

# Predictors of timed picture naming in Chinese

**BRENDAN STUART WEEKES**

*University of Sussex, Brighton, England*

**HUA SHU, MEILING HAO, AND YOUYI LIU**

*Beijing Normal University, Beijing, China*

AND

**LI HAI TAN**

*University of Hong Kong, Hong Kong, China*

We report normative data collected from Mainland Chinese speakers for 232 objects taken from Snodgrass and Vanderwart (1980). These data include adult ratings of concept familiarity, age of acquisition (AoA), printed-word frequency, and word length (in syllables), as well as measures of rated visual complexity, image agreement, and name agreement. We then examined timed picture naming of these objects with native Chinese speakers in Beijing in two experiments using line drawings and colored pictures. In both experiments, the variables name agreement, rated concept familiarity, and AoA made significant independent contributions to naming latency in multiple regression analyses. We observed a correlation of  $r = .85$  between naming latency with line drawings and colored pictures and a reduced effect of image agreement on naming when colored pictures were presented. We discuss the implications of our findings for the study of lexical processing in Chinese. Normative data for 232 Chinese nouns may be downloaded from [www.psychonomic.org/archive](http://www.psychonomic.org/archive).

Many studies have analyzed the properties of objects that influence picture naming speed. These studies are important for developing cognitive models of object recognition and naming and for the study of lexical processing in the brain (Bates et al., 2003; Székely et al., 2003). Up until the late 1990s, research on the key factors affecting picture naming speed was limited to that done in English. However, there has been a steady increase in studies investigating timed picture naming in other languages including Dutch (Jescheniak & Levelt, 1994; Severens, Van Lommel, Ratinckx, & Hartsuiker, 2005), French (Alario et al., 2004; Bonin, Chalard, Méot, & Fayol, 2002; Bonin, Peerean, Malardier, Méot, & Chalard, 2003; Chalard, Bonin, Méot, Boyer, & Fayol, 2003; Kremin, Hamerel, Dordain, De Wilde, & Perrier, 2000; Schwitler, Boyer, Méot, Bonin, & Laganaro, 2004), Icelandic (Pind & Tryggvadottir, 2002), Italian (Dell'Acqua, Lotto, & Job, 2000), Japanese (Nishimoto, Miyawaki, Ueda, Une, & Takahashi, 2005), and Spanish (Cuetos, Ellis, & Alvarez, 1999; Sanfeliu & Fernandez, 1996). All of these studies investigate naming of objects using line drawings first published by Snodgrass and Vanderwart (1980) and locally collected normative data. According to a review by Alario et al., the key variables that predict picture naming latency across languages are name agreement and rated age of acquisition (AoA). Bates and colleagues compared languages in a study of picture naming in Bulgarian, En-

glish, German, Hungarian, Italian, Spanish, and Mandarin (Putonghua). Their data confirmed the importance of name agreement but also revealed a robust effect of printed-word frequency across languages. Bates et al. did not, however, consider the full range of variables for each of the languages examined; for example, they did not consider the effect of rated AoA on picture naming in Chinese. The primary aim of the present study was to investigate the impact of key variables—including AoA—on picture naming in Mainland Chinese speakers.

Name agreement has a significant impact on timed picture naming that is independent of the effects of correlated variables such as word frequency, familiarity, and AoA in all languages. Name agreement can be measured in several ways, but the most common way is to determine the proportion of speakers who assign the target name to a pictured object. Another method is to measure the number of names given to a picture across participants. On both measures, pictures of objects that elicit different names—for example, *gun/pistol/revolver*—have lower name agreement than do those that elicit a single name. If an object has more than one alternative name and participants do not all give the same name to a given picture, retrieval of any one of those names will be slowed. Pictures with a single dominant response are named more quickly and more accurately than pictures with alternative responses. This robust effect is possibly due to competition for tar-

get names competing with alternative names for selection during timed picture naming (Barry, Morrison, & Ellis, 1997; Bates et al., 2003; Snodgrass & Yuditsky, 1996; Vitkovitch & Tyrrell, 1995). According to Levelt (2002), effects of name agreement might also have an impact at the locus of structural representations if certain pictures are more difficult to identify than others.

Many studies also find that frequency of exposure to the names of objects has a reliable effect on picture naming (Alario et al., 2004; Barry et al., 1997; Bates et al., 2003; Cuetos et al., 1999; Ellis & Morrison, 1998; Snodgrass & Yuditsky, 1996; Székely et al., 2004; though see Bonin et al., 2002; Dell'Acqua et al., 2000; Pind & Tryggvadottir, 2002). Word frequency is typically estimated using printed frequency counts based on text corpora, although subjective ratings of word familiarity can be used to estimate spoken word frequency (Gernsbacher, 1984). Across studies, the higher the printed-word frequency and the greater the familiarity of an object name, the sooner the picture is named. It is important to note here that printed-word frequency and rated familiarity can have independent effects on timed picture naming (Alario et al., 2004).

