

## シリトリに関する数量的調査

村田 真樹 井佐原 均

独立行政法人 通信総合研究所

{murata,isahara}@crl.go.jp

### 1 はじめに

誰も子供のころシリトリという遊びを行っていたと思うが、そのときふと必勝法とまで言わなくても戦略のようなものはないかと考えたことはあるだろう。例えば、リ攻めである。その文字から始まる単語を言うのが難しい「り」という文字で終わる単語を連続して言っていき、相手を困らせるのである。しかし本当にリ攻めが有効なのか、もっと有効な文字は本当にないのか、そのような疑問を持ったことはないだろうか。言語処理の研究者の立場から、このシリトリに関する疑問について研究をしてみようと思いたった。本稿はその初歩的な研究である。言語処理で利用される単語の辞書を利用してシリトリについて数量的な調査を行なうのである。ところで、シリトリはこのような興味本位の対象としてのものだけでなく、実際の社会にも役に立つものでもある。例えば、シリトリは日本語の単語を覚えるためのゲームとして利用されることがある。これには幼少期の日本人を教育する場面と外国人が日本語を勉強する場面とがある。また、シリトリは、人間と機械のコミュニケーションをテーマとする研究分野においても研究対象となる場合もある。ここでは、人間がどういうときに楽しさを感じるかについて研究がなされている。このようにシリトリは有用性のあるものでもある。本稿はこのシリトリについて分類語彙表 [1] の単語群を利用して数量的調査を行なったものである。

### 2 シリトリ研究の有用性

本節では、研究背景としてシリトリの学術的有用性について議論する。シリトリの有用性は以下の二つに大別できる。

- 日本語単語の教育
- コミュニケーションのための一場面としての利用

まず、「日本語単語の教育」について説明する。日本語単語の教育の場面としては、幼少期の日本人が日

本語の単語を覚えるためにシリトリをするというものと、外国人が日本語の単語を覚えるためにシリトリをするというものの二つがある。実際にこれらの補助を目指したシリトリゲームのソフトの開発もなされている [2]。

次に「コミュニケーションのための一場面としての利用」について説明する。近年、シリトリは、人間と機械のコミュニケーションをテーマとする研究分野においても研究対象となっている [3]。ここでは、人間がどういうときに楽しさを感じるかについて研究がなされている。その研究ではシリトリの相手がシリトリゲームのソフトであってもそれが人間であると教示されているときに面白さを感じるという結果が得られている。

これらの研究のようにシリトリは、日本語教育、コミュニケーションの研究に役に立つものである。本研究はこのように種々の研究に役にたつシリトリについて数量的調査を行なったものである。本研究で行なった数量的調査の結果は、上述の種々の研究で用いられるシリトリソフトを開発する際の戦法の構築などに役に立つことになる。

### 3 シリトリに関する戦法

本節ではどういう戦略をとるとシリトリで有利となるかについていくつかの仮説を考える。

まず、シリトリの戦略をたてる上で必要となる要素について整理する。シリトリの戦略をたてる上で必要となる要素として以下のもの考える。

- その文字で終わる単語の個数。(この数を *END* と呼ぶ。)
- その文字で始まる単語の個数。(この数を *BEGIN* と呼ぶ。)ただし「ん」で終わる単語を除く。これは、「ん」で終わる単語はシリトリでは反則になるためである。
- その文字から始まり「ん」で終わる単語の個数(この数を *NN* と呼ぶ。)

- その文字から始まりその文字で終わる単語の個数 (この数を *SAME* と呼ぶ.)
- その文字で終わる単語の集合において始まりになりうる文字の種類の数 (この数を *BKIND* と呼ぶ.)
- その文字で終わる2文字の単語の集合において始まりになりうる文字の種類の数 (この数を *BKINDC2* と呼ぶ.)

