

The TAI Corpus : Longitudinal Speech Data of a Japanese Boy aged 1;5.20 – 3;1

Susanne Miyata

History

This data was collected during September, 1993 and June, 1995. Tai was after Ryo, Nao, and Aki (Miyata, 1992, 1993, 1995) the fourth child observed longitudinally. For Tai's observation I applied the same schedule used for the observation of the other children : that is once a week for about one hour at his home while playing with his mother.

In the previous observations it had proved convenient for both mother and observer, to fix weekday and time. In Tai's case, we decided to start about 10 o'clock in the morning. After a short period of preparation (video setting, and the indispensable cup of coffee for the observer), we would start with the recordings about 10:30.

The recordings were done simultaneously, the audio on mini-discs (MD) and the video on 8mm. This was done for two reasons : (1) the sound quality of the MD is considerably better than that available on video, while the video recording preserves the context and thus contains the information necessary to make informed judgements about the utterances of the child; and (2) to countercheck the rather low reliability of the equipment. In fact, out of 75 MD recordings, 3 were not usable as a result of tape damage and battery problems (especially in the cold). The additional video recording could compensate for these lapses.

For the recording, the video camera was placed on top of the TV set in the corner of the 16qm room. Using a fish eye lens, as well as a microphone with a recording angle of 90 degrees, most of the sound and movement taking place in the room could be captured. Unlike Aki, Tai showed no interest in the recording equipment, so it could be left unattended. The living room opened on a kitchen of the same size, which was designated as a "playroom" for the purposes of the observation sessions, and the child quickly

accepted it as such. As he grew older, Tai would prepare his toys and a cushion (zabuton) for the observer, urging us to start with the play session immediately.

The observer would sit in the second corner, as passive as possible, in order not to disturb the mother-child interaction. The setting was free indoor play. The mother was instructed to 'make the child speak'. In order to obtain as many free spontaneous speech from the child as possible, she was told not to entertain 'not too much' story telling and singing. The recording time was a little more than 40 minutes, and was cut down to 40 minutes in the transcriptions. After the recording we would sit down in the kitchen and discuss the development of the child, his friendship relations, and his health, as well as general issues of education.

Transcription

The sound data was computerized, and sound-linked to CHAT files (MacWhinney, 2000). The transcription was done on the base of the beforehand linked sound stretches. The easiness to access the sound (it is possible to listen to an utterance just with one mouse-click) proved to be very convenient during this process.

The transcription was done in Latin script (Hepburn system) following JCHAT 1.0 Hebon (Oshima-Takane & MacWhinney, 1995). Word separation follows WAKACHI98 (Miyata & Naka, 1998). For unclear sound stretches I have used UNIBET for Japanese (Terao, 1995).

Biographical Data

Tai was born on April, 10th, 1992 in Nagoya, the firstborn child. His mother was 28 years old at the time of his birth. Pregnancy and delivery were normal. Tai's birth weight was 3330 g. His physical development was normal, and he was healthy throughout the observation.

Tai was an active, curious, and sensitive child, with a long concentration span. He displayed a high sense of responsibility. His pronunciation was very clear. At present (March, 2000) he is a healthy and awake first grader with excellent records.

Other participants:

TMO Mother, called “Kakka”, 29 years, housewife, former secretary at a University in Nagoya. Educational level 15

TFA Father, called “Totto”, 30 years, research engineer. Educational level 15

SUU Investigator, called “Suuchan”, friend of TMO

Pseudonyms

Tai’s parents gave their kind consent for the publication of this data. Although they consented to the use of their actual names, I have decided to anonymize all last names (except my own) and other identifying information throughout the corpus in order to preserve a certain amount of privacy.

Table of Contents

File No.	File Name	Age	Minutes	MLUm (based on a ll utterances)
1	T930930	1;5.20	40	1.514
2	T931007	1;5.27	40	1.591
3	T931014	1;6.4	30	1.288
4	T931021	1;6.11	40	1.440
5	T931029	1;6.24	40	1.924
6	T931103	1;6.24	40	1.924
7	T931111	1;7.1	40	1.477
8	T931118	1;7.8	40	1.635
9	T931125	1;7.15	40	1.820
10	T931223	1;8.13	40	1.691
11	T940107	1;8.28	40	2.105
12	T940113	1;9.3	40	2.329
13	T940120	1;9.10	40	2.331
14	T940127	1;9.17	40	2.180
15	T940204	1;9.25	40	2.223
16	T940210	1;10.0	40	2.235
17	T940217	1;10.7	40	2.313
18	T940224	1;10.14	40	2.233
19	T940303	1;10.20	40	2.348

