

# CHINESE CAPTIONS AND FOREIGN LANGUAGE (MANDARIN) COMPREHENSION

Paul L. Markham, University of Kansas, pmarkham@ku.edu  
Sheree Willis, Confucius Institute, sawillis@ku.edu  
Connie Hsu, mars18\_jeremy@hotmail.com

---

## ABSTRACT

*The purpose of this study was to examine the influence of using Chinese captions vs. no captions on the Chinese (Mandarin) language comprehension of high intermediate university-level foreign language students. A total of 21 high intermediate (fifth semester) students participated in the study. The passage material consisted of a DVD episode (7 minutes) from a Chinese language interview of a pop music star. A total of 21 high intermediate (fifth semester) students participated in the study. The passage material consisted of a DVD episode (7 minutes) from a Chinese language interview of a pop music star. The students viewed only one of two treatment conditions: Chinese captions, or no captions. After viewing the video, the students demonstrated their comprehension via a ten minute English language written-recall protocol. The second comprehension task involved a 10-item English language written multiple-choice test based on the DVD passage material. The results revealed that the participants in the Chinese language captions-group significantly outperformed the no- captions group on both dependent measures despite the added complexity of processing Chinese characters at a fairly rapid rate.*

**Keywords:** foreign language, captions  
Chinese, comprehension

The purpose of this study was to examine the effects of using Chinese captions vs. no captions on the Chinese (Mandarin) language comprehension of intermediate university-level foreign language students. The theoretical framework was based on the work of Paivio [8] who developed a theory of dual coding connecting imagery (mental images or pictures) and language (oral or written), and the processing of these forms of information to facilitate comprehension. Evidence supporting this theory was provided by Paivio's own research [8] and many other studies such as Sadowski, Kealy, Goetz, and Paivio [10], and

## INTRODUCTION

Paivio and Lambert [9] who extended the inquiry concerning dual coding processes into the second language context.

A brief overview of the second language literature reveals supportive evidence for using captions. Garza [4] found that university ESL students learning English, and native English speakers learning Russian improved their reading and listening comprehension with caption availability.

Garza [1] also introduced the distinction between high audio/video correlation video material and low audio/video correlation material. A high audio/video correlation entails a strong connection between the pictorial content of the film and the audio portion of the film content. High audio/video correlation videos convey a great deal of content to the viewer without having to rely on audio input or captions for understanding. Not surprisingly, second language students benefit more from caption availability when viewing low audio/video correlation videos.

Continuing the favorable results, Neuman & Koskinen [7] determined that middle school-level ESL students dramatically enhanced their English language reading vocabulary knowledge after lengthy exposure to target language captions. Koskinen et al. [4] established that captioned videos considerably improved the incidental reading vocabulary knowledge of adult non-native English speaking correctional facility residents. Koolstra & Beentjes [3] found that primary school Dutch speaking students improved their reading vocabulary knowledge by watching several hours of Dutch-subtitled English language television programs in the home. Shea [11] observed that pace-controlled syntactic chunking of captioned video material significantly increased motivation, learning efficiency, and low-intermediate second language students' understanding of video discourse.

In more recent studies, Markham, Peter, & McCarthy [5] found support for the enhancement of Spanish as a foreign language university student listening comprehension. Markham & Peter [6] determined that Spanish as a foreign language student listening and reading comprehension were clearly improved by caption availability. Taylor [12] observed that third-year native-English speaking Spanish students found the combination of pictures, sound, and Spanish captions to be useful, but first-year students found the mixture of all three sources to be bewildering and unhelpful. Thus, considerable research supports the use of captions to increase both second language listening and reading comprehension. However, very little, if any, empirical research has examined the influence of using multilingual captioned video material on foreign language student comprehension of the Chinese (Mandarin) language.

### **RESEARCH QUESTIONS**

1. Does caption availability allow high intermediate level Chinese foreign language students to perform at a higher level on a written-recall test of a DVD video episode than students at the same level who do not have access to captions?
2. Does caption availability allow high intermediate level Chinese foreign language students to perform at a higher level on a multiple-choice test based on a DVD video episode than students at the same level who do not have access to captions?

