

TED-Talks でチャンク音読をする方法と、 TED 字幕の教育的妥当性について

田淵 龍二 (ミント音声教育研究所)

背景と先行研究

「教育的妥当性」という概念には学習者や学習目的に応じて変動する相対的・主観的・主体的・歴史的な要素を多く含む。ここでは客観的な要素と思われる脳の機能との関係に限定する。具体的には、TED 字幕が作動記憶 (working memory) の制約に適合しているかとの観点である。作動記憶の制約とは「聞き取った音声は、反復されない限り 2 秒前後で減衰する」(Baddeley, 1996; Card, 1983) との理論に由来する。実際、映画音声の観測によると、呼気段落長平均は約 2 秒で、1 ~3 秒の間に 8 割の呼気段落が集中 (湯舟, 田淵 2013) していた。呼気段落 (breath group) とは吸気による無音区間で区分されるひと息の連続音声のことである。

理解を得るための発声では 2 秒前後の音声連続が適切であると考えられることから、筆者らは以下の仮説を立てている。

- (1) 語学教育では、平均 2 秒前後の音声 (あるいは黙読時の内声) が (特に初級者には) 学習対象としてふさわしい。
- (2) 外国語教育の場合には、学習者の負荷に配慮して平均 1.5 秒程度にする。

方法

TED-Talks 所収のビデオから 49 本を選び、その字幕長を解析した。7 本はトーキーズ (Talkies <http://www.mintap.com/talkies/>) 所収であり、残り 42 本は 6 つの分野から 7 本ずつ最新 (2015 年 10 月現在) ビデオ作品を選んだ。6 つの分野とは TED が Popular topics として提示している technology, entertainment, design, business, science, global issues である。字幕ファイル (.srt) には字幕の開始時間と終了時間が記載されているので、その時間幅をその字幕長とした。

結果

49 本のビデオの総再生時間は 563 分 (約 10 時間, 85,147 単語), 10,696 字幕、平均字幕長は 3.0 秒 (標準偏差 1.1 秒), 平均話速は 162wpm (標準偏差 40), 1 秒間の単語数に換算すると 2.5wps (words per second, $sd=0.4$), 1 字幕あたり平均単語数は 7.9 (標準偏差 3.2) であった (表 1, 図 2)。

考察

統計数字だけで判断できないことがある。プレゼンは「ひとり語り」で、映画は会話が主である。また一方は理屈や事例を積み重ねて主張を展開する講演であり、他方は映像に語らせる芸能である。こうした内容以外の問題点として字幕の質 (精度) がある。映画字幕製作は筆者が 10 年かけて詳細に呼気段落に分割した。他方、TED 字幕の製作方針は不明である。が、その出来を見る限り統一性はなく、音声との同期がずれていることが多い。例えば無音区間を無視している字幕もあ

る。総じてTED字幕長は実音声よりも長めに作られている様子がうかがえる。TED字幕は、一般的な映画字幕と同様に、音映像を見聞しながら文脈理解を補助することを主目的としていて、音声との同期は二の次（大雑把）と思われる。

表 1. TED-Talks 字幕の基本統計量と映画（m-Boxed 所収）との比較

	のべ数				1字幕あたり平均と標準偏差					
	作品	時間 min	字幕	単語	字幕長 sec	語数	話速 wps	sd	sd	
TED	49 本	563	1,0696	85,147	3.0	1.1	7.9	3.2	2.5	0.4
映画	11 本	1,277	1,9551	120,885	1.9	0.9	6.2	3.7	3.2	1.2

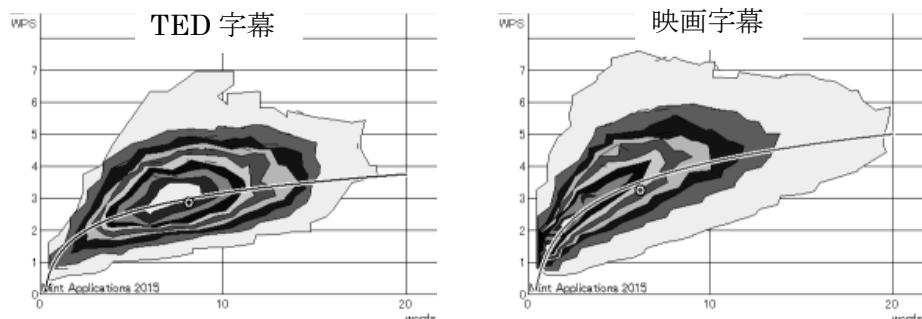


図 1. TED 字幕（左）と映画字幕における単語数（横軸）と話速（wps）の関係を示す 3D 等高線図。等高線は頻度を 10 等分。左下から右上に延びるのは話速曲線。映画に比べて TED 字幕は中央付近への集中が強いことが見て取れる。作図は SpeechContour (Mint Applications) による。

TED の字幕長は、学習に最適な仮説値（平均 1.5 秒）と比べると倍の長さであり、映画の 1.9 秒と比べても 1.5 倍である。語数にすると 1.3 倍で、話速はむしろ 2 割ほど遅いことが分かる。しかし先に述べた理由で、字幕長はもう少し短いと推測され、ニュース報道の平均呼気段落長 2.5 秒（湯舟、田淵 2013）と同じか少し長いと思われる。図 1 の分布比較を見ても、作動記憶の時間制限との関係では、TED-Talks の字幕長は教材として妥当な範囲にあることがわかった。

トーキーズ (<http://www.mintap.com/talkies/>) による字幕単位チャンク音声訓練

外国語学習者、特に初級者には負荷の大きい素材であると考えられることから、教材として利用する場合には配慮と工夫が必要である。トーキーズの字幕 1 回再生などの機能を使うことで 2~3 秒程度の音声に集中して音読やシャドーイングなどのチャンク単位での反復訓練が期待される。

参考文献

- Baddeley, A. D. (1996). *The fractionation of working memory*. Proc. Natl. Acad. Sci. USA Vol. 93, pp. 13468–13472, November 1996 Colloquium Paper
- Card, S. K., & Moran, T. P., & Newell, A., (1983). *The psychology of human-computer interaction*. Lawrence Erlbaum Associates, Inc.
- 湯舟英一, 田淵龍二 (2013). 「映画音声コーパスを利用した Breath Group 長の分析」. *Language Education & Technology*, 50, 23–41.