# Detection of Words in Fluent Chinese by English-acquiring and Chinese-acquiring Infants

Jane Tsay* and Peter Jusczyk**
*National Chung Cheng University and **Johns Hopkins University

To address the question of whether infants can detect familiar words when they are embedded in a sentential context, Jusczyk and Aslin (1995) tested 7.5-month-old English-acquiring infants using the Headturn Preference Procedure (Kemler Nelson, Jusczyk, Mandel, Myers, Turk, and Gerken, 1995) (HPP). Four monosyllabic words, "feet," "cup," "dog," and "bike," were used as test words. Half of the infants were familiarized with "cup" and "dog," and the other half with "feet" and "bike." Four passages, consisting of six sentences, were constructed using these four test words. Test words appeared in a different position in each of the six sentences. After infants had accumulated 30 seconds of listening time to two test words, the infants listened to the four passages. Since each group of infants was familiarized with only two test words, the two passages that contained test words that the infants were not familiarized with acted as controls. They found that infants listened significantly longer to passages containing the familiar test words.

A natural follow-up question is: Do infants need to know the language to recognize the familiar word patterns in continuous speech? Specifically, after being familiarized with words in a foreign language, will they detect them in the fluent speech of that language? Somewhat similar experiments have been run successfully using artificial languages (see e.g. Goodsitt, Morgan, & Kuhl, 1993, and Morgan, 1994, for infant speech perception studies; cf. Dahan & Brent, 1999, for a study using adult subjects to model possible infant behavior). Our study was the first to apply the paradigm established in Jusczyk and Aslin (1995) to address this question using stimuli from a genuine language that the infant subjects had never been exposed to.

We expected that Mandarin Chinese (called "Chinese" in the remainder of this report) would provide a particularly interesting language to test with English-acquiring infants because its phonological characteristics are typologically quite different from English. First of all, although it is not true that Chinese is a truly isolating language (most Chinese words contain more than one morpheme; Zhou & Marslen-Wilson, 1994), the most common words are monosyllabic, as are virtually all morphemes. Second, each syllable carries a lexical tone. Thus stress is not useful for identifying word boundaries. Third, Chinese syllable structure is much simpler than that of English, which means that many morphemes are homophonous with others. Fourth, Chinese does not allow consonant clusters and velar and alveolar nasals are the only consonants that can appear in the coda position, which may make it harder to segment words acoustically. Nevertheless, if infants rely on general algorithms and do not rely on these cues (e.g. Brent and Cartwright, 1996), we might expect that English-acquiring infants will have no serious problem detecting familiarized Chinese words in fluent Chinese speech.

## Experiment 1

The HPP was used to test infants' recognition of Mandarin words in fluent speech.

## *Methods*

*Subjects*. Twenty-four American infants (10 females, 14 males), approximately 7.5 months of age, were tested. The average age of the infants was 33 weeks, 6 days (range: 30 weeks, 3 days to 41 weeks, 2 days). Three additional infants were tested but not included because of being too restless during the test.

*Stimuli*. A female native speaker of Mandarin Chinese in her 30's from Taiwan recorded the stimuli. Four monosyllabic content words with clear onsets were chosen as target words. They were *tou[LH]* "head", *dan[HL]* "egg", *tian[H]* "sky, day", and *bei[H]* "cup".[1] Four passages, each containing six sentences using one of the target words, were prepared. The target word occurred at the beginning, middle, and end positions of sentences two times each. The four different 6-sentence passages (see Table 1) were recorded along with another four filler passages to make sure that the talker did not emphasize the target words. The talker was reminded to read the passages in a lively voice, as if she were talking to a small child.

Table 1 Four test passages

Target word: tou "head"
Meige ren de ***tou*** dou bu yiyang. "Everybody's head is different."
***Tou*** keyi cangzai maozi dixia. "A head can be hidden inside a hat."
Wo de hao pengyou shi Da-***tou***. "My best friend is Big Head."
Shizi de ***tou*** hen qiguai. "The head of a lion is very weird."
***Tou*** buyao shenchu che wai. "Don't stick your head outside of the car."
Linju de nanhai jiao Xiao-***tou***. "The neighbor's boy is called Small Head."