### **METHODOLOGY**

A total of 21 high intermediate (fifth semester) students participated as intact groups in the study. The students were enrolled in two independent sections of a high intermediate level Chinese course. They were placed in this course either by progressing through the normal university course sequence or by taking a placement exam. The data were collected in fall semester (captions-group) and spring semester (no captions-group). This procedure was necessary because of the small number of participants in fall semester (10) and the likelihood that another small group of participants would be available in spring (11). The passage material consisted of a DVD episode (six minutes and thirty five seconds) from an interview with a Taiwanese pop music

star, Jay Chou. As the DVD episode consisted largely of two people talking in an ordinary looking office, the audio/video correlation of the episode was largely quite low. However, in one short segment Chou was depicted performing on stage. This short video segment clearly identified his occupation. The rest of the video was situated in the office.

The audio portion of the episode was always presented in Mandarin. The students were informed in advance that their participation in this research was totally unrelated to their course grades. The students participated in the investigation as intact classes during their normally scheduled class time. After viewing the video, the students demonstrated their comprehension via a ten minute written recall-protocol in which the students were told to write down as much as they could remember about the content of the video material in English. The ten minute timeframe was determined to be more than adequate based on previous pilot-testing. The written recall-protocol allowed the students ample opportunity to demonstrate what they understood and could remember about this particular passage. In order to facilitate scoring the English language written recall-protocol, the passage was translated into English by two raters. One of the raters was a native Chinese speaker with a high level of proficiency in English and the other was a native English speaker who was highly fluent in Chinese. The interrater reliability of their translations was very high at .94. The scoring procedure consisted of counting the number of idea units generated by the students using the parameters presented by Kintsch & Van Dijk [2]. Two raters scored the number of idea units generated by the students. A comparison of their ratings yielded a very satisfactory interrater reliability of .91.

After collecting the first protocol, the second research task involved a 10-item English language multiple-choice test based on the DVD passage material. Thus, the second dependent variable consisted of the multiple-choice test scores of the participants. The participants were given five minutes to complete the multiple-choice test. The multiple-choice test allowed the researchers to pinpoint specific information about the content of the passage without requiring students to write to demonstrate comprehension.

The language of the multiple choice items was drawn from the passage itself so as to be directly

reflective of the difficulty of the passage material itself.

## RESULTS

The participants completed a demographic section of the test instrument before participating in the main data collection phase. Regarding the participants' prior knowledge of Chou, only three participants in the captions group and four participants in the no-captions group reported having heard anything at all about Chou. No participants reported being *Very familiar* with Chou in either group. In addition, three participants in the captions-group and four participants in the no-captions group had lived in a Chinese speaking country from seven to twelve months, and three participants in both groups had lived in a Chinese speaking country from one to six months. Four participants in both groups reported living in a Chinese speaking country less than one month. No participant in either group had lived in a Chinese speaking country more than one year other than two native Chinese speakers. Two participants in the no-captions group identified themselves as being native Chinese speakers and their data were subsequently excluded from the study. Similarly, none of the participants had studied Chinese prior to their college experience except for the two native Chinese speakers. Given this demographic profile, neither group was determined to have an advantage after the decision was made to exclude the Chinese native speakers' data.

The main data collection results revealed that the participants in the Chinese language captions group significantly outperformed their counterparts in the no-captions group on both dependent measures (Tables 1 & 2). Regarding the first research question, the group means of the captions-group (12.60) were substantially higher than the means of the no-captions group (7.72). These differences proved to be statistically significant at the ( $p < .005$ ) level. This finding clearly indicated that the captions-group comprehended and recalled more of the passage material than their no-captions group counterparts.

Regarding the second research question, the captions-group once again surpassed the performance of the no-captions group on the multiple-choice test (Tables 1 & 2). The difference between the groups mean was equally as dramatic in that the captions-group had a mean

of 7.20 and the no-captions group mean was 4.54. This difference in the outcomes was somewhat unexpected as the written recall-protocol was a productive skill (writing) test, whereas the multiple-choice test was a recognition-based test. The results regarding the multiple-choice test were strongly significant favoring the captions-group ( $p < .000$ ).

