

Vermeer, A. (2004). The relation between lexical richness and vocabulary size in Dutch L1 and L2 children. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 173-189). Amsterdam, the Netherlands: John Benjamins.

報告者: 小泉利恵 (常磐大学)

## ■ 概要

即興のスピーチデータで見られる語彙の豊富さ (lexical richness) を測る指標:

- ・その多くは、異なり語 (type) と延べ語数 (token) の分布または関係に基づいて算出されるが、その信頼性も妥当性もないようだ

別の指標の提案: MRL (Measure of Lexical Richness 語彙の豊富さの指標)

- ・使用語彙の難易度 (9 レベルの頻度に基づく) 情報を用いて算出
- ・その妥当性検証
- ・MLR は type/token に基づいた指標よりも妥当なようだ

## 1. Introduction (pp. 174-177)

語彙の豊富さ: 話した、または書いたテキストにおいて、どの程度多様で広い語彙が使われるか

異なり語の割合 (Type Token Ratio; TTR): 異なり語数 ÷ 延べ語数

- ・問題点: ①値は低から高へ変わるとされるが、実際にはそう変わらない。習得の状態によって、パターンが変わりうる。

例: 異なり語数と延べ語数が、線形的に比例して多くなれば、TTR は一定の値をとる

②テキストの長さ、トピックの数に影響される

- ・実証研究で: 明らかに異なる語彙を持つ学習者を弁別しない。縦断的研究で語彙の成長が見られない

異なり語数、レマ数: 能力が高くなるにつれて、値が高くなる (図 1 参照)

- ・しかし、統語的成長や平均発話長 (Mean Length of Utterance) も測っている

Malvern & Richards (1997 など): テキストの長さの違いによる影響を克服した指標 (D) の提唱

①テキストからランダムに語をサンプルし、それぞれに TTR を算出し、グラフ化

②パラメータ値を調整することにより、ソフトウェア (vocd 小泉注: Meara & Miralpeix, 2004 も使える) が実証的データから得られたカーブと、モデルから算出した理論的カーブが最もよくフィットするところを探す

Jarvis (2002)

- ・明らかに語彙に違いのあるグループをあまりよくは弁別しない
  - ・語彙テスト得点・全体的評価と D の相関は、低いまたは中程度
- (Vermeer の意見) よって語彙の豊富さの妥当な指標では全然ない

語の難易度を考慮した指標の開発

- ・語の難易度と習得順序には関係がある
- ・語の知識と、その語が起こる頻度の相関は高い (Vermeer, 2001)
- ・語の頻度は習得順序と関係があるため、より妥当な語彙の豊富さの指標を出すには、頻度 (頻度のグループ) と発話中の語を関連付けられるかもしれない (might)
- ・絶対的な発表語彙のサイズの推定値を出せる

注意: 大人では、子どもより、語の頻度と習得順序は複雑

- ・子どもでは、語の頻度は習得順序と関連があり、語の頻度は語の難易度として考えてもよい

## 2. Towards an alternative measure of lexical richness (p. 178)

MLR: 日常の言語インプットでの相対的な頻度を計算し、それを語の難易度と捉えて指標を算出

頻度: 200 万語の延べ語数 (26000 レマ) を 9 つのカテゴリーに分けた

リサーチ・クエスチョン

(1) MLR は、顕著な語彙の差が見られる DL1 と DL2 を弁別できるか?

(2) MLR の併存的妥当性は何か?

(3) MLR は、(異なり語 小泉注: Results での言及なし) 発話の数とは独立か? 平均発話長とは独立か? (別のものを測っているか)

## 3. MLR procedure (pp. 178-182)

①即興の発話データを自動的に修正するコンピュータプログラムの開発

各語をレマ化して、可能性のある全ての意味を挙げる (例 1 参照)

レマを 1 語と考える

②挙げられた意味の 1 番目が最も頻度の高い意味を使って、MLR1 を算出

③文脈に合う意味を 1 つ 1 つ決める。固有名詞を取り除く。コロケーションなどを 1 つにつなげる作業をする

(the\_cold\_war): manual disambiguation  
→MLR2 を算出

④アウトプットでのレマを、9つの頻度カテゴリと付き合わせる

⑤表1の token coverage の相対的分布をモデルとする。語の分布が頻度1(高頻度)から頻度9までに渡っていたら、その話し手は発表語彙を26000語持ち、MLR得点は9  
語の分布が頻度1(高頻度)のみだったら、その話し手は発表語彙を1000語持ち、MLR得点は1  
(例3参照)

#### 4. Validation study (pp. 182–183)

##### 4.1 Subjects (p. 182)

オランダ語がL1の子ども (DL1):16人

オランダ語がL2の子ども (DL2):16人 (トルコまたはモロッコの背景あり。2世の移民。家では母語を使う)

オランダにて。共に grade 2。1人から約200の発話 [utterance] を収集

##### 4.2 Instruments (pp. 182–183)

30~40分の個別面接で集めた発話。物語を話す、友達など様々なトピックでインタビューを受けた。書き起こし後、分析。

受容語彙タスク:実験者が口頭で意味を言う→受験者が当てはまる絵を指差す  
受容語彙の絶対的なサイズを示す

定義タスク:与えられた語の意味を、受験者が説明・描写。1・0で評価

信頼性:  $\alpha = .95, .90$

各タスクと教師の評価の相関と、各タスクと即興の発話データとの相関が高かった

##### 4.3 Procedure (p. 183)

発話数、延べ語数、異なり語数、レマ数、平均発話長、Guiraud (異なり語/ $\sqrt{\text{延べ語数}}$ ) と比較

#### 5. Results (pp. 183–185)

(1) MLRは、顕著な語彙の差が見られるDL1とDL2を弁別できるか?:t検定(表2参照)

・MLR1では有意差なし。MLR2では有意差あり。よって、MLR1 < MLR2

・他の指標では有意差なし

(2) MLRの併存的妥当性は何か?:指標間の相関(表3参照)

高い相関:MLRと定義タスク ( $r = .61, .71$ )、MLRと受容語彙タスク ( $r = .45, .50$ )

→MLRの併存的妥当性はある

(3) MLRは、発話の数とは独立か?平均発話長とは独立か?(別のものを測っているか)

MLRと発話数・平均発話長の相関はほとんどない ( $r = -.03 \text{ to } .11$ )

→MLRはテキストの長さまたは統語的能力とは独立したものを測っている

レマ数は、受容語彙タスクと高い相関あり ( $r = .66$ )。しかし、平均発話長と発話数と高い相関があり、テキストの長さ・統語的能力と独立したものを測っているとはいえない

#### 6. Conclusions and discussion (pp. 185–187)

更なる研究

・語彙レベルがより高いグループで弁別できるか?

・信頼性のあるMLRの結果を生むには、最低でどのくらいの延べ語数・発話数や、トピック数が必要なのか?

MLRの欠点:算出に時間がかかること。特にMLR2。

MLRはtype/tokenを基にした指標よりも、ずっと良い指標のようだ

#### References

Meara, P., & Miralpeix, I. (2004). *\_lognostics D\_Tools* (Version 1.0) [Computer software]. University of Wales Swansea. Retrieved December 7, 2005, from <http://www.swan.ac.uk/cals/calsres/lognostics.htm>