Target word: dan "egg"
***Dan*** fangzai bingxiang li. "Eggs are put in the refrigerator."
Wo mai da ***dan***-gao gei ni. "I bought a big cake (lit. egg-cake) for you."
Muji xiale haoduo ***dan***. "The hen laid many eggs."
Youxie ren bu ai chi ***dan***. "Some people don't like eating eggs."
Xiaohua ba ***dan*** dapole. "Xiaohua broke the eggs."
***Dan*** gundao zhuo-zi dixia qu le. "Eggs rolled down under the table."

Target word: tian "sky, day"
***Tian*** shang you ji duo baiyun. "There are some clouds in the sky."
Ta tangzai dishang kan lan-***tian***. "He is lying on the ground looking at the sky."
Women xingqi-***tian*** qu kao rou. "We had a barbecue on Sunday (lit. week-day)."
***Tian*** kong feichang de yin'an. "The sky is very dark."
Dajia dou xihuan qing ***tian***. "Everybody likes clear days."
Xiayu ***tian*** bu neng qu dongwu yuan. "One can't go to the zoo in raining days."

Target word: bei "cup"
Milaoshu zai zhao cha-***bei***. "Mickey Mouse is looking for a teacup."
Zhe ge xiao ***bei***-zi hen piaoliang. "This little cup is very pretty."

---

[1] Mandarin has four lexical tones: high level H, low dipping L, rising LH, and falling HL.). There is a tone alternation involving L-tone words in a certain phonological context. To avoid complications, L-tone words were not used as target words.

***Bei***-zi li you guozhi. "There is juice in the cup."
***Bei*** di zuozhe yi zhi qingwa. "There is a frog sitting in the bottom of the cup."
Ta ba yi da ***bei*** niunai heguangle. "He drank up a big cup of milk."
Wo song ni yi ge boli ***bei***. "I'll give you a glass cup."

The recordings were made in a sound attenuated room with a Shure microphone. The critical four passages were digitized via a 12-bit analog-to-digital converter at a 10 kHz sampling rate and stored on a VAX Station Model 3176 computer. The average duration of the passages was 17.37 seconds (ranging from 16.68s for the *dan* "egg" passage to 18.06s for the *tou* "head" passage). The average amplitude of the passages was 75.64 dB (ranging from 75.04 dB to 76.38 dB).

The talker was then asked to record the four target words in isolation for the familiarization phase of the experiment using the same equipment as above. For each target word, the talker was asked to say the word in a lively voice 15 times in a row with some variation each time. These lists were digitized the same way as the passages. The average duration of the word lists was 26.29 seconds (ranging from 25.29s to 26.60s). The average amplitude was 73.19 dB (ranging from 72.36 dB to 73.76 dB). Digitized versions of the passages and word lists were transferred to a PDP 11/73 computer for playback during the experiment.

*Design.* Half of the infants heard the word lists with the words *tou* "head" and *bei* "cup" during the familiarization phase, and the other half heard word lists with the words *dan* "egg" and *tian* "sky, day". During the test phase, all infants heard four blocks of the same four passages. Each block contained a different random ordering of the test passages.

*Apparatus.* A PDP 11/73 controlled the presentation of the stimuli and recorded the observer's coding of the infant's responses. A 12-bit D/A converter was used to recreate the audio signal. The resulting audio signal was then filtered at 4.8 kHz, amplified, and played from two loudspeakers mounted on the walls of the testing booth.

