2018. In Proceedings of the 32nd Pacific Asia Conference on Language, Information and Computation.
The Hong Kong Polytechnic University, Hong Kong.

Communicative Efficiency in Child Mandarin

Jane Tsay
Institute of Linguistics
National Chung Cheng University
Taiwan
Lngtsay15@gmail.com

James Myers
Institute of Linguistics
National Chung Cheng University
Taiwan
Lngmyers@ccu.edu.tw

Abstract

Communicative efficiency in this study is quantified both as transmission efficiency and as representation efficiency, adopting an explicit method proposed by Myers, Tsay and Su (2011). Transmission efficiency is defined as the amount of information transmitted during a specific time and is measured by the rate of propositions per second. On the other hand, representation efficiency is defined as the amount of information represented in a specific grammatical unit (e.g., syllable) and is measured by the rate of propositions per syllable. In this study, we are especially interested in the development of child language in communicative efficiency. Narratives were elicited from 35 seven-year-old Mandarin-speaking children from an elementary school in Taipei using the picture storybook "Frog, where are you?" (Mayer 1969). It was found that, compared with the results of 29 adults (mean age 37 years), the representation efficiency of the children was the same as that of the adults, although their transmission efficiency was significantly lower than that of the adults. The results show that the children's grammatical structure in speech already conveys the same amount of meaning or information as adults by the age of seven when their articulation is still much slower than the adults.

1 Introduction

1.1 Defining communicative efficiency

The primary goal of human language is to communicate. However, how communicative efficiency of a human language can be quantified is not a simple question. For example, articulation rate has been widely used to assess children's speech skills (e.g., Goldman-Eisler, 1968), children with stuttering problems or other specific language impairments (e.g., Hall et al., 1999; Erdemir, et al., 2018), and, second language proficiency (e.g., Bergmann et al., 2015). Since the rate of articulation is measured by the number of syllables per minute of the time spent in vocal activity (pauses subtracted out) (Goldman-Eisler, 1968) or the number of syllables or phones per second excluding pauses and disfluent segments (Erdemir et al., 2018), it is concerned only with the speed of articulation. However, language communication is not just about speed. More importantly, it is about "what" (meaning) is being communicated. The rate of articulation does not reflect how much linguistic information, especially meaning as expressed by propositions, is conveyed or represented in a linguistic unit like a syllable.

Therefore, Myers, Tsay, and Su (2011) propose that in order to capture a complete picture of communicative efficiency, both representation efficiency and transmission efficiency should be taken into consideration. Representation efficiency is defined as the number of propositions "represented" per syllable and transmission efficiency is defined as the number of propositions "transmitted" per second.

1.2 Previous studies on communicative efficiency

The study of Myers et al. (2011) was originally inspired by Bellugi and Fisher (1972) which focuses on a cross-modality comparison between English and American Sign Language (ASL).

Bellugi and Fisher (1972) compared the narratives in English and ASL by three bilinguals

and found that, although a sign in ASL took longer to produce than a spoken word in English, a proposition took about the same amount of time to produce in either language. However, their measurements were focusing only on speed, i.e., the rate of articulation. That is, they only measured the transmission efficiency.

Therefore, Myers et al. (2011) propose to measure both representation efficiency and transmission efficiency, and provide an explicit quantification method for the measurements, as schematized below for both signed and spoken languages. (Figure 1 and Figure 2 are from Myers, et al. (2011), p. 173 and p. 174, respectively.)

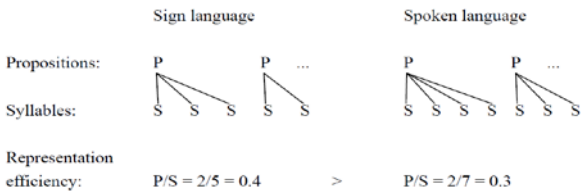

Figure 1. Calculation of representation efficiency

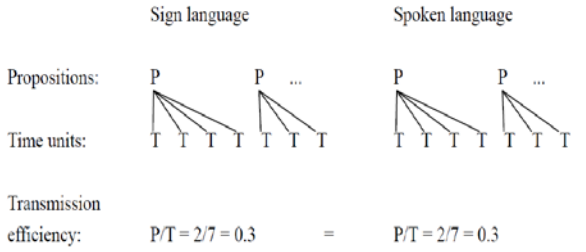

Figure 2. Calculation of transmission efficiency

1.3 The current study

This model has been used to study communicative efficiency in adults for cross-modality comparison between Mandarin and Taiwan Sign Language (TSL) (Myers et al. 2011) and longitudinal comparison in both Mandarin and TSL (Tsay, Myers, and Tai, in press).

