

日本語での文字数は英語での二倍になるか：言語の文字数と情報量

虞 天昀

はじめに

普段話す時に、文字何個が口から出たかを注意する人はいないと思うけれども、文字数の比較は実際に作文、翻訳や出版など様々な場面でよく使われている指標だ。違う言語の間の文字数を換算する方法がわかれば、上記の場面に活用するだけではなく、言語の単位（日本語では漢字と仮名、英語では単語）の性質や文の構造を理解することにもつながっている。

リサーチクエスション：

一定の情報を表すために日本語での文字数は英語での二倍であるかどうかを検証する。

方法

二つステップを分けてクエスションの検証を進める。まずは単語をまわり、英語とそれに対応する日本語の名詞、動詞、形容詞と接続詞それぞれの単語を50個ずつ集め、文字数を比べ、英語の一単語に対し日本語の文字数の平均値を計算する。そして、45個の互いに対応する日本語と英語の文の字数をグラフに書いて、回帰分析を行う。

結果

ステップ1で、英語の一つ単語に対して、日本語は平均的に名詞で1.94、動詞で2.59、形容詞で2.83、接続詞で2.71、まとめて2.52文字が対応する。ステップ2で、回帰分析を通して、文の考察でも、日本語の文字数が英語の2.49倍であることがわかって、相関係数も0.93で、高い相関性を示している。

結論：日本語での文字数は英語での2.5倍である。

表（1）英語一つ単語あたりの日本語単語の平均文字数

名詞	動詞	形容詞	接続詞	合計
1.94	2.59	2.83	2.71	2.52

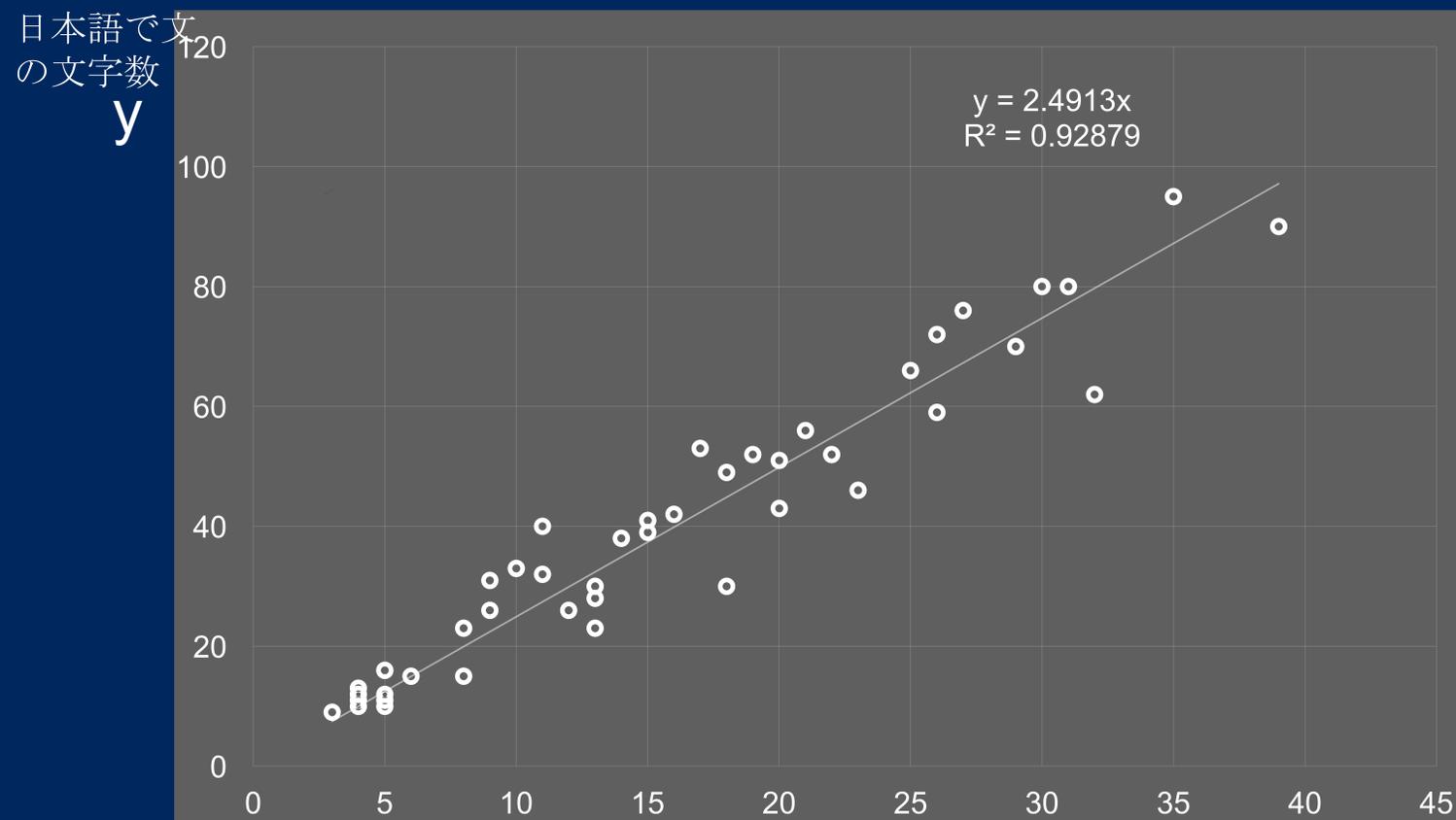


図1：文の文字数の比較

考察

結果としてステップ1も2も日本語の文字数は英語の2.5倍ぐらいであることを示している。

ステップ2で文の比較の結果から、日本語と英語の文字数は高い相関性の正比例を示した。そこで、文字数と情報量の関係を考えたら、言語相対論が述べたように言葉に対する理解と認識は文化と言語によって異なると言っても、情報量に対する理解は共通だと言ってもいいだろう。つまり、日本人が2倍の情報量をもつと思うメッセージは英語を話す人にとっても二倍の情報量を持っている。

一方で、2.5倍という数字を考察したら日本語は英語より文字数が多いと言っても、含まれる情報量が割に少ないというわけではない。なぜなら、英語と日本語の一番基本の単位はキャラクタと仮名だけれども、日本語では漢字という段階があり、漢字と仮名を合わせて英語での単語と相当する単位になるからだ。文字数を計算する際に、日本語で計算されているものは実際に英語の単語より細かい単位であることを考慮したら、かえって漢字を持つ日本語の方が情報の伝達が上手な言語であると考えても良い。

終わりに

日本語と英語の文字数について考察した。作文など文字数を換算する必要がある場合には日本語が英語の2.5倍と考えたらいいいのだ。

文献

フエルディナン・ド・ソシュール
(1972) 『一般言語学講義』、岩波書店。

横原恭士、『日本語における漢字の役割について』相愛大学研究論集 1992年

X 英語で文の文字数