

日英機械翻訳における副詞句翻訳の問題点について

小倉 健太郎 Francis BOND 池原 悟
NTT コミュニケーション科学研究所

1 はじめに

我々は言葉を使用する時、副詞(句)という名のもとに、種々の文法的機能や意味を持たせ、文の豊かな意味内容を表現している。副詞(句)は、その文法的機能の多様性と複雑さのため処理が難しく、また、それが文の意味の骨格をなすものではないということから、自然言語処理において、これまであまり研究が進んでいない。

本稿では、日英機械翻訳における副詞句翻訳に関する問題点について述べる。副詞句翻訳の問題点としては、大きく分けて、(1)副詞句の文法的機能の多様性と意味の多義性(解析での問題点)、(2)副詞的な意味の日本語と英語の表現の違い(変換での問題点)、(3)英語副詞句の語順処理(生成での問題点)、(4)副詞的意味の内部表現(知識表現での問題点)などを挙げる事ができる。ここでは、これらの問題点について記述するが、特に副詞的な意味の日本語と英語の表現の違いに関して英語副詞の観点から見た日英対照データの副詞句の調査結果を中心に、問題点を考察する。

2 日英機械翻訳の副詞句翻訳の問題点

2.1 副詞句の文法的機能の多様性と意味の多義性(解析での問題点)

副詞句の文法機能(付接詞、離接詞、合接詞)や意味(様態、場所、時点など)は非常に多様であり、統一的な取り扱いがなかなか難しいが、言語学者の努力により、英語の副詞に関してはその機能や意味に関してかなり詳細に分類されている[13][12][4]。また、日本語の副詞の意味に関しても、参考のできるものが現れている[2][5]。しかし、副詞に関してはまだまだ包括的な研究は少なく、個別の副詞(even, still, already)に関する研究[1][14]や意味を限定した範囲(時を表す副詞句の時間的な意味など)での研究[15]が多い。

また、英語に関しては辞書(英和辞書、英英辞書)での副詞の取り扱いが不十分である。特に、-ly型の副詞に関しては、記述が少なく、その副詞の文法的機能や用法を調べるのに重要な用例が示されていない場合が多い。よく使用される英語副詞に関して、意味や文法的機能や用法などを詳細に記述したものとしては、小西らによる「英語基本形容詞・副詞辞典」がある。しかし、日英機械翻訳を考えた場合、そのカバー範囲はまだ十分なものとはいえない。

意味の多義性の面を見ると、他の品詞と同様に、副詞も多義性を持つ。例えば、「よく」という副詞は、国語辞典によれば、頻度を表す「しばしば」という意味と、「それをする能力が十分にある」という意味と、「その行為をした労を認める」という意味に使われる。翻訳を考えると、頻度を表す「しばしば」という意味の場合は“often”と訳される場合が多く、2番目の意味では、“well”と訳されることが多く、3番目の意味では、英語では副詞ではなく、“very good”などのように形容詞で訳したり動詞を使って翻訳されることが多い。

語義の絞り込みの条件を考えると、動詞の場合は、その動詞の格要素(主語や目的語)になる名詞との関係で、かなり動詞の意味を限定できるし、名詞の場合も、それを支配する述語(動詞や形容詞)との関係や、修飾部との意味関係などで意味を限定できる。それに対して、副詞は現在のところ、副詞の意味を限定するための条件がきちんと整理されておらず、語義の絞り込みはなかなか難しい。

2.2 副詞的な意味の日本語と英語の表現の違い(変換での問題点)

日本語では副詞的な意味合いは主に副詞で表現されるが、「今日、明日」などの時詞と呼ばれる名詞や、「中心に」などのように「名詞+に」で副詞的な働きをする場合や、「早く」のように形容詞の連用形によ