20	T940311	1;11.1	40	2.467
21	T940324	1;11.14	40	2.739
22	T940330	1;11.20	40	2.529
23	T940407	1;11.28	40	3.306
24	T940414	2;0.4	40	2.519
25	T940421	2;0.11	40	2.471
26	T940428	2;0.18	40	2.689
27	T940505	2;0.25	40	2.929
28	T940512	2;1.2	40	3.042
29	T940519	2;1.9	40	3.004
30	T940526	2;1.16	40	3.248
31	T940602	2;1.23	40	3.737
32	T940609	2;1.30	40	3.368
33	T940616	2;2.6	40	3.178
34	T940623	2;2.13	40	3.178
35	T940630	2;2.20	40	3.016
36	T940707	2;2.27	40	3.609
37	T940714	2;3.4	40	3.413
38	T940721	2;3.11	40	2.831
39	T940728	2;3.18	40	3.288
40	T940804	2;3.25	40	2.998
41	T940813	2;4.3	40	3.102
42	T940825	2;4.15	40	2.934
43	T940831	2;4.21	40	3.158
44	T940909	2;4.30	40	3.425
45	T940916	2;5.6	40	2.916
46	T940922	2;5.12	40	3.325
47	T940929	2;5.19	40	3.564
48	T941006	2;5.26	40	3.134
49	T941013	2;6.3	40	3.486
50	T941020	2;6.10	40	3.688
51	T941028	2;6.18	40	4.036
52	T941103	2;6.24	40	3.182
53	T941110	2;7.0	40	3.252
54	T941117	2;7.7	40	3.563

55	T941123	2;7.13	40	3.516
56	T941201	2;7.21	40	4.006
57	T941208	2;7.28	40	3.932
58	T941215	2;8.5	40	4.486
59	T941222	2;8.11	40	4.040
60	T950112	2;9.2	40	4.175
61	T950119	2;9.9	40	4.779
62	T950127	2;9.17	40	3.806
63	T950202	2;9.23	40	3.133
64	T950209	2;9.30	40	4.261
65	T950216	2;10.6	40	3.286
66	T950223	2;10.13	40	4.085
67	T950302	2;10.20	40	3.663
68	T950310	2;11.0	40	4.059
69	T950324	2;11.14	40	5.003
70	T950330	2;11.20	40	5.672
71	T950413	3;0.3	40	4.227
72	T950504	3;0.24	40	5.058
73	T950511	3;1.1	40	4.923
74	T950518	3;1.8	40	4.133
75	T950608	3;1.29	40	3.787

Warnings

- a) Reliability was not checked.
- b) Comments and descriptions concerning child activities are not yet supplied. They will be added in a later version.

File Sample

@Begin

@Participants : TAI Taishoo Child, TMO Kakka Mother, SUU Suuchan
Investigator

@Filename : Tai941117. cha

@Age of TAI : 2;7. 7

@Date : 17-NOV-1994

[. . abbreviated. .]

- *TAI : koko ni osuwari shite!
- *TMO : e@i?
- *TAI : osuwari , , tsukareta kara +. . .
- *TMO : un .
- *TAI : +, osuwari shite!
- *TAI : koko .
- *TAI : osuwari .
- *TMO : nande?
- *TMO : nande suwaru no@fp?
- *TMO : xxx .
- *TAI : anoo koarachan mottekiteageru kara .
- *TMO : e@i?
- *TMO : ii yo .
- *TAI : nande?
- *TMO : n?
- *TAI : nande?
- *TMO : oide yo , , Taishoo , , koko ni!
- *TAI : anoo kok(o) [//] Taishookun wa Dakkochan de +. . .
- *TAI : Kakka wa # koarachan de +. . .
- *TAI : de # yonin de # xxx de +. . .
- *TMO : nani [>] [= ! laughing]?
- *TAI : de : [<] +. . .
- *TAI : sonde oshimai da : [?] .
- *TMO : nani ga oshimai na no@fp?
- *TAI : xxx .
- *TAI : koarachan .
- *TAI : ima kara Dakkochan nyuuin shinan yo .
- *TMO : nyuuin suru no@fp?
- *TAI : n(y)uuin +. . .
- *TAI : ashita kara n(y)uuin suru no@fp .
- *TAI : ashita kara n(y)uuin suru no@fp .
- *TMO : nyuuin suru no@fp?
- *TAI : un .