次にシリトリの戦法を考える。本稿では、ある特定の文字を選択しその文字で終わる単語を連続して言ってい、相手の使える単語をある特定の文字から始まるものに限定させ、その文字から始まる単語を減らしていき相手を負けに追い込む作戦を基本戦法とする。以下ではこの戦法での攻めのことを、その選択した特定の文字での攻めと言う。この戦法のもとでシリトリで有利となるための制約条件を考える。制約条件としては以下のものが考えられる。この制約条件はこの条件を満足しなければ有利にすることができないというものである。

- 制約1 (*END > BIGIN*)  
はじまる単語の数よりも終わる単語の数の方が  
多い文字でなければ攻めが続かない。

次にシリトリで有利となるための選好を考える。選好としては以下のものが考えられる。この選好はその選好の傾向が強いほど有利になるであろうというものである。

- 選好1 ( $\max END/BIGIN (END \gg BIGIN)$ )  
はじまる単語の数に比べて終わる単語の数の方が  
比率的に多い文字である方が攻めが続きやすい。
- 選好2 ( $\min BIGIN$ )  
はじまる単語の数が小さい文字で攻めるほど、相手は回答に困り勝ちやすい。(文献 [2] のシステムでは、この数に従って言う単語を決定している。)
- 選好3 ( $\max NN/(BIGIN + NN)$ )  
その文字から始まる単語群の総数に対して、「ん」で終わる単語の数が大きい文字で攻めるほど、相手は「ん」で終わる単語をいいやすく、相手は反則負けしやすく、勝ちやすい。

- 選好4 ( $\min SAME$ )  
はじまりとおわりが同じ文字である単語が多いときのそのはじまりとおわりの文字で攻める場合、この種の単語により回答され、きりかえしを受けやすく攻めが続きにくくなる。
- 選好5 ( $\max BKIND$ )  
いろいろな文字ではじまる単語で攻めることができる文字ほど攻めが続きやすい。
- 選好6 ( $\max BKINDC2$ )  
特に2文字の単語であると想起がしやすいので、いろいろな文字ではじまる2文字の単語で攻めることができる文字ほど攻めやすい。

#### 4 シリトリに関する数量的調査

本稿ではシリトリのルールとしては以下のものを採用する。

- 使用可能な語は名詞のみ。
- 「ん」で終わる単語を使用すると負け。
- 濁音、半濁音で終わる単語の次には、その濁音、半濁音をとった語で始めてもよい。また、前の単語の最後の文字に濁音、半濁音をつけた語で始まる語を次に言ってもよい。
- 「しゃ」など小さい「や」で終わる単語の次には「や」から始まる単語をいう。「しゃ」から始まる語は使えない。
- 長音はひらがなに戻しその文字とする。例えば「カバー」は「かばあ」となり、「あ」が終りの文字となる。

本稿での調査には分類語彙表で“\*”の印がついている単語を用いた。これは、現代雑誌九十種の語彙調査において標本使用度数が7以上であった語である。この単語群をシリトリに用いる単語として利用したのは、この単語群が、例えば国語辞書中の全単語を用いるよりは、普通に人間が思いつく単語群に似ているであろうと考えたためである。本稿ではこの単語群を利用するが、例えば小学生用の教科書で頻出の語を使うなど、この単語群の選定を改良することは可能であると思われる。分類語彙表の名詞で“\*”の印がついてい

る単語の総数は、4,718語であった。多義性のある語の語数はその多義の個数ではなく1個とした。

この単語群を用いて6つの選好の値を計測した。その結果を表1に示す。表では選好1の値が大きい文字から順に示している。それぞれの選好の上位10個(選好2,4のように値が小さい方がよいものについては小さい10個)を太字で表示している。それぞれの選好でよいものが攻めるのに用いるとよい文字であると判断する場合は、太字の欄が多い文字ほど攻めるのによいということになる。ところで、ここでは制約1を示していないが、これは選好1の値が1よりも大きいことを意味していて、この表の上から「め」の行までの文字が制約1を満足していることになる。

次にこの表を見ながらどの文字で集中的に攻めるとよいかを検討する。まず、制約1を検討する。制約1は、「はじまる単語の数よりも終わる単語の数の方が多い文字でなければ攻めが続かない」ので、はじまる単語の数と終わる単語の数の比の選好1の値が1よりも大きくなければならないというものである。ある文字で集中的に攻めようと考え、この条件は守らなければならない。このことにより表で「む」よりも下の文字は攻める際に利用できないことがわかる。

