

Markednessを使った省略の一考察

成山重子

奈良先端科学技術大学院大学 / メルボルン大学
shigecko@is.ait-nara.ac.jp / shigecko@unimelb.edu.au

1. はじめに

ゼロ代名詞に代表される言語の省略現象を把握し先行詞を同定することは、言語処理の研究においても言語学の研究においても重要な課題である。しかし、周知の通り、省略は統語論のみならず意味論そして語用論が深く関与しており、それらを包括的にとらえることが必要な極めて複雑な言語現象であるといえる。そのため、Centering Theory [5,8,13,14]のように語用論まで考慮した論説は、日本語も含め、広く言語処理の研究において応用されてきた。しかし、複文の扱いは不十分で、特に日本語では単文のみに摘要されることや[8,11]、一文に1つの照応詞しか同定できないなどの未解決問題も多くある[11]。さらに機械翻訳などの言語間の言語処理においては、同定した先行詞をゼロ代名詞に訳すか、他の代名詞にすべきか、所有格や自分などの再帰代名詞はどう扱うかなどの諸々の問題もでてくる[15]。

本稿は、従来とは違った観点から省略現象を観察し、新たな見解から省略現象の理解を深めることを目的とする。そのため、ゼロ代名詞の先行詞の同定という処理の観点から離れて、話者/筆者/生成の立場から、どんな時省略をするのかを、言語学で諸々の言語現象の説明に使われているMarkednessの概念を用いて考察する。機械翻訳も考慮にいれ、英語と日本語の両方のゼロ代名詞の事例をとりあげながら所見を述べることにする。

2. Markedness

Markednessは、言語現象にある諸々な不均衡な関係を説明するのに大変有効な概念であり、言語学では19世紀から使われている[1,4]。また、様々な言語にみられる言語現象で普遍性もあると指摘されている[3,6]。

Markednessは、Unmarked(標準/基本/default)とMarked(非標準)という二極を総称する概念である。その有効性から言語学では広範に使用されているため、定義や使い方は多様性極まりないが、大きく次の二つに分類する事ができる。①両極性に分割する現象をとらえるもので、Unmarked(文字通り、無標記)とMarked(有標記/明記)からなる。②傾向性があるもので、Unmarked(例えば、自然/普通な文)、対、Marked(より非自然/特別/非標準な文)の関係をもつ。以下にそれぞれについて詳しく述べる。

2.1. 二種の Markedness

一つは、両極性をもつ言語現象をとらえるのに有効で、Unmarked(文字通り、無標記)とMarked(有標記/明記)に分かれる。例えば、無声音('s)[-voiced]を基本として、それに対し有声音('z)は[+voiced]と扱ったり、単数をUnmarkedとし複数(単数+'s)をMarkedとすることなどである。下記の表1にさらに例をあげる。これらの例より、Unmarkedは文字通り無標記であるのに対して、MarkedはUnmarkedになんらかの形態素を付加していることがわかる。日本語においても、「肯定(Unmarked)+ない=否定形」などの例からこの現象があることがわかる。

表 1 : UnmarkedとMarkedの英語例

FEATURES	UNMARKED	MARKED
VOICING	voiceless	voiced
NUMBER	singular e.g. <i>dog</i> + \emptyset <i>the</i> \emptyset <i>girl</i> + \emptyset	plural (dual) e.g. <i>dog</i> +s <i>the</i> <u>two</u> <i>girls</i>
GENDER	masculine \emptyset + <i>man</i> \emptyset + <i>male</i> <i>prince</i> + \emptyset	feminine e.g. <u>wo</u> + <i>man</i> <u>fe</u> + <i>male</i> <i>prince</i> + <u>ss</u>
NEGATION	positive e.g. \emptyset + <i>happy</i>	negative e.g. <u>un</u> + <i>happy</i>
TENSE	present e.g. <i>walk</i> + \emptyset	past e.g. <i>walk</i> + <u>ed</u>

もう一つは、傾向性をもつもので、例えば言語学では、例文が非文かどうかの判断をするとき、自然（普通／よくある）な文をUnmarkedとし、それに対し、非標準的な（より不自然な／特別な／含みがある／少々曖昧な／強調や比較を表現する、などの）文をMarkedとする扱いである。他の具体例を以下にいくつか示す。

