052	終わらさしていただきます
S07M1366	選ばさせていただきました
S09M0974	終わらさしていただきます
S09M1274	終わらさしていただきます
S11M1257	動かさしてくれと言っても

**Abstract****A Quantitative Analysis of *Sa*-insertion in *Corpus of Spontaneous Japanese***

SHIN-ICHIRO SANO

*Graduate student, Sophia University*

This research statistically investigates the recent synchronic variation in Japanese causative forms involving a process called *Sa*-insertion, which adds an extra *-as-* to the causative morpheme, as in *yar-asase-ru* vis-à-vis the standard *yar-ase-ru* ‘let someone do’ and *hair-asase-ru* vis-à-vis the standard *hair-ase-ru* ‘let someone enter’. Although previous studies have elucidated the language-internal properties of this phenomenon especially with respect to the length of stem verbs and the types of the following elements (Sano 2008, among others), language-external factors such as gender and style, as well as the interaction of those language-internal factors, remain to be explored.

The present paper aims to verify the claims of the previous studies and reveal the language-external factors as well as the effects of the interaction of major factors governing the distribution of *Sa*-insertion. For this purpose, *Corpus of Spontaneous Japanese (CSJ)* is employed which has rich annotations concerning language-external factors. An exhaustive examination of *CSJ* brought forth a total of 42 causative forms with *Sa*-insertion, as opposed to a total of 1,498 standard causative forms; thus, the rate of *Sa*-insertion amounts to 2.73%. The data are then subjected to a factor by factor analysis, followed by a multivariate analysis.

The result shows that the distribution of *Sa*-insertion is significantly affected by language-external factors in addition to language-internal factors, and that the factors that govern the distribution interact with each other in an intricate way.