るもの、「きれいに、立派に」などの形容動詞や形容動詞型名詞に「に」がついて副詞的な役割をするものなどいろいろに表現される。それらは英語では主に副詞として表現されるものである。翻訳では、言語の表現習慣の違いなどから、日本語では副詞的な表現ではないものが、英語では副詞として表現される場合も多い。英語で副詞として表現されるものが日本語表現でどのようなものになっているかは3節に詳しく述べる。逆に日本語で副詞として表現されるものも英語では副詞の他にいろいろに表現される。人間の翻訳では日本語の副詞の半分程度しか英語では副詞として表現されない[11]。

日本語の副詞的な表現を英語で副詞に翻訳する場合でも、副詞的な表現自体だけでなく、その副詞的な表現を修飾する要素なども考慮しなければならない。例えば、「この順に」という日本語は、単純に「順に」を副詞的表現として捉え“in order”と訳し、「この」を「順に」の修飾部とすると、“this in order”のように訳される危険がある。また、「この」を「順」という名詞を修飾する連体詞として捉えたと、全体として副詞句として捉えることがなかなか難しく“on this order”のように生成すべき前置詞を間違えるような現象が現れる。

2.3 英語副詞句の語順処理 (生成での問題点)

英語の副詞句は他のものと異なり比較的いろいろな位置に生成できると言われているが、副詞の文法的機能や意味が決まると生成できる位置もかなり限定できる。また、適切な位置に生成しないと、違う意味に解釈される危険がある。我々は、我々が開発している日英機械翻訳システム[6][8]で英語副詞を適切な位置に生成できるように、副詞の文法的機能と意味と単語による優先的な出現位置(文頭、中位、文末、前置、後置)の傾向から、英語の副詞を43種に分類することにより、英語の訳語選択が正しく行なわれている条件で、97%副詞が適切な位置に生成できることが分かった[7]。当然のことであるが、文中での副詞の位置を決定する場合、他の文要素との関係(文頭副詞の場合は接続詞などのと位置関係、文末副詞の場合他の格要素との語順関係、格要素がthat節やto不定詞

句となっている場合の語順など)や、同じ優先的な出現位置を持つ副詞同士での語順関係も考慮する必要がある。

また、副詞を正しい位置に生成するためには副詞のスコープを正しく認定する必要があるがこれも現在の解析方法ではなかなか難しい。さらに、現在は実現するのはなかなか難しいが、話題化による文頭への倒置なども考慮する必要がある。

2.4 副詞的意味の内部表現 (知識表現での問題点)

副詞的の意味を計算機内部でどのように表現すべきかも問題となる。我々は文法的な機能の違いも文の表す意味の大きな要素であると考え、変換後の英語構造では副詞的な表現を、叙述内容の心的な態度を表す様相構造、格要素構造、修飾構造、接続構造など意味の違いを構造上の違いとして明確に表現する方法をとった[9]。また、文の部分構造に適切に副詞の意味構造を持たせることにより、副詞のスコープを明確に表現できるようにした。

副詞的なものを辞書でどのように表現するかというのも問題となる。「とても」のように固定した表現は、副詞として辞書に登録すればよいし、「大幅に」のようなものは、形容動詞型の名詞として「大幅」で登録すればよいが、「非常に」の場合は「非常」という名詞が存在するし、「順に」の場合は、「順」という言葉にはまだ名詞的な意味合いが残っているので、どのように表現すべきかはなかなか難しい。

3 副詞的意味の日本語と英語の表現の違い - 英語副詞の観点から -

日英翻訳で英語が副詞として表現される日本語表現は非常に多彩である。我々は、新聞記事(経済産業分野)1,000文とその英訳(プロの翻訳者による翻訳)を、分析し、どのような日本語表現が英語の副詞¹として翻訳されるかを調査した。

¹“at least”や“in advance”などのイディオム副詞も含む

3.1 英語副詞の使われ方の傾向

新聞記事1,000文の英訳に現れる総副詞数は585語、異なり語数は145語であった。同じ副詞が多用される傾向が見られた。高頻度の副詞を頻度順にあげると、also:79, about:28, more:26, now:24, only:22, besides:21, out:19, up:17, already:17, actively:13, greatly:12, before:9, especially:9, in the future:9, both:8, directly:8, around:7, mainly:7, approximately:6, easily:6 のようになる。