The experiment took place in a three-sided test booth constructed out of 4 ft x 6ft pegboard panels on three sides and open at the back. An experimenter was located behind the front wall of the booth, and looked through the existing holes in the pegboard to observe the infant's headturns. The remainder of the pegboard panels were backed with white cardboard to guard against the possibility that the infant might respond to movements behind the panels. The test booth had a red light and a loudspeaker mounted at eye level on each of the sides. There was a green light in the center of the front panel, and directly below the center light there was a 5-cm hole for the lens of a video camera. The video camera was used to provide a permanent record of each session. A white curtain was suspended from the ceiling and prevented the infant from seeing over the top of the booth. A computer terminal and 6-button response box were located behind the front panel of the booth. The response box, which was connected to the computer, was equipped with buttons that started and stopped the flashing of center and side lights. It also recorded the direction and duration of headturns and terminated a trial when the infant looked away for more than 2 second.

*Procedure*. Test sessions used the Headturn Preference Paradigm (Kemler Nelson, Jusczyk, Mandel, Myers, Turk, & Gerken, 1995), and were modeled after prior work by Jusczyk and Aslin (1995). During the test session, infants sat on their caregiver's lap in the test booth. During the familiarization phase, infants heard two of the target words on alternating trials until they accumulated at least 30 seconds of listening time to each word. Listening time was assessed by the amount of time the infant spent looking at the "source" of the sound (the flashing light on one of the side walls). The loudspeaker used for each word was randomly varied from trial to trial. A trial was terminated when the infant looked away for more than 2 seconds or when the end of the word list was reached. Information about the direction and duration of headturns and the total trial duration were stored in a data file on the computer. The test phase began immediately after the familiarization criterion was reached. The test trials were blocked in four groups so that each test passage occurred once in a given block, although the order of the four passages within each block was randomized. Each infant was tested on a total of 4 blocks, for a total of 16 test trials.

Both familiarization and test trials began by blinking the green light in the center of the front panel. Once the infant had oriented in that direction, the light was turned off and one of the two red lights began to flash. Once the infant had oriented towards that light, the stimulus for that trial (regardless of whether it was a familiarization stimulus or a test stimulus) began to play from the loudspeaker on the same side. The stimulus continued to play until its completion, or until the infant had looked away for 2 consecutive seconds, whichever came first. Any time the infant spent looking away (whether it was 2 seconds or less) was not included when measuring the total listening time. The red light continued to flash for the duration of the entire trial.

The experimenter behind the center panel pressed a button on the response box whenever the infant looked at or away from the flashing light. The experimenter was not told ahead of time which items were used in the familiarization phase, and both the experimenter and the caregiver listened to masking music over headphones.

### *Results and discussion*

Mean orientation times were calculated for each of the 24 subjects across the four blocks of trials. The mean length of orientation across subjects was 7.32 sec (s.d. = 2.39 sec) for the Familiarized passages and 7.43 sec (s.d. = 1.93 sec) for the Unfamiliarized passages, a nonsignficant difference ($t(23) = -0.29$, $p > 0.7$).

The null results of this experiment could have any of a number of causes. On the one hand, it is possible that word segmentation relies very heavily on learning to detect cues in the ambient language, so general algorithms based on distributions of word-sized units do not work. The subjects in Experiment 1 had no chance to pick up the relevant cues for Chinese, and so were not able to detect the words embedded in fluent sentences. On the other hand, it could be that this kind of task does not work for Chinese materials in general. That is, it could be that nature of Chinese words and Chinese phonology make it too difficult for infants at this age to detect them in fluent speech. To help distinguish the two very different sorts of interpretations, we began Experiment 2, which used the same materials but presented them to infants acquiring Chinese in homes where Chinese was the dominant language.

## Experiment 2

### *Methods*

*Subjects*. Due to difficulties in finding appropriate subjects in the area of the laboratory (Amherst, NY), only four Chinese infants, of approximately 7.5-months of age, provided usable results in this experiment.

*Stimuli*, *Design, Apparatus*, *Procedure*. The stimuli, design, apparatus, and procedure were identical to Experiment 1.

