

# インターネット上の自発述語 におけるサ抜き現象\*

佐々木 冠<sup>†</sup>

キーワード：自発、サ抜き、韻律的最小性、インターネット

## 1 はじめに

筆者は、これまで、北海道方言における自発述語の統語的性質や音韻的性質についてネイティブスピーカーへのアンケートを通して調査を行ってきた。しかしながら、アンケート調査には、北海道方言話者が自然に産出した例文を収集できないという問題点がある。

この問題を解消する方法としては、北海道方言話者の自然談話を録音しその中から自発構文を抽出する方法、北海道方言の話者が作成した電子的な文書から自発構文を抽出する方法が考えられる。前者は、母語話者の産出する用例を収集する方法として伝統的なもので、イントネーションなどの文法と密接にかかわる音声情報が付随する情報の豊かな資料を作るうえで欠かせない方法である。その一方で、短時間に大量のデータを利用できるようにすることが困難であるという欠点がある。一方、後者は、イントネーションなどの文法と密接にかかわる音声情報が欠落するという欠点はあるものの、文字化された資料が多い言語の場合、短時間で大量のデータを利用できるようにすることが可能であるというメリットがある。

---

\* (14) のグラフに示したデータは、2008 年度前期に札幌学院大学で開講された全学共通科目「人間の言語のしくみ」受講者の協力のもと収集したものである。この場を借りて、協力してくださった受講者に謝意を表したい。

<sup>†</sup> 札幌学院大学

方言の語形に関しては、後者の方法がとられることはあまりなかった。これは、方言を使った文学作品やジャーナリズムの文章が標準語のそれと比べて少なかったことによるものと思われる。しかし、状況は変化している。インターネットが普及したことにより文字化された発言を個人が自由に発信できるようになった。このような状況の下では、文字化された方言の量がかつてより増えている可能性がある。試験的に自発語形を Yahoo や Google といった検索サイトで検索したところ、かなりの量の自発語形がインターネット上に存在することがわかった。そこで、母語話者が自然に産出した自発構文の用例を収集するためにインターネット上のデータを利用することにした。

本稿は、インターネット上の自発語形調査の中間報告である。本稿の構成は次のとおりである。第2節では、調査方法を示す。そして、その方法に基づいて行ったデータ収集の中間結果を第3節で示す。第4節では、収集したデータを利用し、佐々木 (2007) および Sasaki (2008) で行ったサ抜き語形に関する一般化の検証を行う。第5節では今後の課題について論じる。

## 2 調査方法：データ収集の方法

2007 年前半に Yahoo および Google を使って「押ささる」「押ささって」「押ささり」などの自発語形を検索したところ、一つひとつの語形に数百件の検索結果が出た。これを手作業で収集整理することは不可能と判断し、研究室にサーバを設置し、業者に Yahoo の検索結果を自動的に蓄積するプログラムを作ってもらった。

2007 年 7 月 31 日から自発語形収集のためのプログラムを走らせ、現在もこのプログラムによる語形収集を継続している。このプログラムは、登録しておいた語形が Yahoo の検索結果にあった場合、検索されたウェブページの語形の前後 100 字を収集してきて蓄積するものである。

プログラムに登録する語形は、次のようなかたちでリストアップした。『計算機用日本語基本動詞辞書 IPAL』にある他動詞のうちサ変動詞と最重要動詞を検索する動詞とした。これらの動詞の自発語形を作り、さらに自

発語形の活用形を作った。その際、五段活用動詞で /s/ で終わる語根の動詞については、サ抜き語形と非サ抜き語形の両方を作った。また、サ変動詞については、リスト作成の前に北海道方言の母語話者から「～する」に対応する自発語形が「～ささる」とあるという情報を得たので、「～ささる」の活用形を作った。合計で 3842 語形をプログラムに登録し自発構文の収集を行った。