次に選好について検討する。選好についてはどれが特によいかは明らかではない。このため、総合的に考察する必要がある。

ここではまず考察のきっかけとして一般に言われるり攻めの「り」の行を考察しよう。ここで特に目を引くのは選好3の0.23である。制約1を満足する文字の中では最も大きい値である。選好3は「ん」で終わる単語の文字の割合を意味するが、相手がうっかりしていれば「り」で攻めていると毎回2割程度の確率で「ん」をいう可能性があり、「ん」を言わせて勝つという可能性が高いのである。これは一つの興味深い点である。

次に「り」において選好5,6をながめてみる。これらの値が一番ということはないが比較的大きな値である。これらは、BKINDのような「その文字で終わる単語の集合において始まりになりうる文字の種類の数」、また特に選好6は二文字単語におけるその数を意味するが、これらの値が大きいとさまざまな語から「り」で終わる単語を言うことができるのである。濁点や「ん」などを除く、シリトリの単語の先頭となりうる仮名の種類は44であり、そのうち38個の文字から「り」にできるのである。さらに選好6が20であることから、「蟻(あり)、檻(おり)、狩り(かり)」など

表1: 数量的調査の結果

文字	選好1	選好2	選好3	選好4	選好5	選好6
る	<b>21.00</b>	2	0.00	0	23	8
う	<b>12.40</b>	60	0.10	5	<b>41</b>	11
く	<b>5.40</b>	67	0.13	3	<b>39</b>	<b>23</b>
い	<b>4.23</b>	122	0.13	19	<b>43</b>	<b>20</b>
つ	<b>3.41</b>	56	0.07	0	<b>37</b>	<b>17</b>
ら	<b>3.00</b>	14	0.18	1	21	10
り	<b>2.71</b>	49	<b>0.23</b>	3	<b>38</b>	<b>20</b>
の	<b>1.80</b>	20	0.13	2	21	3
ろ	<b>1.56</b>	25	<b>0.22</b>	2	23	6
れ	<b>1.52</b>	29	0.09	0	22	<b>14</b>
み	1.48	50	0.15	7	26	11
き	1.30	189	0.15	19	<b>41</b>	<b>27</b>
ち	1.09	87	<b>0.19</b>	4	<b>35</b>	<b>20</b>
め	1.06	<b>31</b>	0.09	4	23	11
む	1.00	<b>28</b>	0.07	1	17	4
す	1.00	63	<b>0.19</b>	4	29	12
え	0.97	40	0.18	3	22	9
と	0.90	148	0.11	4	<b>38</b>	<b>16</b>
わ	0.85	39	0.05	3	20	7
や	0.77	<b>31</b>	0.16	3	16	7
ま	0.65	60	0.06	2	22	12
ね	0.58	31	0.16	2	13	11
ふ	0.58	123	0.16	8	24	8
て	0.56	85	0.14	3	25	11
け	0.51	125	<b>0.20</b>	2	32	13
し	0.47	479	0.18	37	<b>40</b>	<b>26</b>
も	0.43	<b>30</b>	0.17	1	8	5
ひ	0.42	110	0.10	9	22	9
こ	0.38	199	0.17	13	27	12
な	0.38	55	0.11	2	16	10
か	0.35	288	0.15	14	<b>34</b>	<b>14</b>
た	0.34	157	0.11	4	26	11
さ	0.33	96	0.18	2	20	7
は	0.28	152	<b>0.20</b>	8	23	6
に	0.27	44	<b>0.19</b>	2	10	7
ぬ	0.25	8	0.00	0	2	2
ゆ	0.22	37	0.10	1	8	3
そ	0.17	89	0.13	1	12	7
ほ	0.17	90	<b>0.22</b>	1	12	3
へ	0.15	33	<b>0.28</b>	0	5	2
せ	0.13	148	<b>0.22</b>	2	12	5
お	0.12	98	0.05	1	10	4
よ	0.09	74	0.14	3	4	1
あ	0.07	106	0.05	2	6	2

(表の上から「め」の行までの文字が制約1を満足している)

のような2文字からなるシリトリにおいて人間にとって非常に思いつきやすいものが20個あるのである。このことから「り」での攻めは続きやすく有利であることがわかる。

次に「り」において選好4をながめてみる。選好4は始まりと終わりが同じ文字になる単語の数であるが、「り」は3で特に少ないわけではないがそれほど多くなく、切り返しをうける可能性は大きくないことがわかる。

