

**PERMISSION TO DISTRIBUTE TRANSCRIPT DATA
THROUGH TALKBANK
Carnegie Mellon University**

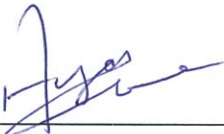
Carnegie Mellon University is participating in a data-sharing program known as TalkBank that was established through federal and foundation grants. For the program to function properly, many contributors will be asked to and will give permission for their work to be made available to other researchers. With a full understanding of the aforementioned, I hereby give permission to Brian MacWhinney, Director of the Center of the Child Language Data Exchange System at Carnegie Mellon University, to make and circulate electronic copies of the language transcripts and media that I describe below. These copies may be distributed to scholars and other responsible parties. I warrant that this use of the data is in accord with Human Subjects review procedures at my institution. I also warrant that there is no copyright restriction over the transcripts and media being circulated. Any further restrictions that I wish to place on the use of these data are listed under (2) below. I do not hold Brian MacWhinney or Carnegie Mellon responsible for the enforcement of these further restrictions and indemnify and render harmless both Brian MacWhinney and Carnegie Mellon University against any actions at law or in equity or in similar courts of any jurisdictions arising from violations of these restrictions.

1. General description of the data set.

*Please see separate word document.
(attached as appendix here)*

2. Restrictions to be placed on the use of the data:

No special restrictions.

Signed _____  Date 16 Oct 2017

Please mail this form to:
Brian MacWhinney, CMU-Psychology, 5000 Forbes Ave. Pittsburgh, PA, 15213, USA

HKPU (L2 Mandarin, L1 English)

Angel Chan

Department of Chinese and Bilingual Studies

The Hong Kong Polytechnic University

Hong Kong

angel.ws.chan@polyu.edu.hk

This corpus builds on a completed MA student dissertation by Jeff Zhen-Hui Feng with technical support by a doctoral student Wenchun Yang under the supervision of Angel Chan. It is a small video-linked L2 spoken Mandarin Chinese corpus featuring 14 adult subjects whose L1 is English and 6 L1 adult participants as controls. It is hoped that the addition of this corpus could further raise the visibility of SLA learner corpora featuring Chinese as the target language at Talkbank in this collaborative data-sharing international consortium.

Participants

Fourteen L2 adult learners of Mandarin (L1: English) and six L1 native Mandarin speakers (as L1 controls), aged between 20 and 70, were recruited. The 6 L1 participants were selected to match 6 L2 participants based on gender, age and education level to enable systematic L1-L2 comparisons for future research. Tables 1 and 2 below provide the background information of our L2 participants and the L1 participants respectively. All the L2 participants are able to communicate verbally in Mandarin Chinese at least at simple sentence level. All participants have given their written consent to participate in this project.

Table 1. Background Information of the 14 L2 Subjects (L1 English)

No.	Subject	Age	Gender	Education Level	Age at which learning of Chinese started	Contexts of acquisition	Other languages
1	Aa	23	M	Bachelor	22	Classroom/Conversation	French, Swedish
2	Al	34	M	Bachelor	30	Classroom/Conversation/ Reading/Self-learning	None
3	Ba	20	M	Bachelor	18	Classroom/Conversation	Spanish, German
4	Ga	42	M	Doctor	19	Conversation/ Reading/Watching TV and movies	Spanish
5	Ge	34	M	Master	27	Conversation	Spanish
6	Ja	31	M	Master	24	Classroom/Conversation/ Self-learning	Spanish, French
7	Je	38	M	Doctor	30	Conversation	French
8	Jo	37	M	High School	34	Classroom/Conversation	None
9	Mi	24	F	Master	18	Classroom/Conversation/ Reading	German, French
10	Mo	70	M	Doctor	22	Classroom/Conversation/ Reading/Self-learning	French
11	Na	23	M	Bachelor	20	Classroom/Conversation	French
12	Pa	48	M	Master	28	Classroom/Conversation/ Self-learning	German, French
13	Ph	24	M	Bachelor	20	Classroom/Conversation/ Self-learning	None
14	Ta	36	F	High School	33	Classroom/Conversation	None

Table 2. Background Information of the 6 L1 Mandarin Subjects

No.	Participant	Age	Education Level	Gender	L2 Matching Subject
1	Do	35	Doctor	M	Je
2	Gu	38	Bachelor	M	Ga
3	Qi	49	Bachelor	M	Pa
4	Wu	35	Doctor	M	Jo
5	Ya	24	Master	F	Mi
6	Zh	32	Master	F	Ta

Procedures

Speech samples were collected from each of the 20 participants on an individual basis. Each participant engaged in a structured narrative task, retelling in Mandarin Chinese the classic frog story “*Frog, Where Are You?*” (Mayer 1969, Berman & Slobin 1994) commonly used in cross-linguistic research. The process was videotaped with a high-quality audio track.

As is typical for studies using the frog story research tool, each participant reads the standard “frog story” storybook that tells a story in 24 pictures with no words, and then is asked to tell or retell the story in the target language. In addition, in this project, after reading the *wordless* storybook once, each participant (L1 and L2 participants alike) would listen once to the story narrated and audio-recorded in English with a standard story script to ensure that s/he became familiar with the story contents, before s/he retold the story in the target Mandarin language. The selected L1 Mandarin participants have the level of L2 English proficiency being able to understand the story narration in English.

The procedures in constructing the corpus follow the Talkbank format. We also conducted inter-person reliability checks of the transcriptions, the video-linking and synchronization of the data, as well as manual disambiguation of the automatic tagging.

Acknowledgment

The construction of this corpus was supported by a grant (project code: 1-ZVAQ) from the Hong Kong Polytechnic University to Angel Chan.

Please cite the following when using data from this corpus:

Chan, A., Feng, Z-H, Yang, W-C. (2013). A new multimedia shared L2 spoken Mandarin Chinese corpus: construction and linguistic analyses. Paper presented at the 21st Annual Meeting of the International Association of Chinese Linguistics (IACL 21).7-9 June 2013. Taiwan.