Since transmission efficiency reflects the speed of articulation, it might be expected to be affected more by maturing motor control. By contrast, representation efficiency reflects a language's structural complexity and hence might show the process of children's development in grammatical knowledge. Therefore, the current study investigates Mandarin speaking children's communicative efficiency in comparison with that of the adults. To our knowledge, this is the time that this model has been used in child language.

2 Methods

The methods are the same in both the adult study and the child study. The steps are as follows: (1) data collection; (2) transcription of the recorded narratives; (3) measurements; (4) calculation of representation efficiency and transmission efficiency.

2.1 Data collection: elicitation of narratives

Narratives are ideal texts with naturally expressed utterances. The picture storybook "Frog, where are you?" (Mayer 1969) was used to elicit narratives. There are several advantages of using a picture storybook. First of all, it does not contain written words and is neutral to speakers of different languages. Second, it provides a topic for the speakers to talk about. Third, the storybook shows a coherent story for the speakers to elaborate. Moreover, the content of the narratives will be more consistent across speakers, compared with free narratives which might show very high variations.

Thirty-five Mandarin-speaking first graders from an elementary school in Taipei were recruited for recording the narratives of the frog story. They aged from 6 years 7 months to 7 years 6 months with an average of 7 years and one month.

Each child was recorded separately. The child was shown with the picture book and was allowed to look through the whole book to understand the story. The child was then asked to tell the story page by page.

2.2 Transcription of the recorded narratives

The recorded narratives were transcribed using the transcription tool CHAT of the Child Language Data Exchange System (CHILDES, MacWhinney, 2014), with one tier (*CHI or *INV) in romanization (Mandarin Pinyin), one tier (%ort) in Chinese characters, and a third tier (%cod) with morph-syntactic annotations.

*CHI: xiao3 nan2hai2 zuo4 zai4 na4bian1 kan4 zhe0 qing1wa1.
%ort: 小 男孩 坐 在 那邊 看 著 青蛙.
%cod: VH Na VA P Ncd VC Di Na
*INV: ran2hou4 ne0?
%ort: 然後 呢?
%cod: D T
*CHI: xiao3 nan2hai2 shui4jiao4 le0.
%ort: 小 男孩 睡覺 了.
%cod: VH Na VA Di
*CHI: qing1wa1 pao3chu1lai2 le0.
%ort: 青蛙 跑出來 了.
%cod: VH Na VA Di

Figure 3. A sample transcript

2.3 Measurements

Following Bellugi and Fisher (1972) and Myers et al. (2011), propositional utterances were first identified by main verbs and/or adjectival predicates, and the number of syllables in each propositional utterance was counted.

The durations of propositional utterances were then measured using the software Praat (Boersma and Weenink, 2018). Pauses between utterances were excluded following the literature mentioned above.

The following utterance *ta ba boli dapo le* 他把玻璃打破了"He broke the glass" contains one proposition (the main verb *dapo* "break") and has 7 syllables. The duration of this utterance is 2.15 seconds.

In counting the number of syllables in each propositional utterance, in addition to aspect markers, such as *le* 了 in the example, final particles are also included because they carry pragmatic meanings or functions.

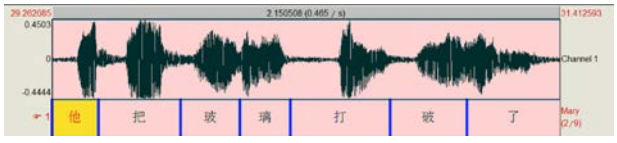

Figure 4. Measurement of duration and segmentation of syllables

2.4 Calculation of representation efficiency and transmission efficiency

Representation efficiency was calculated by dividing the number of propositions by the number of syllables in each propositional utterance and transmission efficiency was calculated by dividing the number of propositions by duration (in seconds), as demonstrated in Figure 1 and Figure 2, respectively.

For example, in the above one-proposition utterance with 7 syllables and 2.15 seconds in duration, the rate of representation efficiency (propositions per syllable) is 1/7=0.14 and the rate of transmission efficiency (propositions per second) is 1/2.15=0.47.

3 Results

The results of the 35 children are as follows: the average rate of representation efficiency is 0.12 and the average rate of transmission efficiency is 0.38.

To understand children's language development, these results are compared with the results of the 29 adults reported in Myers, et al. (2011), where the same method was used in eliciting narratives and measuring adults' representation efficiency and transmission efficiency. These 29 adults were aged from 17 to 61 years old with an average age of 37 years.