### *Results and discussion*

Mean orientation times were calculated for each of the 4 subjects across the four blocks of trials. The mean length of orientation across subjects was 8.21 sec (s.d. = 3.27 sec) for the Familiarized passages and 9.15 sec (s.d. = 3.07 sec) for the Unfamiliarized passages. Because all four subjects oriented longer during the Unfamiliarized passages, this difference was significant by two-tailed paired *t* test ($t(3) = -7.27, p < 0.01$) and nearly so by Wilcoxon signed rank test ($p = 0.068$), though of course with this small sample size the results must be taken with a great deal of caution. Indeed, the first four subjects run in Experiment 1 also showed longer orientation times during the Unfamiliarized passages, but when the sample size was increased to 24, this illusory effect entirely disappeared. Unfortunately only four subjects were run in Experiment 2 by the time the experiment was terminated.

## General Discussion

Why could English-acquiring infants not detect familiarized Chinese words in fluent speech? One possibility may lie in the phonological differences between the two languages. Because of the relatively simple syllable structure of Chinese, there are not as many syllable types in Chinese as there are in English. Therefore, Chinese syllables might not sound distinct enough from each other to English-acquiring infants.

Moreover, obstruents have advantages over sonorants in acoustically demarcating word boundaries. Jusczyk & Aslin (1995) used CVC words, where both C's were obstruents, and all but one of the C's were stops. By contrast, Mandarin Chinese only allows velar and alveolar nasals to appear in coda position, which may have made the Chinese words blur into their neighbors more. Partly due to the simpler syllable structure of Chinese compared with English, Cheng (2002) found in a simulation study that an algorithm relying solely on distributional information, which worked relatively successfully with English (Brent & Cartwright, 1996) did not help much in segmenting Mandarin Chinese.

Prosodic differences are also crucial. Because pitch is not used to make lexical distinctions in English, the number of different syllable types in Chinese may perhaps seem even smaller to English-acquiring infants, making the task of picking out just the right one in fluent speech even harder.

As to the Chinese-acquiring infants, it has been argued that most words in Chinese are disyllabic due to a disyllabic foot bias (e.g., Duanmu 2002). The large number of homophones in Chinese may further encourage this bias. Thus even Chinese infants may have be conditioned to expect words to be disyllabic, making it hard to detect the monosyllabic words used in both experiments. Of course, with data currently from only four Chinese infants, who were probably also exposed to a certain amount of English as well given that they were growing up in the US, it is certainly too early to draw conclusions about the natural patterns of Chinese infant speech perception. Hopefully, further research can be conducted along this line after infant speech labs have been established in Chinese-speaking countries.

Acknowledgements

This study was conducted with support from NIDCD Training Grant (5T32DC00036-02) to the second author. The first author was supported as a postdoctoral trainee by the same grant. We thank Ann Marie Jusczyk for her very generous and professional help with the experiments. The first author may be contacted at Graduate Institute of Linguistics, National Chung Cheng University, Min-hsiung, Chia-yi, 621, Taiwan, or by email at Lngtsay@ccu.edu.tw.

References

Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition, 61*, 93-125.

Cheng, T. (2002). *Testing a model of word segmentation in the acquisition of Mandarin*. Unpublished MA thesis, National Chung Cheng University. Chia-Yi, Taiwan.

Dahan, D., & Brent, M. R. (1999). On the discovery of novel wordlike units from utterances: An artificial-language study with implications for native-language acquisition. *Journal of Experimental Psychology: General, 128* (2), 165-185.

Duanmu, S. (2000). *The phonology of Standard Chinese*. Oxford: Oxford University Press.

Goodsitt, J. V., Morgan, J. L., & Kuhl, P. K. (1993). Perceptual strategies in prelingual speech segmentation. *Journal of Child Language, 20,* 229-252.

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology, 29,* 1-23.

Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. A. (1995). The Headturn Preference Procedure for testing auditory perception. *Infant Behavior and Development, 18*, 111-116.

Morgan, J. L. (1994). Converging measures of speech segmentation in prelingual infants. *Infant Behavior and Development, 17*, 387-400.

Zhou X., & Marslen-Wilson, W. (1994). Words, morphemes and syllables in the Chinese mental lexicon. *Language and Cognitive Processes, 9* (3), 393-422.