

接続詞は一定の順序をもって現れるか？

東京工業大学 大学院理工学研究科 数学専攻

木村 一輝

はじめに

ニュースや小説など、一定の長さをもつ文章は、接続詞を用いて展開されている。接続詞は、メッセージを結びつける情報源であると考えられていて [1]，文章展開の一端を担っていると考えられる。

それゆえ、文章展開の傾向が、接続詞の一定の順序として現れるのではないかと予測できる。次のことを問う：順序をもって現れる接続詞はあるか？

方法

英語のテキストを集めた約一億語を含むコーパスである Brown コーパスを対象とする。

コーパスから、出現頻度の高い従属接続詞 4 つを候補として選び、その接続詞だけを取り出したテキストをつくる。そのテキストから、すべての接続詞組 (a, b) の順序指数を計算する。

順序指数は次のように定義する。接続詞組 (a, b) に対し、

$f(a, b) = (a, b)$ が順に出現する頻度

$$\text{順序指数} = \frac{|f(a, b) - f(b, a)|}{f(a, b) + f(b, a)} \times 100$$

順序指数の定義で、組の総頻度 $f(a, b) + f(b, a)$ で割っているのは、異なる接続詞間で比較をするためである。

順序指数の低い接続詞対は、順序を入れ替えても、出現頻度が余り変わらない。順序指数の高い接続詞対は、順序を入れ替えると、出現頻度が大きく変わる。したがって、順序指数の高い接続詞対は、一定の順序をもって現れやすいとかがえることができる。

結果

出現頻度の高い従属接続詞 4 つは、順に、”as”，”when”，”if”，”so” であった。

以下に、接続詞の順序対の出現頻度の表、接続詞組に対する順序指数の図を載せる。

表 1 接続詞の順序対の出現頻度

順序対	頻度	順序対	頻度
("as", "when")	840	("when", "as")	848
("as", "if")	757	("if", "as")	705
("as", "so")	849	("so", "as")	893
("when", "if")	226	("if", "when")	222
("when", "so")	276	("so", "when")	288
("if", "so")	260	("so", "if")	204

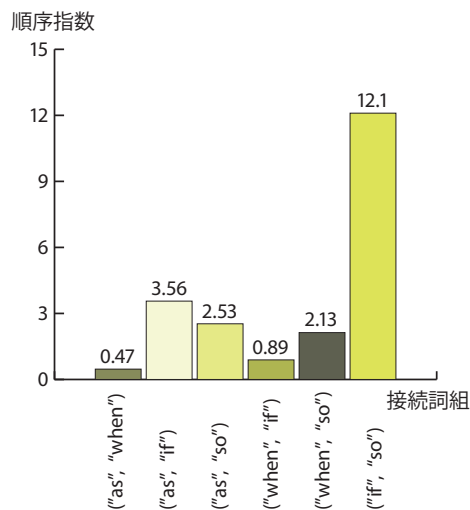


図 1 接続詞組に対する順序指数

考察

順序をもって現れる接続詞は確かに見つかった．しかしながら，(“if”, “so”) の順序が多いのは，文章展開による出現傾向というよりは，文単体での出現傾向を反映していると考えられる．これは，文の区切りをなくして，隣接した接続詞組のみを数えた方法によるので，文の区切りを捉えた方法であれば，文章展開の傾向に影響を受けた接続詞の出現傾向がわかる可能性がある．

おわりに

順序をもって現れる接続詞はあるかどうかを調べた．順序をもって現れる接続詞の組と，そうではない組が見つかった．(“if”, “so”) の組 1 つだけが，他の接続詞 5 組に比べ，際立った順序をもって現れている．

参考文献

- [1] Deborah Schiffrin, Deborah Tannen, Heidi E. Hamilton (2003) “The Handbook of Discourse Analysis” p.38. Wiley-Blackwell