The results for the two native Chinese speaking participants proved to be useful as a baseline for comparative purposes. They both recalled at least twice as many idea units as their counterparts in the no-captions group (14 and 15 respectively). Similarly only one of the two native Chinese speakers missed a single item on the multiple choice test. This finding verified that high target language proficiency makes a clear difference on performance in the caption-less context.

The passage material consisted of a DVD episode (six minutes and thirty five seconds) from an interview with a Taiwanese pop music star, Jay Chou. As the DVD episode consisted largely of two people talking in an ordinary looking office, the audio/video correlation of the episode was largely quite low. However, in one short segment Chou was depicted performing on stage. This short video segment clearly identified his occupation. The rest of the video was situated in the office.

## LIMITATIONS

Despite the fact that the data collection included two semesters of evidence, there were only 21 total participants in the study. The reason for this relatively small number of participants is because it is very difficult to find large samples of students who are studying Chinese at the high intermediate level. At this large Midwestern institution of higher education, it would not have been difficult to find a very sizeable number of foreign language students studying Spanish, for example, at the high intermediate level in any given semester except summer. However, not too many students choose to study Chinese in the first place and an even smaller number of students opt to study Chinese long enough to achieve a high intermediate level of proficiency in the language. Regardless of this factor, there were enough students in the captions-group (10) and the no-captions group (11) to justify using ANOVA to compare the group means using the F-test.

To make matters even more complicated, two participants in the captions-group had only learned simplified Chinese characters prior to transferring to the Midwestern institution where the study was conducted. This factor was important because the other students at the Midwestern institution had been learning only traditional Chinese characters prior to the study. Thus, the transfer students' lack of ability to read a few of the traditional Chinese characters used in the captions probably slightly lowered the group means of the captions-group. In addition, as mentioned earlier, two participants in the no-captions group self-reported that they were native Chinese speakers who had immigrated to this country either as a young teenager or as a high school graduate respectively. Though neither participant had access to Chinese captions in this study, they certainly possessed a higher level of oral Chinese language fluency than their counterparts in the no-captions group or the captions-group. Thus, the scores of the two native Chinese speaking individuals could not be used in this study.

## **DISCUSSION**

The consistent pattern of results favoring the captions-group is noteworthy because of the paucity of prior research regarding the effects of caption availability on high intermediate level foreign language students' comprehension of non-alphabetic languages such as Mandarin. Despite a restricted, emerging knowledge of Chinese characters, the participants in the captions-group consistently outperformed the no-captions group at a clearly significant level.

Limited Chinese character reading ability could easily have been a factor negating the captions-group's advantage as there were no pre-determined limits on the number of traditional Chinese characters used in the video and no limitations on what the two native Chinese speakers might want to talk about on the video. Based on feedback from the instructor, students at this level typically can read approximately 2000 Chinese characters, whereas a native Chinese speaker of the same age could be expected to be able to read approximately 3000 characters at a minimum. Regardless of the participants' limited ability to read Chinese captions, having access to captions clearly improved the participants' comprehension of the passage.

Another factor might have assisted the captions-group to the detriment of the no-captions group. Clearly, most intermediate level foreign language students find rapid, colloquial target language speech to be difficult to understand via listening comprehension only. This challenging aspect of the study was not unexpected. However, the interview format with Jay Chou could also have made the task slightly more difficult. Chou's native language is Taiwanese (Minnan) and the interview was conducted in Mandarin. Though Minnan speakers such as Chou learn Mandarin throughout their formal schooling, there were occasional minor pronunciation differences in Chou's version of Mandarin. The extent to which Chou's minor pronunciation differences actually influenced the outcome of the study is unknown.

Pedagogically speaking, captions can clearly aid the development of both listening comprehension and reading comprehension based on the extensive research cited earlier. However, at a given moment in time, the second language teacher who is primarily interested in listening development might want to begin by providing the target language captions in the beginning with a truly difficult passage for the student level, and then remove the captions during a subsequent viewing of the video if the pedagogical aim is primarily to develop listening comprehension. Students at this level can use captions as an aid to listening word identification in the early stages of the acquisition process as found by Markham, Peter & McCarthy [5]. The flexibility of caption availability is an obvious strength of this technology that only enhances the value of captions as an aid for second language listening and/or reading development.