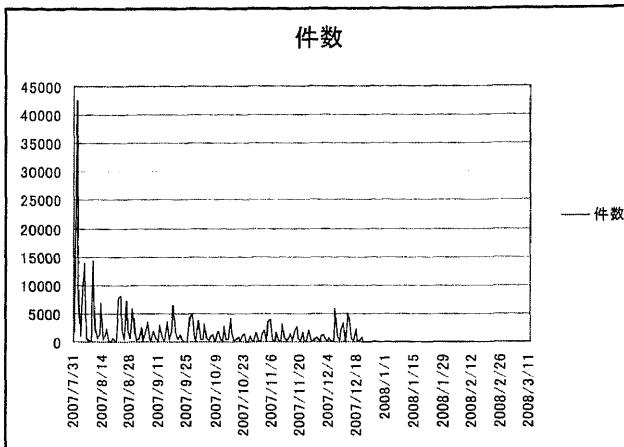
プログラムは収集した語形に関してカンマで区切られた五つのフィールドからなるレコードを作る。五つのフィールドは「検索語形」「ウェブページのタイトル」「ウェブページの URL」「検索された語形の例」「収集した日時」からなる。

このプログラムによる語形収集は現在も継続中である。次節では、2008 年 3 月 11 日までに収集した語形の概要を報告する。

### 3 中間報告

2007 年 7 月 31 日から 2008 年 3 月 11 日までの 7 ヶ月で収集した語形の総数は、277,600 であった。日ごとに収集した自発語形の件数を示したグラフが(1)である。2007 年 8 月 1 日に 42,530 件の自発語形を収集した後、収集する語形の件数は減り始め、2007 年 12 月 21 日に 0 件を記録してから、12 月 31 日に 1 件の語形を変策したほかは収集した語形の件数はゼロであった。

## (1) 収集した語形の日ごとの件数



収集したデータの中には、自発語形の用例としては使えないものも含まれていた。具体的には、次のようなデータである。

## (2) 検索語形の内部に記号が混ざったレコード

検索語形「押さら」に対して「押）さら」のような文字列が収集されたレコードが存在する。このレコードを排除するため、`gawk`で「検索された語形の例」のフィールドに「検索語形」が含まれているレコードのみを残すプログラムを作った。`gawk`のプログラム作成に当たっては奥田統己氏から教示を受けた。

## (3) 古語を拾ってきてしまったレコード

Yahooで「外さる」を検索すると「外さるる」のような古語の連体形などを含むウェブページがヒットする。北海道方言とは異なる言語体系のデータなので、こうした要素を含むレコードは削除する必要がある。一定のパターンを見つけた後は、`gawk`のプログラムで類例を削除できるが、半ば手作業の側面があり、こうした語形を含むレコードを削除する作業にはかなり時間を要した。

(3) のタイプのレコードの削除がかなりやっかいであり、北海道方言とは無関係なデータを完全には削除し切れていないのが、現状である。しかし、これから研究を進めることにより、(3) のタイプのレコードを完全に排除することは可能と思われる。ただ、(3) のタイプのデータを完全に排除しても、残された語形が北海道方言の語形といえるかどうかについては、別な形で検証が必要である。収集したデータを概観しただけでも、青森県の方言として記載されてあるデータもある。これは、北海道方言と北東北方言の連続性を考えると当然といえる。集めた語形が北海道方言のデータであることを検証するためには、ウェブページにその言語データを記載した人間の属性を確認する必要があるが、それをどのようなかたちで効率的に実現できるか、現在検討中である。書き手の属性の確認ができるようになるまでは、自発語形収集プログラムで集めたデータは、北海道方言の自発語形のデータというよりは、インターネット上の自発語形のデータということになる。本稿のタイトルが北海道方言を冠したものではなく、「インターネット上の」という限定になっているのはこのためである。

#### 4 分析：サ抜き語形形成の条件

自発語形を含む用例の収集は、最終的には自然に産出された自発構文の意味的特徴や統語的特徴を解明することを目的とするものである。自発述語と共起する名詞句の内在的な意味や格形式に関して収集した用例から一般化を行うことも可能なはずであり、将来は是非それを実現したい。しかし、残念ながら時間的な制約からこの報告書の作成期限までにそれを実現することは不可能である。しかし、自発述語の形式そのものに関する傾向性を収集したデータから導くことは現時点でもある程度可能である。そこで、サ抜き語形形成の条件を収集したデータから考察したい。

筆者はこれまで、佐々木 (2007) および Sasaki (2008) において、アンケート調査をもとに、自発述語のサ抜き語形の形成に関して分析を行ってきた。「サ抜き」とは、自発述語の通常の形態法から期待される語形が「～ささる」である例で、「さ」が一つ脱落し、「～さる」になる現象である。

