EXPLORING INDIVIDUAL DIFFERENCES AND CONTEXTUAL VARIATIONS IN CHILD LANGUAGE CORPORA

Hintat Cheung[*]   **Jing-chen Yang**
Hong Kong Institute of Education   *National Taiwan University*

ABSTRACT

Corpora are usually built to serve specific purposes. Child language corpora are constructed mainly for examining the course of development of a target group of children. As in other developmental studies, individual differences are commonly found. Individual differences can lead to growth curves of different slopes and unexpected plateaus. These inherent variations raise the question of how representativeness of a child language corpus can be determined. To this end, the present study examined the range of variations that are inherent to contextual variations. Child language samples archived in Taiwan Corpus of Child Mandarin (TCCM, http://taiccm.org/) were analyzed. Two types of language samples were compared: spontaneous conversational samples and narratives elicited in experimental settings. D, an index of lexical diversity in child language samples, as well as several other indices on language development were computed. Our findings suggested that conversational samples and narrative samples are quite different in their capacities in gauging linguistic development. D showed sensitivity to the early stages of language development in typically developing children while Verb Type showed age effect in children with Specific Language Impairment (SLI).

---

[*] Corresponding author: hintat.htc@gmail.com

探索兒童語料庫中的個別差異與語境變異

張顯達　　　　　　　　**楊靜琛**

香港教育學院　　　　　國立台灣大學

提要

兒童語料庫是以探究兒童語言發展為建置目的。兒童發展研究中經常報告個別差異的現象，如不同的成長曲線和預期之外的發展停頓。這一類的個別差異涉及如何判斷兒童語料樣本是否具備代表性。本文嘗試透過語料庫語料分析去探索這個方法學上的議題。語料是來自台灣兒童語料庫 (TCCM, http://taiccm.org/)，包括自發對話和誘發敘事說話兩種樣本。分析聚焦在詞彙的量化指標。結果顯示自發對話和誘發敘事說話這兩種不同語境中取的樣本在詞彙量化指標上表現不盡相同。詞彙多樣性指標較能反映初期的正常兒童語言發展，而動詞類別指標能夠呈現出語言障礙組的年齡差異。

關鍵詞

兒童語料庫　　**個**別差異　　**語**境變異　　**語**言發展

Linguistic Corpus and Corpus Linguistics in the Chinese Context