- 語彙選択: 例えば、どんなに短くても英語では“*How long ...?*” (Unmarked) と質問し、“*How short ...?*”とは普通言わない。
同様に、たとえ一歳であっても、“*How old ...?*” とときき、“*How young ...?*” という聞き方は英語にない。これは、慣用化された副詞に多くみられる現象で、もちろん全ての副詞が対をなすわけではない。例えば、“*How good is it?*” も “*How bad is it?*” もどちらも使われる。
- 構文: 生成文法では、能動文をUnmarked基本形とし、受身文をMarked（生成して導き出すもの）としている。

2.2. もう一種の Markedness

上記の二種のMarkednessには相関関係がみられる。つまり、Unmarkedとは、②自然な（頻度の高い）言語表現は①無標記されるという関係である。ここから、省略現象の把握に有効な、三つ目の分類を導き出す事ができる。③(①+②)「省略は同一性を表す。」換言すれば、「同一であれば明記しない。同一でなかったら、明記する。」よく出てくる主語は繰り返さずよりも、省略した方が簡素／経済的で文の結束性も高まる。実際、統計的にも同一主語の複文の割合は、日本語で64%、英語でもそれにとても近い66%と報告されており、同一主語文を生成する偏りがあることがわかる[10]。

以下、どんな時省略が現れるか、著者／話者は省略するのか生成の立場から、日本語と英語の事例を考察する。

3. 主語の同一性

統語論[2, 他]では、英語の複文で省略された従属節の主語は、主節の主語と同一であるという(1a)。これは、Markednessの③(①+②)「省略は同一性を表す」の主旨と一致する。そして、異なっていれば明記する(1b)。所有格も同様に、主語と同一の場合は省略され(2a)、違う場合は明記される(2b)。

(1a) *I want* \emptyset *to go.* [unmarked = SS (same subject, 同一主語)]

(1b) *I want* him *to go.* [marked = DS (different subject, 異主語)]

(2a) *I will do it, upon \emptyset return.* [unmarked = SS]

(2b) *I will do it, upon his return.* [marked = DS]

日本語でも、英語のように普通、省略された従属節の主語は、主節の主語と同一になり(3a)、異なる場合は明記される(3b)。ただし、日本語では助詞に表れるトピック性もこれに関与してくるため、(3c)のように従属節の主語が省略されていても、主節の主語が「が」をともなった時には、文脈により、SS,DSのどちらの解釈も成り立つ[10]。

(3a) \emptyset 家に帰ると、太郎はすぐ寝てしまった。 [unmarked = SS]

(3b) 次郎が家に帰ると、太郎はすぐ寝てしまった。 [marked = DS]

(3c) \emptyset 家に帰ると、太郎がすぐ寝てしまった。 [marked = SS/DS]

また、所有格では英語と同様に、主語と同一の場合は省略され、違う場合は明記される。

(4a) 太郎は \emptyset 車を売った。 [unmarked = SS]

(4b) 太郎は次郎の車を売った。 [marked = DS]

4. 主語の非同一性

しかし、省略が主節主語と同一なのに明記されたり、また逆に異なっているのに省略されるときもある。まず前者の場合だが、例えば、再帰代名詞「自分」がある場合は、同一人物なのに明記されることになる。

(4c) 太郎は自分の車を売った。 [marked meaning = SS]

これは話者が強調を表したり、「自分」を第三者的に描写している[7]ときに現れる。これをMarkednessの②と③により説明すると、Unmarked③(同一性)に他の意味を付加して基本の文を逸脱したこと(Marked②摘要)から、明記されたわけである。

対話においても同様で、「明日行きますか。」という質問に対する答として下記のような文が可能だ。(5a)はUnmarked③の例であるが、同一であっても、強調したり(5b)、比較を表すとき(5c)は、Marked②のため明記される。

(5a) はい、 \emptyset 行きます。 [unmarked = SS]

(5b) はい、私が行きます。 [marked = SS 強調]

(5c) はい、私は行きます。 [marked = SS 比較]

次に、主節の主語と異なっているのにもかかわらず、下記のように省略されることもある。

(6) \emptyset_i *Looking back at it now, John_j was wrong.* [DS: $\emptyset=I$]