北海道方言の自発述語は動詞語根に自発接尾辞 /rasar/ が後接する構成になっている。そして、動詞語根と自発接尾辞の境界に子音連続が生じる場合、自発接尾辞の先頭の子音が脱落する。これは、子音連続を回避した結果であり、受動接尾辞や使役接尾辞などの態接尾辞でも見られる現象である。動詞の語根が/s/で終わる場合、通常の形態法では、/hagemas-rasar-ru/ →\*[hagemasaru]のように、「～ささる」で終わる形式が生じるが、アンケート調査を行うと、特殊な条件下以外では、実際には[hagemasaru]のように「さ」が一つ脱落した形式が現れることが多い。特殊な条件とは、語根の母音の数に関するものである。動詞語根に含まれる母音が一つ以下の場合には、サ抜き現象がブロックされ、「～さる」ではなく「～ささる」という語形になる傾向がある。Sasaki (2008) は、このサ抜き現象のブロックを韻律的最小性 (Ito 1990) によるものとする分析を行っている。

アンケートに現れたサ抜き現象に関する傾向性が、インターネットで収集したデータにも現れるか否か、以下で検証を行うことにする。

まず、インターネットで検索した /s/ で終わる語根をベースにした自発述語のうち、語尾ごとに検索された用例を示すことにする。収集した撥音便の語形を示したものが (4) である。[] 中の数字は当該語形の件数である。この語形には「ない」「ねえ」といった否定の接辞が後接する。サ抜き現象が起きていない語形、つまり「～ささん」で終わる語形は、語根に含まれる母音の一つだけの和語の動詞とサ変動詞に限定されている。語根に含まれる母音の一つだけの動詞の中には、非サ抜き語形のみなもの (出ささん、干ささん) と非サ抜き語形とサ抜き語形が共存しているもの (差(さ)さん、押(さ)さん、刺(さ)さん、消(さ)さん) があつた。撥音便に関しては、語根に含まれる母音の一つ以下の動詞ではサ抜き現象がブロックされることがあるという一般化が成立する。

#### (4) 撥音便 (487 例)

ずらさん[1]、引き起こさん[1]、引き返さん[1]、隠さん[1]、益さん[2]、汚さん[1]、押ささん[176]、押さん[4]、卸さん[1]、過ごさん[2]、外さん[2]、覚さん[1]、覚まさん[1]、干ささん[5]、起こさん[2]、起さん[1]、減らさん[4]、現わさん[4]、差ささん[1]、差さん[5]、刺ささん[21]、刺

さん[90]、指さん[5]、出ささん[1]、出さん[1]、消ささん[1]、消さん[3]、  
 伸ばさん[1]、浸さん[1]、醒さん[1]、捜さん[1]、挿さん[108]、増やさん  
 さん[2]、探さん[3]、倒さん[2]、動かさん[1]、反らさん[1]、反映ささん  
 [2]、飛さん[1]、沸かささん[2]、返さん[1]、勉強ささん[5]、暮らさん[3]、  
 目指さん[1]、戻さん[2]、落とさん[1]、乱さん[1]、冷やさん[1]、話さ  
 さん[2]、惑わさん[3]、騙さん[3]

語根に含まれる母音が一つ以下の動詞ではサ抜き現象がブロックされることがあるという一般化は未然形に関しても成立する。(5)の未然形の例を参照。(4)の撥音便形の場合でも(5)の未然形の場合でも同様だが、語根に含まれる母音が一つの動詞の中にはサ抜き現象がブロックされない例もある。ただ、サ抜き現象のブロックが生じるのが語根に含まれる母音の一つに限られていることには変わりはない。

#### (5) 未然形 (317例)

押ささら[20]、荒さら[9]、差さら[16]、刺ささら[9]、刺さら[203]、指  
 ささら[5]、指さら[7]、射さら[1]、消ささら[1]、蒸さら[2]、挿さら[41]、  
 直さら[2]

次に已然形のデータを示す。収集したデータに含まれる已然形の例は少ない。また、語根に含まれる母音が一つの動詞でもサ抜き現象がブロックされている例が見当たらず、「満たす」のような語根に母音を二つ含む動詞でサ抜き現象がブロックされている例があった。ここでは、サ抜き現象のブロックと語根の母音の数に関する一般化が成立していない。なお、已然形に関してはサ変動詞の例が収集されていなかった。