3.2 英語で副詞として表現される日本語表現

上記のデータに対して、どのような日本語表現が英語の副詞に訳されているかを調査した(表1)。

日本語の副詞句(時詞を含む)が英語の副詞になる場合が最も多いが、それでも17%程度であり、いろいろな日本語表現が英語では副詞として訳されることがわかる。また、助詞も副詞として訳されることが多いが、最も多いのは係助詞の「も」がalsoと訳される場合で、42回(7%)もあった。次に多いのは接続表現であった。接続表現には、接続詞のほか、「... (する) ほか... (する)」といった接続的な表現や、「読点」なども含む、純粋に接続詞が副詞となるものは19回(3%)とそれほど多くない。名詞が副詞に訳される場合は、「に」がついて副詞的な用法となるものと、それ例外がある。前者は21回で後者は38回であり、名詞の副詞的な表現はそれほど多くはない。後者の場合は、「大半」、「程度」など量的な名詞表現が副詞となるものや、日本語と英語の言語表現習慣の違いや発想の違いから文構造に変化が生じている場合など、そのほかいろいろの場合があり、かなり複雑である。形容動詞(形容動詞型名詞を含む)が、副詞に訳される場合も、大きく分けて2種類ある。一つは語幹(または形容動詞型名詞)に「に」がついたもので、形容動詞の副詞的な用法である。もうひとつは、日本語と英語の言語表現習慣の違いや発想の違いから文構造に変化が生じている場合である。前者が26回、後者が12回で、人間が翻訳する場合文構造自体が大きく変化する場合が比較的多いことがわかる。

英語が動詞+副詞になる場合というのは、他のもの

表1: 英語の副詞に対応する日本語表現

日本語表現	日英対訳例	頻度	比率
副詞	将来(in the future), さらに(even)	99	16.9
助詞、助詞 相当語	だけ(only), も(also)	62	10.6
接続表現	また(also), “、”(as well as)	62	10.6
名詞	中心に(especially), 程度(about)	59	10.1
形容動詞	大幅に(greatly), 特殊な(specially)	50	8.5
接辞	約(about), 新(newly)	38	6.5
動詞	引き続き(continuously)	26	4.4
複合語	活発化(actively), 本格(actively)	25	4.3
様相・ 時制表現	そうだ(probably), てきた(before)	14	2.4
形容詞	素早く(quickly), 近い(almost)	9	1.5
連体詞	大きな(greatly)	1	0.2
意識		20	3.4
該当日本語 なし		66	10.3
英語が動詞 + 副詞	設置する(set up), 行い(carry out)	54	9.2
計		585	100.0

とは少し基準が違うものであるが、量としては54回(9%)とかなり多い、元となる日本語は、動詞かもしくは述語性を持つ名詞である。

4 考察

副詞の文法的機能、意味、用法などを決定し、辞書情報として組み込む場合、その用例を含む文を調査することが重要である。包括的に副詞の用例を収集する基本的な方法は、大量の言語コーパスから用例を抽出し、分類整理することである。しかし、用例を抽出することはなかなか難しい作業であるし、翻訳ということ考えた場合、言語間の表現の対応がとれる対

訳データであることが望ましいが、大量で適切な対応データを入手することも現在はなかなか難しい。

副詞のようなものの用例を効率良く入手する方法としては、言語現象のエッセンスを人間が長い時間をかけて集めたものである辞書や文法書の用例を利用する方法が考えられる。和英辞典には、英和辞典では載っていないような副詞の用例も載っており、有効である。我々は、現在この方法で英語の副詞に関して、およそ語彙数1500、用例文数5,000の言語データを収集し、機械翻訳用辞書の英語の副詞(15,000 エントリ)の機能種別をつけるのに利用している。

