

台灣兒童語言語料庫(TCCM)語料說明

語料庫現有 362 份語料檔案，年齡層範圍由一歲半至八歲半。其中有 159 份的語料(3 歲至 8 歲)是來自特定型兒童語言障礙 (Specific Language Impairment) 兒童。以內容來區分，有自發性對話、故事共讀、看圖說故事三種。

為便利使用者引用資料來源，目前語料庫的檢索分類是以語料提供者分類：

1. CJC01 – 親子共讀 20 份 (張鑑如老師提供)
2. CJC02 – 親子對話 16 份 (張鑑如老師提供)
3. HTC01 - 自發性對話 127 份 (張顯達老師提供)
4. HTC02 – SLI (含對照組) 自發性對話、看圖說故事 199 份 (張顯達老師提供)

語料統計資料

Age\GRP	HTC01	CJC01	CJC02	HTC02/SLI	HTC02/C	total
1;0~1;11	18	/	/	/	/	18
2;0~2;11	58	20	/	/	/	78
3;0~3;11	47	/	16	3	5	71
4;0~4;11	4	/	/	12	12	28
5;0~5;11	/	/	/	47	12	59
6;0~6;11	/	/	/	46	11	57
7;0~7;11	/	/	/	43	/	43
8;0~8;11	/	/	/	8	/	8
File	127	20	16	159	40	362
Child utt.	33842	1065	831	1896	2328	39962
Adult utt.	63881	2398	1593	3991	5114	76977
Total utt.	97723	3463	2424	5887	7442	116939