(7) \emptyset_i *Having bought her_j a bunch of flowers, she_j is happy.* [DS: $\emptyset=I$]

(8) \emptyset_i 小さな体で体力的にもたないだろうが、 \emptyset_j 手術をすることにした。

このような文は、統語論では説明できないが、②Unmarkedを使うと、「その解釈が普通/自然だ/理に当るから」ということになる。(6)(7)は省略された従属節の主語は一人称であるが、一人称はその文を生成する張本人であることから認知的に特別扱いされ、このような「例外」的文も発生する[9]。例えば、(6)では「後になって考えてみると、」という認知的行為は、話者本人にしかできない

ことから、それが自然な解釈になる。また、(8)では「手術をするのは医者」という世界知識から省略される。

5. おわりに

本稿では、生成の観点からMarkednessという新たな概念を用い複雑な省略現象を把握すべく一考察を行った。まず、従来言語学で使われているMarkednessの概念の二種、すなわち、①両極性による、Unmarked(無標記)とMarked(有標記)と分割する現象と、②傾向性による、Unmarked(自然/普通)とMarked(より不自然/特別)の関係をみた。これに加え、三つ目の③(①+②)「省略は同一性を表す」を新たに提唱した。これは、至極当たり前のことで、「同一であれば、あえていわない。違っていたらいう。」というのである。

これにより、基本的に同一のものはゼロ代名詞で表し(もちろん、英語の場合は他の代名詞もかわる)、異なる場合は明示する。同一であるのに明記するときは、何らかの違い(普通と違う、他に含みがある、など)があることを示唆しているのである。今後、このMarkedになる事例を大量に分析し具体化することでより深い理解が見出されるであろう。また、主節の主語と異なっているのにもかかわらず省略される例などは、(6)のような例はある程度は解決できるが、世界知識を必要とするところは、周知の通り、言語処理では非常に厄介であり、Markednessの概念をもって今のところ得策はない。

参考文献

- [1] Battistella, E. (1996). *The logic of markedness*. Oxford: Oxford University Press
- [2] Chomsky, N. (1995). *The minimalist program*. Mass.: MIT Press
- [3] Croft, W. (1990). *Typology and universals*. Cambridge: Cambridge University Press
- [4] Eckman, F.R.- Moravcsik, E. and Wirth, J.R. (eds). (1986). *Markedness*. New York: Plenum Press
- [5] Grosz, B. A.K. Joshi, & S. Weistein, (1983), Providing a unified account of definite noun phrases in discourse, In Proceedings of the 21st Annual Meeting of the American Association for Computational Linguistics, Cambridge, MA: ACL. pp. 44-50
- [6] Greenberg, J. (1966). *Language universals, with special reference to feature hierarchies*. The Hague: Mouton
- [7] Iwasaki, S. (1993). *Subjectivity in grammar and discourse: theoretical considerations and a case study of Japanese spoken discourse*. Amsterdam: John Benjamins
- [8] Kameyama, M. (1998). Intrasentential Centering: a case study. In M. Walker, K. Joshi and E. Prince (eds). *Centering theory in discourse*. Oxford: Clarendon Press. 89-112
- [9] Nariyama, S. (forthcoming). Subject ellipsis in English. *Journal of Pragmatics*
- [10] Nariyama, S. (2002). The WA/GA distinction and switch-reference for ellipted subject identification in Japanese complex sentences. *Studies in Language* 26:2, 369-431, John Benjamins
- [11] Nariyama, S. (2001). "Multiple argument ellipses resolution in Japanese", In Proceedings of Machine Translation Summit VIII, Spain, 241-245
- [12] Strube, M. (1998). Never look back: An alternative to Centering. In proceedings of the 36th Annual Meeting of the ACL: 1251-1257
- [13] Walker, et al. (1994). Japanese discourse and the process of Centering. *Computational Linguistics*. 20(2):193-231
- [14] Walker, M., K. Joshi and E. Prince (eds.). (1998). *Centering theory in discourse*. Oxford: Clarendon Press.
- [15] 吉見毅彦. (2001). 英日機械翻訳における代名詞翻訳の改良, 自然言語処理, vol.8, No.3, pp.87-106