Paivio's dual coding theory [8] seems to be strongly supported by this research evidence in that the pictures, captions, and the audio seemed to be simultaneously processed by the students without great difficulty. Whether the students were actually simultaneously processing the different sources of information or rapidly changing back and forth between information sources is not something that was investigated in this study. For pedagogical purposes, perhaps it actually doesn't matter whether simultaneous processing or rapid task-switching is taking place as long as the comprehension benefits for language students are clearly manifested.

The educational importance of this study is that intermediate, university level foreign language students clearly demonstrated comprehension

gains based on caption availability with a non-alphabetic, logographic language. Despite the added difficulty of having to decode and comprehend Chinese characters at a fairly rapid rate, the captions were clearly helpful to comprehension based on this research. This outcome confirms and extends earlier findings with foreign language students of alphabetic languages. Clearly caption availability is a very helpful factor to high intermediate foreign language students of a non-alphabetic language as well.

#### REFERENCES

1. Garza, T. (1991). Evaluating the use of captioned video material in advanced foreign language learning. *Foreign Language Annals, 24*, 239-258.
2. Kintsch, W. & Van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review, 38*, 363-394.
3. Koolstra, C. & Beentjes, W. Children's vocabulary acquisition in a foreign language through watching subtitled television at home. *Educational, Technology, Research, and Development, 47*, 51-60.
4. Koskinen, P., Knable, J., Markham, P., Jensema, C. & Kane, K. Captioned television and the vocabulary acquisition of adult second language correctional facility residents. *Journal of Educational Technology Systems, 24*, 359-373.
5. Markham, P., Peter, L. & McCarthy, T. (2001). The effects of native language vs. target language captions on foreign language student dvd video comprehension. *Foreign Language Annals, 34*, 439-445.
6. Markham, P. & Peter, L. (2003). The influence of english language and spanish language captions on foreign language listening/reading comprehension. *Journal of Educational Technology Systems, 31*, 331-341.
7. Neuman, S. & Koskinen, P. Captioned television as comprehensible input: Effects of incidental word learning in context for language minority students. *Reading Research Quarterly, 27*, 95-106.
8. Paivio, A. (1971). Imagery and deep structure in the recall of english nominalizations. *Journal of Verbal Learning and Verbal Behavior, 10*, 1-12.
9. Paivio, A. & Lambert, W. (1981). Dual coding and bilingual memory. *Journal of Verbal Learning and Verbal Behavior, 20*, 532-539.
10. Sadowski, M., Kealy, W., Goetz, W. & Paivio, A. (1991). A critique of schema theory and a dual coding alternative. *Reading Research Quarterly, 26*, 463-484.
11. Shea, P. (2000). Leveling the playing field: A study of captioned interactive video for second language learning. *Journal of Educational Computing Research, 22*, 243-263.
12. Taylor, G. (2005). Perceived processing Strategies of students watching captioned video. *Foreign Language Annals, 38*, 422-427.

**Table 1**

**ANOVA**

		Sum of Squares	df	Mean Square	F	Sig.
Propositions	Between Groups	124.371	1	124.371	10.248	.005
	Within Groups	230.582	19	12.136		
	Total	354.952	20			
MCTest	Between Groups	36.911	1	36.911	20.430	.000
	Within Groups	34.327	19	1.807		
	Total	71.238	20			

**Table 2**

**Descriptive Statistics**

Captions		N	Minimum	Maximum	Mean	Std. Deviation
Captions	Propositions	10	6.00	18.00	12.6000	4.32563
	MCTest	10	4.00	9.00	7.2000	1.61933
	Valid N (listwise)	10				
No Captions	Propositions	11	3.00	11.00	7.7273	2.49363
	MCTest	11	3.00	6.00	4.5455	1.03573
	Valid N (listwise)	11				