3.1 Results of representation efficiency

The mean representation efficiency was 0.12 per syllable for the children, the same as that of the adults.

	Children (N=35)	Adults (N=29)
Representation Efficiency	0.12	0.12

Table 1. Representation efficiency

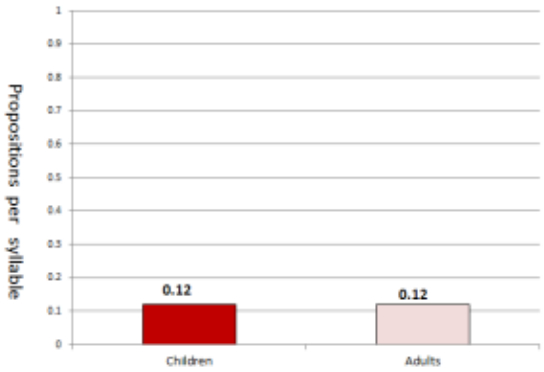

Figure 5. Representation efficiency

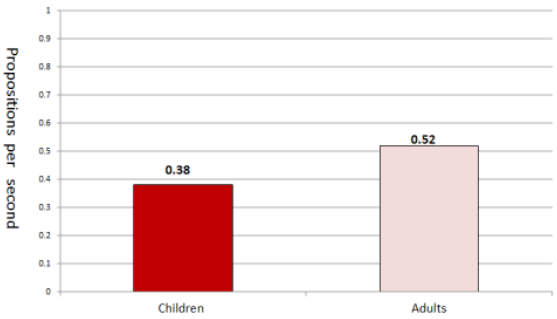

Figure 6. Transmission efficiency

The results show that the 7-year-old children could convey as much information or meaning as the adults within the same linguistic representation.

Nevertheless, the children showed significantly greater variance (0.0004) than the adults (0.0001) (one-tailed variance test: $F(34,28) = 2.59$, $p < .01$). Even though the children managed to encode linguistic information in a similar way as the adults, their greater variation suggests that they were not as consistent in how they did this.

3.2 Results of transmission efficiency

The mean transmission efficiency of the children was 0.38 per second and was significantly lower than the 0.52 mean of the adults.

	Children (N=35)	Adults (N=29)
Transmission Efficiency	0.38	0.52

Table 2. Transmission efficiency

Since variance in transmission efficiency was marginally higher in the adults (0.008) than in the children (0.004) (one-tailed variance test: $F(28,34) = 1.71$, one-tailed $p = .07$), we compared the means using an unpaired t test not assuming equal variance. This showed that mean transmission efficiency was significantly higher for the adults (0.52) than for the children (0.38) ($t(52) = 7.13$, $p < .0001$).

The lower transmission efficiency of the children can be explained by the slower articulation of the children. This can also be demonstrated by the rate of articulation which is calculated by the number of syllables per second.

The children's articulation rate was 2.86 syllables per second, compared to 4.20 syllables per second by the adults.

	Children (N=35)	Adults (N=29)
Syllable per second	2.86	4.20

Table 3. Rate of articulation

4 Discussion and conclusion

The results show that these seven-year old children already had adult-like representation efficiency, although they were still significantly lower in transmission efficiency than the adults.

This model can be used to supplement the mean length of utterance (MLU) measure which has long

been considered as a reliable index for children's grammatical complexity since Brown (1973).

As a comparison with our efficiency measures, we also calculated the MLU of the children and the adults in this study. The number of syllable per utterance (MLUs) was counted, instead of the number of word per second (MLUw), following Cheung (1998), which found that there was a high correlation between using the syllable and the word as the counting unit in Mandarin Chinese.

The mean MLU of the children in our study was 10.96, significantly lower than the 13.76 of the adults ($t(58) = 5.6$, $p < .0001$). Child MLU did not predict the variation in their transmission efficiency ($r(34) = -.15$, $p = .4$)

Despite the children's utterances being shorter than those of the adults, they still seem long enough to have conveyed the same amount of information, even beyond the main verbs and adjectival predicates that we counted.

Curiously, the utterances of the 7-year olds in our study seem to be longer than what has been found in the literature for Mandarin-speaking children. For example, Cheung (1988) found that the mean MLUs for the 7-year olds was 6.78, and Tsou and Cheung (2007) found that the normal (control group) 5-6 year olds' MLUs was 5.71. This difference may relate to discourse context: describing a detailed picture story may simply require more words than everyday conversation.

There are other considerations for future research. One concerns the narrative style of storytelling. Would a different narrative style make a difference either for the children or for the adults? Or, would the results be different in a typologically different language? Another consideration is to study even younger children to find out when children start having adult-like linguistic representation.