#### (6) 已然形 (22例)

覚まされ[1]、及ぼされ[1]、空かされ[1]、降され[2]、差され[1]、射され  
 [1]、取消され[1]、挿され[10]、反され[1]、満たさされ[1]、満たされ  
 [1]、目指され[1]

連用形に関しては、語根に含まれる母音が一つ以下の動詞ではサ抜き現象がブロックされることがあるという一般化が成立する。(7) の例を参照。連用形に関してもサ変動詞は収集されていなかった。

#### (7) 連用形 (1749 例)

押ささり[62]、押さり[3]、過ぎさり[1]、過さり[1]、干ささり[3]、干さり[3]、差さり[206]、刺ささり[249]、刺さり[456]、指ささり[1]、指さり[40]、施さり[3]、試さり[1]、写さり[1]、射さり[349]、取り消さり[1]、蒸さり[21]、挿ささり[7]、挿さり[279]、増さり[45]、注さり[1]、超ささり[1]、超さり[2]、飛ばさり[1]、崩さり[4]、冒さり[1]、満たさり[2]、励まさり[4]、話さり[1]

命令形は7例しかなく、検索した語形のうち収集できた例が最も少なかった。なお、このうち「殺さろ」と「潰さろ」で収集された例は「実の母親にでも殺さろ」「不埒な言動を弄する小集団を潰そうとして逆に潰さろ」であり、それぞれ「殺されろ」「潰されろ」の「れ」が抜けた脱字の例である可能性がある。また、「刺さろ」は「指す」の相対自動詞である「刺さる」の活用形である可能性がある。相対自動詞の活用である可能性は、(4) から (6) に示した語根の母音が一つの動詞についても当てはまる。(8) には、サ抜き現象がブロックされた例が全くないが、これは、「刺さろ」が相対自動詞の活用形であるためかもしれない。

#### (8) 命令形 (7 例)

殺さろ[1]、刺さろ[5]、潰さろ[1]

次に已然形+「ば」の縮約形「～さりゃ」の例を(9)に示す。ここでもサ抜き現象のブロックされている例は見られない。ただ、(8)の場合と異なり、相対自動詞であるためということでは説明できない例もある。「押さりゃ」は「押されば」の縮約形である。他動詞「押す」には、少なくとも現代語では、「押さる」という相対自動詞はない。そうするとサ抜き現象が



語根の母音が一つの動詞にも適用された例と見なければならぬ可能性がある。

(9) 「～さりゃ」 (45 例)

隠さりゃ[1]、越さりゃ[1]、押さりゃ[4]、壊さりゃ[1]、降ろさりゃ[1]、差さりゃ[6]、済まさりゃ[1]、殺さりゃ[6]、治さりゃ[1]、出さりゃ[1]、挿さりゃ[3]、探さりゃ[1]、潰さりゃ[1]、倒さりゃ[3]、動かさりゃ[2]、燃やさりゃ[1]、返さりゃ[1]、崩さりゃ[2]、満たさりゃ[2]、流さりゃ[2]、話さりゃ[2]、騙さりゃ[2]

(10) に過去形の用例を示す。語根に含まれる母音が一つ以下の動詞ではサ抜き現象がブロックされることがあるという傾向性はあるが、例外もある。「溶かささった」「話かさった」のように語根に含まれる母音の数が2以上の動詞の場合でもサ抜き現象がブロックされている例がある。ただし、このような例外は少ない<sup>1</sup>。三つの例外を非サ抜き語形と比べると、「溶かささった」(1例)、「溶かさった」(1例)、「話かさった」(3例)、「話さった」(7例)であり、用例が一つしかない「溶かす」から派生した語形以外では、サ抜き語形のほうが多い<sup>2</sup>。

(10) 過去形 (2,064 例)

移さった[2]、隠さった[1]、映さった[1]、押ささった[338]、押さった[27]、過ごさった[1]、解かさった[1]、回さった[2]、壊さった[2]、外さった[1]、干ささった[29]、干さった[26]、帰さった[4]、記さった[3]、起こさった[10]、起さった[1]、減らさった[1]、荒ささった[1]、荒さった[2]、差ささった[1]、差さった[265]、済まさった[1]、殺さった[9]、散らさった[1]、残さった[2]、刺ささった[253]、刺さった[410]、指ささった[13]、指さった[81]、実現ささった[2]、写さった[2]、射ささ

