

# 多言語コーパスを用いた幼児語彙獲得時期での男女間相関の特性

藤田 浩貴<sup>†</sup> 南 泰浩<sup>†</sup> 小林 哲生<sup>‡</sup> 奥村 優子<sup>‡</sup>

<sup>†</sup> 電気通信大学 大学院情報理工学研究科

<sup>‡</sup> NTT コミュニケーション科学基礎研究所

## 1. はじめに

我々はこれまで、幼児の語彙発達現象を大量のデータを使って調査してきた[1-3]。語彙発達の分野では、性別が幼児の語彙発達に及ぼす影響について様々な議論が行われている[4]。本稿では、10ヶ国語（日本語[5,6]・米語[7-9]・イタリア語[10,11]・クロアチア語[12]・ケベック語[13,14]・スペイン語[15,16]・スロバキア語・デンマーク語[17]・トルコ語[18]・ノルウェー語[19]・ラトビア語[20]）の横断データを用いて、幼児の理解語の獲得時期と発話語の獲得時期での男女間相関の特性を調べる。

## 2. 方法

### 2.1. 使用データ

幼児の理解語・発話語の習得時期を求めため、マッカーサー乳幼児言語発達質問紙 (MacArthur-Bates Communicative Developmental Inventories; 略して CDI) を用いて収集されたデータベースを利用する[21]。このデータベースを男女別に分割した。

### 2.2. 理解語・発話語の習得日齢の計算

語ごとに理解日齢・発話日齢・理解-発話日齢（理解から発話までの日数）を計算する。まず男女別に分割した CDI データを使って各語の男女別の語彙習得曲線を求める、その各々の曲線を次の二式で示すロジスティック関数でモデル化する。

$$f(x) = \frac{ae^{cx+b}}{1 + e^{cx+b}}$$

$$f'(x) = \frac{a'e^{c'x+b'}}{1 + e^{c'x+b'}}$$

ここで、 $f(x)$ は日齢  $x$  においてその語を理解した幼児の割合（理解率）を、 $f'(x)$ は日齢  $x$  においてその語を発話した幼児の割合（発話率）を示す。 $a, b, c, a', b', c'$ のパラメータは非線形の最小二乗法によって求める。図1に、米語の「elephant」という語の語彙習得曲線をモデル化した例を示す。図中の□は CDI のデータから求めた実際の理解率を、○は同様に発話率を表す。点線は□のデータから得られた  $f(x)$  の曲線で、実線は○のデータから得られた  $f'(x)$  の曲線を示す。図中の\*は、縦軸の値が 0.5 となる点である。ここで、 $f(x)=0.5$  となるときの  $x$  を「理解日齢」、 $f'(x)=0.5$  となるときの  $x$  を「発話日齢」、発話日齢と理解日齢の差を「理解-発話日数」と定義する。理解日齢および発話日齢はニュートン法によって計算した。

### 2.3. 男女間相関係数の算出

各言語データ中にある全ての語で、2.2 で定義した理解日齢・発話日齢・理解-発話日数を算出し、三つそれぞれの項目で男女間相関をとった。

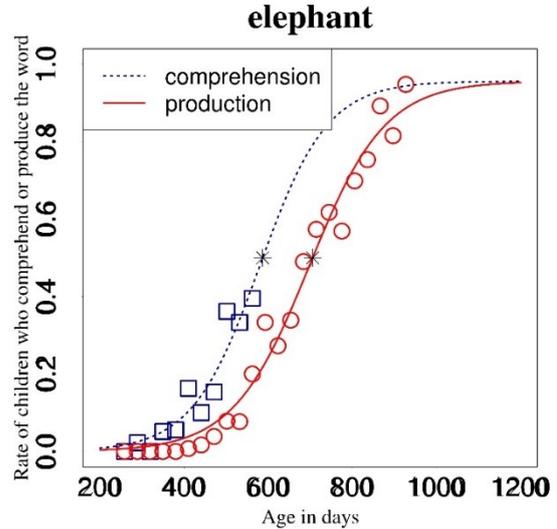


図1 「elephant」の語彙習得曲線をモデル化した例

## 3. 男女間相関結果

表1に各言語の語彙習得日齢を示す。この結果男女間相関はどの語彙習得日齢でも比較的高いことが分かる。

表1 各言語の語彙習得日齢の男女間相関係数

言語	理解	発話	理解-発話
Norwegian	0.96	0.98	0.92
Danish	0.93	0.98	0.91
French (Quebec)	0.85	0.96	0.86
Latvian	0.82	0.89	0.73
Spanish	0.89	0.94	0.68
English (American)	0.89	0.94	0.86
Croatian	0.82	0.89	0.72
Slovak	0.90	0.94	0.82
Italian	0.88	0.95	0.73
Japanese	0.82	0.87	0.61
Turkish	0.88	0.98	0.82

ここで注目して欲しいことは、発話日齢が必ず理解日齢の相関係数を上回ることである。このことを図2に示す。図から、この差が有意であることが分かる。

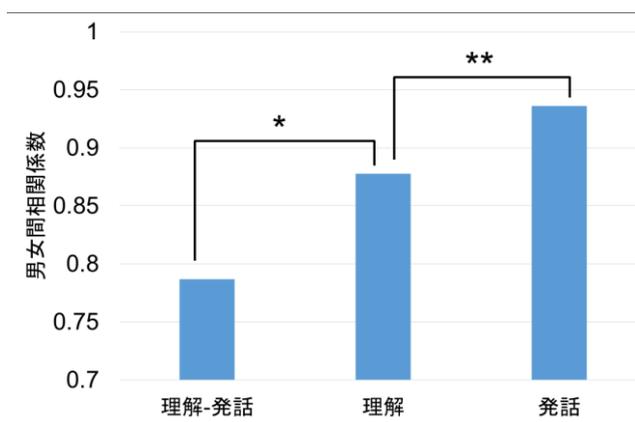


図2 男女間相関係数の平均値の比較  
(\* :  $p < 0.05$ , \*\* :  $p < 0.01$ )