There are also methodological issues to take into consideration. For example, the identification of propositional content by main verbs and adjectival predicates might be re-examined in languages with rich nominalization (e.g., gerunds in English, Zucchi 2013).

Acknowledgments

The child data were based on transcripts from a previous study of the first author funded by the National Science Council, Taiwan (NSC 94-2420-H-194-003-KD). Research assistant Yi-Chih Chan helped with the recordings and transcriptions of the children's narratives. The measurements were done with the help of research assistants Shihkai Liu and Yushan Yen. This study was also supported by a research grant from the Ministry of Science and Technology, Taiwan (MOST 104-2410-H-194-079-MY3). The reviewers' comments are also acknowledged. Any defects are our sole responsibility.

References

Bellugi, Ursula and Susan D. Fischer. 1972. A comparison of sign language and spoken language. *Cognition* 1. 173-200.

Bellugi, Ursula, Susan D. Fischer, and Donald Newkirk. 1979. The rate of speaking and signing. The Signs of Language, ed. by Edward S. Klima and Ursula Bellugi, 181-194. Cambridge, MA: Harvard University Press.

Bergmann, C., Sprenger, S. A., and Schmid, M. S. 2015. The impact of language co-activation on L1 and L2 speech fluency. *Acta psychologica, 161*, 25-35.

Boersma, Paul and Weenink, David. 2018. Praat: doing phonetics by computer [Computer program]. Version 6.0.40, from http://www.praat.org/

Brown, R. 1973. *A first language: The early stages.* Harvard U. Press.

Cheung, Hin-Tat. 1998. Pingjun yuju changdu zai zhongwen de yingyong [The application of MLU measurement to Mandarin Chinese]. *Ting Yu Hui Kan [Journal of Hearing Science]* 13 (1998): 36-48. 張顯達. 平均語句長度在中文的應用.《聽語會刊》(13)，36-48.

Erdemir, A., Walden, T. A., Jefferson, C. M., Choi, D., and Jones, R. M. 2018. The effect of emotion on articulation rate in persistence and recovery of childhood stuttering. *Journal of fluency disorders, 56*, 1-17.

Goldman-Eisler, Frieda. 1968. Psycholinguistics: Experiments in spontaneous speech. New York: Academic.

Hall, K. D., Amir, O., and Yairi, E. 1999. A longitudinal investigation of speaking rate in preschool children

who stutter. Journal of Speech, Language, and Hearing Research, 42, 1367–1377.

Kuo, Lin-Ju. 2014. *Taiwanese Elderly Adults' Narrative Ability in Southern Min - A Study on Mean Length of Utterances*. MA thesis, National Chung Cheng University, Taiwan. 郭玲汝. 2014. 從平均語句長度探討台灣高齡者閩南語敘述能力. 碩士學位論文. 國立中正大學語言學研究所, 臺灣.

MacWhinney, Brian. 2014. *The CHILDES project: Tools for analyzing talk, Volume I: Transcription format and programs*. Psychology Press.

Mayer, Mercer. 1969. Frog, where are you? New York: Dial Press.

Myers, James, Jane Tsay, and Shiou-fen Su. 2011. Representation efficiency and transmission efficiency in sign and speech. In *Language and Cognition: Festschrift in honor of Professor James H.-Y. Tai on his 70th birthday*, ed. by Jung-hsing Chang. 171-199. Taipei: The Crane Publishing.

Rice, M. L., Smolik, F., Perpich, D., Thompson, T., Rytting. N., Blossom. M. Mean Length of Utterance Levels in 6-month Intervals for Children 3 to 9 Years with and without Language Impairments. *Journal of speech, language, and hearing research : JSLHR*. 2010; 53(2):333-349.

Tsay, Jane, James Myers, and James Tai. (in press) A Longitudinal Study of Communicative Efficiency in Taiwan Sign Language and Mandarin. In *Minpaku Sign Language Linguistics Series*. National Museum of Ethnology (Minpaku), Japan.

Tsou, C., and Hintat Cheung. 2007. Narrative storytelling of high-functioning children with Autism spectrum disorders. *Bulletin of Special Education* 32.3: 87-109. 鄒啟蓉,張顯達. 2007. 高功能自閉症兒童說故事能力與相關影響因素研究. 特殊教育研究學刊，32 (3): 87-109.

Zucchi, Alessandro. 2013. The language of propositions and events: Issues in the syntax and the semantics of nominalization. Vol. 51. Springer Science & Business Media.