<sup>1</sup>佐々木 (2008) で例外として挙げていた「飛ささった」(5例)は張飛という三国志の登場人物の名前の後に「ささった」がついたものであったことがわかったので、除外した。

<sup>2</sup>「話かさった」は「話(はなし)する」というサ変動詞の自発形過去である可能性がある。そうであればサ抜き現象がブロックされるのはアンケート調査をもとにした一般化からむしろ期待される場所である。下に出てくる「話かさって」も同様である。

た[1]、射さった[45]、出さった[21]、消ささった[6]、消さった[4]、蒸ささった[4]、蒸さった[40]、浸さった[1]、推ささった[1]、挿ささった[5]、挿さった[295]、増さった[55]、注さった[1]、超ささった[12]、直さった[1]、通さった[4]、渡さった[2]、倒さった[5]、発さった[1]、発表ささった[1]、飛ばさった[5]、飛ばさった[6]、沸かさった[6]、返さった[4]、暮さった[1]、崩さった[1]、鳴らさった[2]、滅ぼさった[1]、溶かささった[1]、溶かさった[1]、落とさった[6]、流さった[8]、励まさった[5]、話ささった[3]、話さった[7]、騙さった[14]

「～て」形の例を (11) に示す。この形式は、今回分析の対象からはずした終止・連体形について収集された用例が多かった。ここでも、語根に含まれる母音の一つ以下の動詞ではサ抜き現象がブロックされることがあるという傾向性はあるが、例外もあった。(10) の場合と異なるのは、「外ささって」「探ささって」「返ささって」のような対応するサ抜き語形がない長い(語根が母音を二つ以上含んでいる)例があった点である。サ抜き語形と非サ抜き語形が共存する例をみると、「隠ささって」(4例)「隠さって」(2例)では、サ抜き現象がブロックされた例のほうが多かったが、他では、サ抜き現象が適用されている例のほうが多かった<sup>3</sup>(「回ささって」(1例)「回さって」(5例)、「乾ささって」(1例)「乾さって」(7例)、「直ささって」(1例)「直さって」(14例)、「話ささって」(1例)「話さって」(13例))。

(11) 「～て」 (3,191 例)

ずらさって[3]、もてなさって[7]、遺さって[1]、隠ささって[4]、隠さって[2]、汚さって[2]、押ささって[325]、押さって[64]、解さって[1]、回ささって[1]、回さって[5]、外ささって[2]、確保ささって[1]、乾かさって[1]、乾ささって[1]、乾さって[7]、干ささって[111]、干さって[194]、記さって[14]、記録ささって[8]、起こさって[4]、許さって[5]、協力さ

<sup>3</sup>佐々木(2008)で挙げていた「乾かささって」は、筆者自身が2005年に行ったアンケート調査表をウェブページにアップロードしていたものを検索プログラムが拾ってきたものであったため、外した。

さって[2]、強調ささって[9]、空かさって[6]、減らさって[2]、荒さって[34]、差ささって[3]、差さって[340]、済まさって[1]、晒さって[6]、散らさって[1]、残さって[7]、刺ささって[293]、刺さって[445]、指ささって[40]、指さって[192]、指定ささって[2]、試さって[1]、示さって[1]、射さって[65]、出ささって[1]、出さって[26]、消ささって[3]、消さって[12]、照らさって[1]、紹介ささって[4]、蒸さって[126]、伸ばさって[1]、正さって[9]、挿ささって[54]、挿さって[340]、増さって[232]、促さって[1]、探ささって[1]、注さって[4]、超ささって[4]、超さって[3]、直ささって[1]、直さって[14]、通さって[3]、潰さって[2]、提出ささって[1]、展示ささって[5]、転がさって[1]、渡さって[10]、努力ささって[1]、当選ささって[2]、透かさって[2]、動かさって[6]、擦ささって[7]、剥がさって[1]、発見ささって[2]、発表ささって[1]、反らさって[1]、反映ささって[5]、飛ばさって[13]、評価ささって[1]、沸かさって[4]、返ささって[1]、勉強ささって[4]、崩さって[7]、鳴らさって[4]、滅さって[1]、目指さって[1]、落さって[1]、離さって[1]、流さって[21]、冷やさって[4]、話ささって[1]、話さって[13]、騙さって[13]