#### 4. まとめと考察

本報告では、幼児の語彙発達機構の解明を目指すために、幼児の様々な語彙習得に対する性別の影響を調べた。この手段として、言語10種類分の横断データから各語男女ごとの理解日齢・発話日齢・理解-発話日数を計算し、性別間でこれらの値の相関係数を求めた。その結果、理解日齢・発話日齢・理解-発話日数のいずれも概ね高い男女間相関が見られた。また、理解日齢よりも発話日齢の方が、相関係数が大きくなることも確認された。

#### 謝辞

本研究の一部は、科学研究費助成事業基盤研究(B)17H02190による援助を受けた。

#### 文献

- [1] 南泰浩, 小林哲生, "幼児語彙理解・発話時期の言語共通性について", 電子情報通信学会技術研究報告, HCS2013-78, pp. 49-54, Feb.2014.
- [2] 南泰浩, 小林哲生, "幼児語彙習得順序における性別の影響について", 電子情報通信学会技術研究報告, HCS2014-8, pp. 61-66, May.2014.
- [3] 南泰浩, 小林哲生, "日本幼児の語彙習得順序に関する性別依存性について", 電子情報通信学会技術研究報告, HCS2014-103, pp. 173-176, Jan.2015.
- [4] Marjanovic-Umek, Ljubica; Fekonja-Pekljaj, Urška, Gender differences in children's language: a meta-analysis of Slovenian studies, CEPS Journal, vol.7, no. 2, pp. 97-111, 2017.
- [5] 小林哲生, 永田昌明, "ウェブ上で収集した幼児語彙発達データの信頼性検証", 言語処理学会第16回年次大会発表論文集, pp. 403-406, 2010.
- [6] 小椋たみ子, 綿巻徹, "日本の子どもの語彙発達の規準研究: 日本語マッカーサー乳幼児言語発達質問紙から", 発達療育研究, 24, pp. 3-42, 2008.
- [7] Fenson, L., Marchman, V. A., Thal, D., Dale, P., Reznick, J. S. & Bates, E., MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual. 2nd Edition. Baltimore, MD: Brookes Publishing Co, 2007.
- [8] Thal, D. J., Marchman, V. A. & Tomblin, J. B., Late talking toddlers: Characterization and prediction of continued delay. In L. Rescorla & P. Dale (Eds.). Late Talkers:

Language Development, Interventions, and Outcomes. Baltimore, MD.: Brookes Publishing, 2013.

- [9] Fernald, A., Marchman, V. A., & Weisleder, A., SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, vol. 6, no. 2, pp. 234-248, 2013.
- [10] Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., & Weir, J., A cross-linguistic study of early lexical development. *Cognitive Development*, vol. 10, no. 2, pp. 159-199, 1995.
- [11] Caselli, M. C., Rinaldi, P., Stefanini, S., & Volterra, V. Early action and gesture 'vocabulary' and its relation with word comprehension and production. *Child Development*, vol. 83, no. 2, pp. 526-542, 2012.
- [12] Kovacevic, M., Babic, Z., & Brozovic, B., A Croatian language parent report study: Lexical and grammatical development. Paper presented at the VIIth International Congress for the Study of Child Language, July 1996, Istanbul, Turkey.
- [13] Trudeau, N., & Sutton, A., Expressive vocabulary and early grammar of 16-to 30-month-old children acquiring Quebec French. *First Language*, 0142723711410828, 2011.
- [14] Boudreault, M. C., Cabirol, E. A., Poulin-Dubois, D., Sutton, A., & Trudeau, N., MacArthur Communicative Development Inventories: Validity and preliminary normative data. *La Revue d'orthophonie et d'audiologie*, vol. 31, no. 1, pp. 27-37, 2007.
- [15] Weisleder, A., & Fernald, A., Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24, pp. 2143-2152, 2013.
- [16] Jackson-Maldonado, D., Thal, D., Marchman, V., Newton, T., Fenson, L., & Conboy, B., MacArthur Inventarios del Desarrollo de Habilidades Comunicativas. User's Guide and Technical Manual. Brookes, Baltimore, 2003.
- [17] Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T. & Basbøll, H., The Danish Communicative Development Inventories: validity and main developmental trends. *Journal of Child Language*, 35, pp. 619-650, 2008.
- [18] Acarlar, F., Aksu-Koç, A., Küntay, A.C., Maviş, İ., Sofu, H., Topbaş, S., Turan, F., Adapting MB-CDI to Turkish: The first phase. In S. Ay, Ö. Aydın, İ. Ergenç, S. Gökmen, S. İşsever, and D. Peçenel (Eds.) *Essays on Turkish linguistics: Proceedings of the 14th International Conference on Turkish Linguistics*, Harrassowitz Verlag: Wiesbaden, Germany, August 6-8, 2008.
- [19] Simonsen, H. G., Kristoffersen, K. E., Bleses, D., Wehberg, S., & Jørgensen, R. N., The Norwegian Communicative Development Inventories: Reliability, main developmental trends and gender differences. *First Language*, vol. 34, no. 1, pp. 3-23, 2014.
- [20] Urek, Olga, Anna Vulāne, Roberts Dargis, Agrita Tauriņa, Tija Zīriņa, Hanne Gram Simonsen (to appear) *Latvian CDI: methodology, developmental trends and cross-linguistic comparison*.
- [21] <http://wordbank.stanford.edu/>