- *TMO : nande?
 *TAI : u'un .
 TAI : koarachan itakunai tte ittoru mon da [] tte .
 *TMO : ha@i?
 *SUU : ha@i?
 TAI : ittoru mon da [] tte .
 *TMO : < nani ga > [//] nan [//] nan t(t)e ittoru?
 *TMO : ja < Koarachan doozo > [>] .
 *TAI : < koko ni chiku:@o > [<] .
 *TMO : janjakajan@o xxx .
 *TAI : ashita kara n(y)uuin sun no@fp .
 *TMO : he@i?
 *TAI : ashita kara .
 *TMO : ashita kara?
 *TAI : un .
 *TMO : nyuuin suru toki kore noru no@fp?
 *TAI : anoo ashita kore noru no@fp .
 *TAI : daka(ra) konaide!
 *TMO : n?
 *TAI : anoo ne +. . .
 *TAI : anoo ashita nyuuin suru kara ne .
 *TMO : un .
 *TMO : koarachan@v .
 *TMO : doo shita no@fp -? , koarachan@v .
 *TMO : daijoobu?
 *TMO : ne .
 *TMO : erai tte -? , , koarachan?
 *TAI : un .
 *TAI : anoo Dakkochan ne +. . .
 *TMO : un .
 *TAI : +, sugoi batsu [/] batsu toru nda tte .
 *TMO : ha@i?
 *TMO : nani [/] nani?
 *TAI : gatsun@o tte kore &otot:e [?] kuru nda tte .
 *TAI : omocha asobu nda tte .

*TMO : hontoo [>] ?
*TAI : iku [<] nda tte .
*TMO : doko e ?
*TAI : iku nda tte , , Taishookuntachi .
*TMO : un .
*TMO : asoboo ka , , ja .
[. . abbreviated. .]
@End

Acknowledgments

I gratefully acknowledge the support of this research by the Ministry of Education, Science, Sports and Culture through the Grant-in-Aid for Scientific Research on Priority Areas 10114104 entitled “Development of Mind”, and through the Grant-in-Aid for Scientific Research (Database) 184. I would like to thank Brian MacWhinney (Carnegie Mellon University) for his understanding technical support during the various phases of transcription, the members of the JCHAT Project for their encouraging supportment, and the numerous students who helped with the transcription, especially Yumiko Naganawa and Naomi Hamasaki. My special thanks go to Beverley Curran (Aichi Shukutoku University) for the emotional support and encouragement throughout this work. My warmest thanks though go to Tai and his mother. Without their understanding collaboration, this project would not have been possible.

Literature

- MacWhinney, B. (2000). *The CHILDES Project : Tools for Analyzing Talk*. 3rd ed. Mahwah, N. J. : Lawrence Erlbaum Assoc.
- Miyata, S. (1992). Wh-Questions of the Third Kind : The Strange Use of Wa-Questions in Japanese Children, *Bulletin of Aichi Shukutoku Junior College* No. 31, p. 151-155
- Miyata, S. (1993). *Japanische Kinderfragen : Zum Erwerb von Form - Inhalt - Funktion von Frageausdruecken*, Hamburg : OAG.
- Miyata, S. (1995). The Aki Corpus. Longitudinal Speech Data of a Japanese Boy aged 1.6-2. 12. *Bulletin of Aichi Shukutoku Junior College* No. 34, 183-191
- Miyata, S. (in press). Assigning MLU stages in Japanese. *Journal of Educational Systems and Technologies. The Audio Visual Center, Chukyo University Nagoya* Japan.

Vol. 9.

- Miyata, S. & N. Naka. (1998). Wakachigaki Gaidorain WAKACHI98 v. 1.1. *Educational Psychology Forum Report* No. FR-98-003. The Japanese Association of Educational Psychology.
- Oshima-Takane, Y. & B. MacWhinney (eds.) (1995, 2nd ed. 1998). *CHILDES Manual for Japanese*. Montreal : McGill University / Nagoya : Chukyo University.
- Sugiura, M. , N. Naka, S. Miyata & Y. Oshima. (1997). Nihongo Shutoku Kenkyu no tame no Joho Shisutemu CHILDES no Nihongoka. *Gengo*, 26, 3, 80-87.
- Terao, Y. (1995). Nihongo no tame no UNIBET. Oshima-Takane, Y. & B. MacWhinney (eds.) (1995). *CHILDES Manual for Japanese*. Montreal : McGill University. 97-100.