「～てしまう」の縮約形である「～ち (まう、ゃう)」形の用例を (12) に示す。ここでは、語根に含まれる母音の一つ以下の動詞ではサ抜き現象がブロックされることがあるという一般化が成立する。

(12) 「～ち」 (382 例)

押ささち[8]、押さち[1]、起こさち[2]、差ささち[1]、差さち[4]、殺さち[3]、刺さち[314]、指さち[1]、出さち[4]、挿さち[9]、増さち[3]、潰さち[10]、返さち[1]、崩さち[2]、流さち[14]、話さち[3]、騙さち[2]

最後に、終止・連体形「～さる」のデータを (13) に示す。終止・連体形は収集できた例文の件数が最も多かった語形である。

## (13) 「～さる」 (6531 例)

もてなさる[8]、圧ささる[3]、圧さる[2]、移さる[39]、潰さる[6]、逸らさる[4]、引き起こさる[157]、引き伸ばさる[3]、隠さる[15]、映さる[7]、洩らさる[1]、益さる[1]、越さる[11]、延ばさる[1]、汚さる[19]、押ささる[329]、押さる[52]、果さる[1]、果たさる[50]、過ごさる[17]、過さる[1]、解さる[3]、回ささる[14]、回さる[12]、壊ささる[1]、壊さる[40]、廻さる[7]、外さる[57]、確認ささる[1]、覚まさる[2]、活かさる[48]、乾かさる[1]、乾さる[1]、完成ささる[6]、干ささる[29]、干さる[19]、期待ささる[7]、帰さる[7]、記さる[25]、記録ささる[9]、起こさる[168]、起さる[86]、及さる[8]、及ぼささる[2]、及ぼさる[2]、許さる[100]、協力ささる[1]、強調ささる[12]、脅かさる[36]、興さる[3]、驚かささる[6]、驚かさる[163]、空かさる[2]、繰り返ささる[2]、繰り返さる[148]、繰返さる[11]、決定ささる[1]、検討ささる[3]、見逃さる[5]、見做さる[7]、減さる[3]、減らさる[64]、現さる[5]、現わさる[1]、荒さる[5]、降さる[5]、降ろさる[32]、差ささる[7]、差さる[229]、催さる[57]、採用ささる[1]、済さる[1]、済まさる[26]、殺さる[144]、晒さる[72]、散さる[2]、散らさる[5]、残さる[41]、刺ささる[305]、刺さる[492]、指ささる[12]、指さる[134]、施さる[51]、試さる[80]、示さる[50]、質さる[3]、実現ささる[6]、実施ささる[2]、写さる[8]、射ささる[3]、射さる[85]、赦さる[20]、取り消さる[11]、取消さる[4]、充たさる[3]、出ささる[10]、出さる[73]、消ささる[3]、消さる[49]、照さる[3]、照らさる[69]、蒸ささる[10]、蒸さる[188]、伸ばさる[12]、侵さる[24]、心配ささる[2]、浸さる[1]、尽くさる[8]、尽さる[10]、推さる[54]、成功ささる[55]、正さる[13]、醒まさる[1]、説明ささる[10]、遷さる[10]、捜ささる[1]、捜さる[1]、挿ささる[12]、挿さる[302]、増さる[149]、増やさる[8]、促さる[44]、打ち消さる[19]、対応ささる[5]、探さる[20]、注さる[2]、注目ささる[1]、著さる[2]、著わさる[1]、超ささる[21]、超さる[10]、直さる[21]、通さる[3]、潰さる[93]、展開ささる[3]、展示ささる[1]、転がさる[6]、渡さる[52]、登場ささる[10]、倒さる[108]、逃さる[3]、透かさる[3]、透ささる[1] (とおささる)、透さる[1]、動かささる[1]、動かさる[56]、擦ささる[15]、擦さる[1]、濡らさる[7]、燃やさる[14]、剥がさる[30]、剥