副詞の意味の日本語と英語の表現の違いを表す言語現象(言語対応データ)を検討するにあたって、日本語の副詞的な表現が英語でも副詞的な表現になる場合と、言語の表現習慣や発想の違いによって、日本語と英語で文構造が大きく変わる場合を、分けて考える必要がある。副詞的な日本語表現が英語で副詞となる場合は、かなり規則性がはっきりしているので、取り扱いやすいが、日本語と英語で文構造が変わる場合は、どういう条件の時に構造変換をしなければならないかが、分かっていない場合が多く、また、変換の条件がわかった場合でも、周辺の修飾要素などをどのように変換すべきかなど、変換処理はかなり複雑なものとなる。

5 おわりに

日英機械翻訳におけるいろいろな副詞句翻訳の問題点を述べたが、これらの問題点を解決するためには、実際の言語現象を詳細に検討し、副詞の訳語選択条件や構造変換の条件などを決める必要がある。このためには、大量の日英対応言語データ[3]や、これを検索するシステム[10]が必要である。また、辞書や文法書の用例の有効活用が重要であると考え。最後に、英語の副詞に対応する日本語表現のデータ作成および整理をしていただいた小見佳恵さんをはじめとするNTTアドバンステクノロジーの皆様にご感謝致します。

参考文献

[1] Paul Berckmans. The quantifier theory of even. *Lin-*

guistics and Philosophy, Vol. 16, pp. 589-611, 1993.

- [2] 茅野直子, 秋元美晴, 真田一司. 副詞. 外国人のための日本語例文・問題シリーズ1. 荒竹出版, 1987.
- [3] 江原暉将, 小倉健太郎, 篠崎直子, 森元逞, 榎松明. 電話またはキーボードを介した対話に基づくデータベースADDの構築. 情報処理学会論文誌, Vol. 33, No. 4, pp. 448-456, 1992.
- [4] S. グリーンボーム. 英語副詞の用法. 研究社出版, 1983. 翻訳本.
- [5] 日向茂男, 日比谷潤子. 擬音語・擬態語. 外国人のための日本語例文・問題シリーズ1 4. 荒竹出版, 1989.
- [6] S. Ikehara, S. Shirai, A. Yokoo, and H. Nakaiwa. Toward an MT system without pre-editing — Effects of new methods in ALT-J/E —. In *Proceedings of MT Summit-IV*, pp. 101-106, 1991.
- [7] K. Ogura, F. Bond, and S. Ikehara. English adverb generation in Japanese to English machine translation. In *4th Conference on Applied Natural Language Processing*, Stuttgart, Germany, 1994.
- [8] K. Ogura, A. Yokoo, S. Shirai, and S. Ikehara. Japanese to English machine translation and dictionaries. In *44th Congress of the International Astronautical Federation*, Graz, Austria, 1993.
- [9] 小倉健太郎. 英語意味構造表現における副詞(句)の扱い. 情報処理学会第43回全国大会, 1992. 2H-2.
- [10] 小倉健太郎, 橋本一男, 森元逞. 言語データベース統合管理システム. 情報処理学会自然言語処理研究会, Vol. 88, No. 69, 1988. NL-69-4.
- [11] 小倉健太郎, 白井論, 池原悟. 日英機械翻訳の副詞翻訳. 電子情報通信学会言語理解とコミュニケーション研究会, Vol. 95, 1995. NLC-95-?? (forth coming).
- [12] 岡田伸夫. 副詞と挿入文. 新英文法選書第9巻. 大修館書店, 1985.
- [13] R. Quirk, S. Greenbaum, G. Leech, and J. Svartvik. *A Comprehensive Grammar of the English Language*. Longman, 1985.
- [14] Johan van der Auwera. 'already' and 'still': Beyond duality. *Linguistics and Philosophy*, Vol. 16, pp. 613-653, 1993.
- [15] Frank Vlach. Temporal adverbials, tenses and the perfect. *Linguistics and Philosophy*, Vol. 16, pp. 231-283, 1993.