次に「り」において選好1をながめてみる。選好1ははじまる単語の数と終わる単語の数の比であり、選好1の値が大きいほど、その文字での連続した攻めが続きやすい。この値は2.71であり、1より十分大きく「り」でも攻めが続くであろうことが予想される。

最後に「り」において選好2をながめてみる。選好2はその文字から始まる単語の数であるがそれは49であり、極端に小さいわけではなく、選好2だけで考えると「り」がいいとは言えない。選好2を考えると、むしろ、「る」や「ら」の方が有望である。他の選好、例えば、選好5,6も組み合わせることで「り」が有効になると判断できるのであろうと思われる。

以上より、もし「り」が有効であるならば、その理由は主に以下のものであると予想できる。

- 選好3が大きく相手が「ん」で終わる単語を言いやすい。
- 選好5,6が大きく特に選好6が大きく容易に「り」で終わる単語を連想しやすい。
- 選好1が2.71で大きく、「り」で始まる単語数より「り」で終わる単語数の方が圧倒的に多く攻めが続きやすい。

以上では「り」を中心に考察してきたが、ここで他の語でも有利にならないかを検討してみよう。表で「り」よりも上部にあって太字が多い文字に「く」「い」「つ」がある。これらは選好5,6が大きいので比較的攻めが続きやすくこれらも有効であろうと思われる。しかし、選好3がそれほど大きくなく相手が「ん」を言いやすいということはなく、また、選好2の値は「り」の値よりも大きく相手が言える語がネタギレになるのも「り」より遅いという欠点がある。これを考えると、「く」「い」「つ」も有力だろうけれども「り」ほどではないだろうと予想される。その他では「う」も有力である。「う」は選好5が41ありまた選好1が12.40もあるので、かなり攻めが続きやす

いと思われる。しかし、「う」で終わる単語が多いとはいえ、「愛情」のように難しい熟語で後ろの漢字の長音の「う」になる場合が多く、2文字単語の選好6が小さく、単語を思い浮かべるのは少々苦勞する。それでも、新たな作戦として使えるかもしれない。その他で表の上位で太字が多くあるのは、「り」以外のら行の「ら」「る」「れ」「ろ」である。これらは、選好2が小さく相手が答えるのが難しいのでそれなりに使えろとは思われる。

現状の調査では、数量的な調査を行なっただけであり、これですべてがわかる、また、これが必ず正しいということではないが、それなりに「り」が有効であろうと予測できる結果、また、他の有効な文字の洗い出しに役に立つものであると思われる。

## 5 おわりに

本稿ではシリトリで有力となる戦法の調査のために、シリトリについて分類語彙表[1]の単語群を利用して数量的調査を行なった。この結果、り攻めが有効である理由となりえるいくつかの説明をつけた。また、「り」以外にも有力な文字の洗い出しを行なった。

今後の課題としては以下の二つがある。一つは被験者を使った調査である。本稿の調査では単語群を用いた数量的調査であり、必ずこうだという結果を得ることは難しい。また、本当にり攻めが有効かどうかかわからない。被験者を使っていろいろな文字で攻めを試してみてもどの文字が有効かを確かめる実験を行なうと、よりはっきりした結果が得られることになるとと思われる。もう一つは単語辞書の変更である。本稿の調査では分類語彙表を利用したが、これが実際に人が思い浮かべやすい語群と一致するとは限らない。また、戦法の対象とする人を小学生などに限定すれば、小学生が頻度多く使う語に限って調査するなどのことが必要である。このどのような単語群を利用するとよいかについても今後の課題となっている。

## 参考文献

- [1] 国立国語研究所, 分類語彙表, (秀英出版, 1964).
- [2] Tiemi Christtine Sakata and Masahide Sugiyama, Internet shiritori using java, 情報処理学会第59回全国大会予稿集, 5N-10, (1999).
- [3] 山本吉伸, 開一夫, 松井孝雄, 梅田聡, 安西祐一郎, コミュニケーションの面白さ-シリトリはなぜ面白いのか-, 情報処理学会ヒューマンインターフェース研究会 47-13, (1993).