さる[1]、曝ささる[2]、曝さる[11]、発さる[8]、発見ささる[1]、判断ささる[3]、反映ささる[3]、犯さる[31]、飛さる[7]、飛ばささる[12]、飛ばさる[133]、表さる[72]、表わさる[16]、評価ささる[1]、沸かさる[10]、返さる[84]、勉強ささる[10]、暮さる[4]、暮らさる[8]、崩さる[29]、放さる[6]、冒さる[6]、満ささる[2]、満さる[9]、満たさる[94]、鳴らさる[12]、滅さる[14]、滅ぼさる[43]、目指さる[1]、戻さる[49]、諭さる[7]、溶かさる[12]、要求ささる[1]、落ささる[1]、落さる[14]、落とささる[12]、落とさる[107]、乱さる[9]、利用ささる[8]、離さる[13]、流さる[39]、冷さる[1]、冷やさる[8]、励まさる[85]、漏らさる[2]、話ささる[3]、話さる[32]、惑わさる[70]、斃さる[2]、騙さる[143]

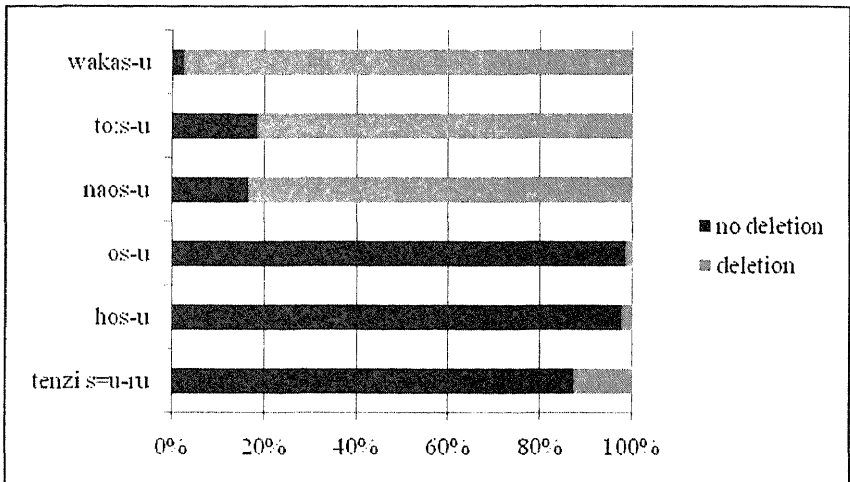
送り仮名違いの検索項目 (たとえば「冷やさる」と「冷さる」) を1項目と数えると、連体・終止形で検索することができた自発述語の形式は、149項目あった。このうち母音を二つ以上含む動詞語根をベースとした自発述語は、109項目あり、そのうち96項目はサ抜きをこうむった形式のみであった。一方、13項目ではサ抜きを被った語形とサ抜きを被らない語形が共存している。なお、母音を二つ以上含む動詞語根をベースとした自発述語では、サ抜きを被らない語形のみ項目は存在しなかった。母音を一つ以下しか含まない動詞語根をベースとした自発述語は、14項目あり、サ抜きを被らない語形のみ項目が2項目あり、残る12項目ではサ抜きを被った語形とサ抜きを被らなかった語形の両方が存在している。

サ抜きのブロックされた語形を含む項目は、二つ以上の母音を含む動詞語根で構成される自発述語の12%であり、母音を一つ以下しか含まない動詞語根で構成される自発述語の86%であった。連体・終止形においても母音を一つ以下しか含まない動詞語根で構成される自発述語の場合のほうがサ抜き現象がブロックされやすいといえる。

アンケート調査とインターネット調査のデータの違いについて考察する。「動詞語根に含まれる母音が二つ以上のときにサ抜き現象が生じやすく、サ抜き現象のブロックは動詞語根に含まれる母音が一つ以下のときに生じやすい」という傾向は、アンケート調査で得たデータでもインターネット調査で得たデータでも見られる。しかしながら、傾向の度合は異なる。

(14) のグラフは2008年5月23日に札幌学院大学で行ったアンケート調査をまとめたものである。この調査では北海道方言話者120名から回答を得た。このグラフが示すように、アンケート調査では動詞語根に含まれる母音の数により、サ抜き現象の生起に関する傾向性にかかなり明確な違いが出る。

(14) アンケート調査で調べたサ抜き現象の分布



これに対して、インターネット調査では、一般化の例外となるデータがかなり出現する。連体・終止形の場合、母音を一つ以下しか含まない動詞語根で構成される自発述語の86%でサ抜き語形と非サ抜き語形の両方が出ることを上で記したが、母音を一つ以下しか含まない動詞語根で構成される自発述語の場合でも非サ抜き語形が圧倒的に多いというわけではない。

