

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

問題

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。特に、従属接続詞は、文と文とをつなぐ役割をもち、文章展開の一端を担っている。

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1億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。

$f(a,b)$ = 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、そのような接続詞対は、一定の順序を強く持っていると考えることができる。

結論

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。

結果

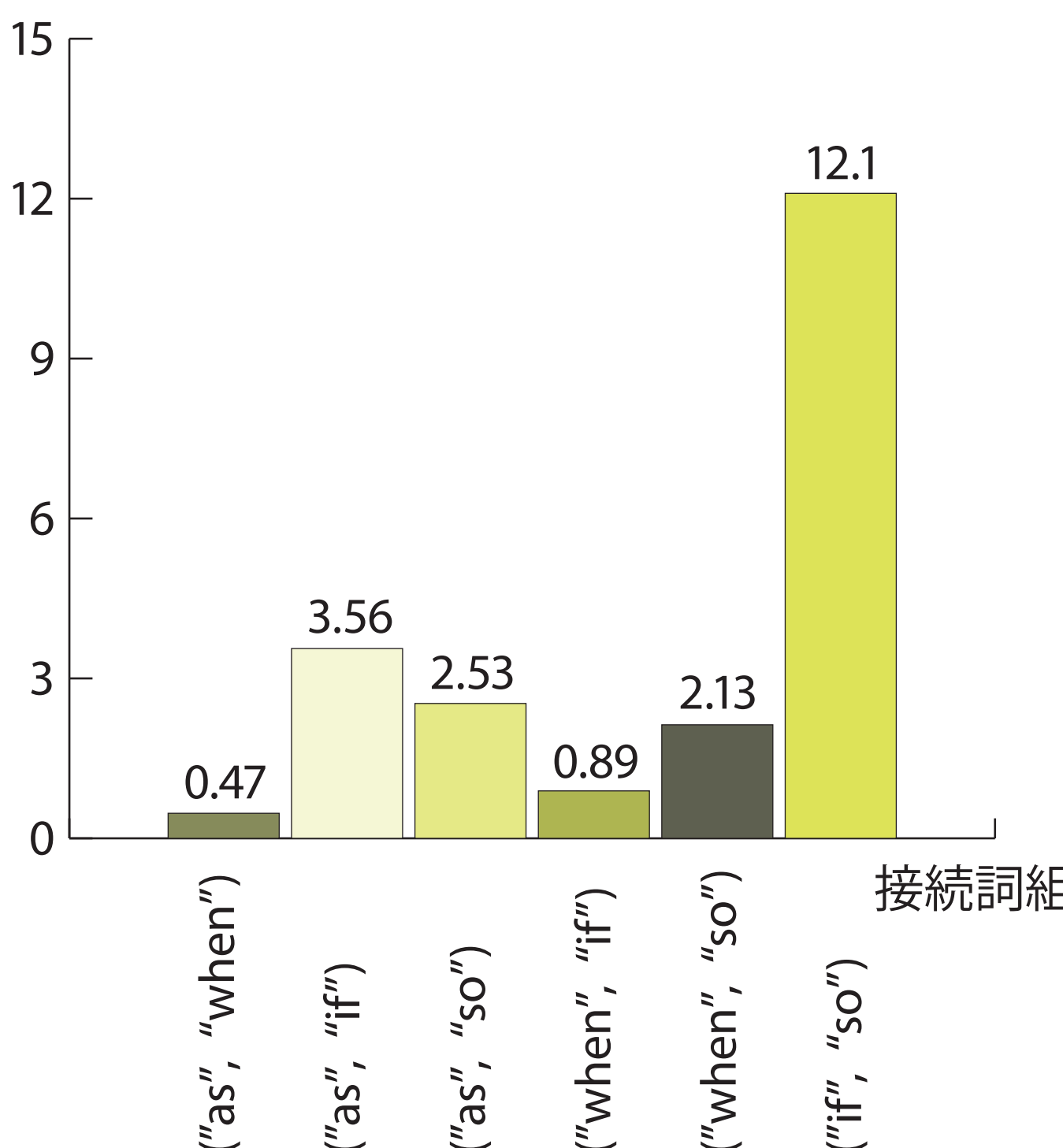
出現頻度の高い従属接続詞4つは、順に、**"as", "when", "if", "so"**であった。

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。

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

表：接続詞の順序対の出現頻度

順序指数



図：接続詞組に対する順序指数