サ抜き語形と非サ抜き語形が共存する項目は12あるが、そのうち非サ抜き語形が多いのは、「圧(さ)さる」「押(さ)さる」「捺(さ)さる」「干(さ)さる」「超(さ)さる」の5項目だけであり、残りの7項目ではサ抜き語形のほうが多い。サ抜き語形が多い7項目のうち「刺(さ)さる」「指(さ)さる」「射(さ)さる」「挿(さ)さる」「差(さ)さる」に関しては、サ抜き語形が相対自動詞と同形なので、プログラムが標準語などの相対自動詞の例

を捨ててしまった可能性がある。これらの項目を除外してもサ抜き語形が優位な項目として、「消 (さ) さる」「出 (さ) さる」が残る。

このように母音を一つ生かししか含まない動詞語根で構成される自発述語においてもサ抜き語形が無視できない割合で存在するのは、東北地方の自発述語のデータの混入によるものである可能性もある。例えば、山形市の方言の場合自発接尾辞が /rasar/ ではなく /rar/ であり (森山・渋谷 1988)、通常形態法で生じる語形が北海道方言のサ抜き語形と同じ形式になる。こうしたデータが混ざれば、母音を一つ生かししか含まない動詞語根で構成される自発述語においてもサ抜き語形は含まれることになる。

第3節で述べたように、今回のインターネット調査では書き手の属性を同定していないので、収集してきたデータが、どの方言を反映しているか不明である。実際には、自発接尾辞が /rasar/ の方言だけでなく /rar/ の方言のデータも混ざっているのだろう。このような状態であるにもかかわらず、母音を二つ以上含む語根で構成される自発述語の場合、アンケート調査の場合と同様に高い割合でサ抜き語形が現れる。これは、比較的長い語形の自発述語の場合、北海道方言を含む自発接尾辞が /rasar/ の方言においてサ抜き現象が高い生産性を持っていることを示す傾向と考えられる。

インターネットで収集したデータとアンケート調査で収集したデータには傾向性の度合の差がある。この度合いの差は、書き手の属性を絞り込むことができていないことによるものと思われる。しかしその程度差を除けば、概ねインターネットから収集したデータでもアンケート調査で得たデータと同様の傾向が見られるといえるだろう。

## 5 課題

前節の分析から、インターネット上の方言語形が全く無秩序なものではなく、アンケート調査の結果とほぼ同じ傾向性を持っていることを示すことができたのではないかと思う。これが、方言文法研究資料としてのインターネットの可能性を示すものであるならば、方言文法に質的にも量的にもこれまでとは異なる資料が加わることになるだろう。

インターネット上のデータには、フィールドワークで集めたデータと異なり、いくつかの難点がある。イントネーションなどの文法と密接にかかわる音声情報が欠落していること、そして話者の属性が掴みにくいことである。前者は文字情報である以上どうすることもできないのかもしれない。後者は、方言という地域的なバリエーションを研究していく上ではかなり痛いデメリットである。後者については何とか克服する方法を探す必要がある。インターネット上のデータを産出した話者がどの方言の話者であるかは、使われている語彙などからわかる場合もあるだろう。ブログなどの場合、プロフィール欄を利用することが可能な場合もあるだろう。インターネット上のデータがどの方言のデータであるかを確定する効果的な方法の確立が今後の課題となる。話者の属性を確定し、特定の方言のデータだけを抽出した場合、本稿で示したものと異なる傾向が現れるか否か、今後検証していきたい。

#### 【参考文献】

- Ito, Junko (1990) 'Prosodic Minimality in Japanese'. *CLS* 26: 213-240.
- 佐々木冠 (2007) 「北海道方言における形態的逆使役の成立条件」『他動性の通言語的研究』角田三枝・佐々木冠・塩谷亨編: 259-270. くろしお出版
- 佐々木冠 (2008) 「インターネット上の自発述語形収集調査」『日本語方言の態に関する通言語学的観点からの研究』(科学研究費補助金(基盤研究(c)) 研究成果報告書) 佐々木冠編: 102-110. 札幌学院大学.
- Sasaki, Kan (2008) 'Syllable deletion as a prosodically conditioned derived environment effect'. A paper distributed at the 18th Japanese / Korean Linguistics Conference, The City University of New York.
- 森山卓郎・渋谷勝己 (1988) 「いわゆる自発について: 山形市方言を中心に」『国語学』 152: 47-59.