The other reliable predictor of timed picture naming speed is AoA. This is an estimate of the age at which the name of an object is learned, measured with subjective or objective values. Many studies report effects of subjective AoA on picture naming using retrospective adult rated estimates (Carroll & White, 1973; Gilhooly & Gilhooly, 1979; Morrison, Ellis, & Quinlan, 1992) and some studies also report effects of objective AoA—based on children's naming performance—on picture naming (Pind & Tryggvadottir, 2002). Morrison, Chappell, and Ellis (1997) observed a high correlation between rated and objective measures of AoA, and Ellis and Morrison (1998) found that both measures predict adult naming in their own study and in data published by Barry et al. (1997) and Snodgrass and Yuditsky (1996).

The question of whether AoA has an effect on timed picture naming independent of frequency and other variables is more controversial. Studies using multiple regression report that subjective and objective measures of AoA have a significant effect on picture naming independent of printed-word frequency and rated familiarity in English (Barry et al., 1997; Bates et al., 2003; Ellis & Morrison, 1998; Snodgrass & Yuditsky, 1996), French (Alario et al., 2004; Bonin et al., 2002; Bonin et al., 2003), Icelandic (Pind & Tryggvadottir, 2002), Italian (Dell'Acqua et al., 2000), and Spanish (Cuetos et al., 1999; see also Pérez & Navalón, 2005). Zevin and Seidenberg (2004) also raised concerns about use of objective measures of AoA to predict naming times. These estimates represent the average age at which children use a word correctly in picture naming tasks (e.g., Morrison & Ellis, 2000). Thus, objective AoA is a measure of performance, unlike rated AoA, which is based on a subjective evaluation. According to Zevin and Seidenberg (2004), the problem with using objective AoA to predict timed picture naming latency is that correlating one performance measure with another does not reveal why some names take longer to name than others. Therefore, we did not measure objective AoA here.

We do not know whether AoA has an independent effect on picture naming in Chinese. Bates et al. (2003) examined picture naming performance of Mandarin speakers in Taiwan using 520 objects and found that goodness of depiction (GoD), printed-word frequency, and length in syllables predict naming latency. GoD refers to subjective ratings of 520 object pictures obtained from a group of 50 U.S. college students, who were asked to rate how well a picture fit with its dominant name according to a 5-point scale. Word frequency values were taken from the Chinese Knowledge Information Group (1997), a Taiwanese source based on the frequency of printed complex characters in Taiwanese texts. The results from Bates et al. show clearly that objects with less common names and poor quality representations of objects take longer to name in Chinese. However, because printed-word frequency is likely to be correlated with AoA, it is not clear whether the effects of word frequency reported in the Bates et al. study are the results of differences in AoA or rated familiarity, which is correlated with both frequency and AoA.

We report below the timed picture naming performance of Mainland Putonghua speakers from Beijing who named pictures depicting objects from the Snodgrass and Vanderwart (1980) set. Shu, Cheng, and Zhang (1989) reported normative data on rated concept familiarity for these objects, as well as measures of name agreement, image agreement, visual complexity, and word length. Their methods and data (published in Chinese) are reproduced in this report. In addition, we collected data on the rated AoA of object names, as well as fresh norms for rated familiarity of objects, since older normative data may misrepresent the current familiarity of objects (e.g., *roller skate*), given the rapid economic growth in China over the past decade. To preface the results: We found effects of name agreement, rated familiarity, and AoA on timed picture naming, but, unlike Bates et al. (2003), we found no significant effect of printed-word frequency when the effects of AoA, name agreement, and familiarity were removed, using multiple regression.

## PREPARATORY STUDY

The stimuli for the present study were 232 black-and-white drawings of objects taken from Snodgrass and Vanderwart (1980). (The names of objects are reported in the Appendix, available in the Psychonomic Society online archive at [www.psychonomic.org/archive](http://www.psychonomic.org/archive).) For each object, ratings of name agreement, concept familiarity, visual complexity, and image agreement from Shu et al. (1989) are also reported in the Appendix (see archive). Shu et al. used identical procedures to Snodgrass and Vanderwart to obtain ratings from undergraduate Putonghua speakers from Beijing Normal University (BNU). Participants were shown individual pictures on a slide projector and were asked to (1) write down the name of each picture given three choices and respond with the name of the object shown, or the words *don't know the object* or the words *don't know the name of the object*; (2) make ratings of the familiarity of the object, defined as the degree to which they come in contact with or think about the

concept depicted by the picture, on a 5-point rating scale in which 1 indicates *very unfamiliar* and 5 indicates *very familiar*; (3) make ratings of the visual complexity of each picture, using a 5-point scale in which 1 indicates *very simple* and 5 indicates *very complex*; and (4) make ratings of image agreement—i.e., how closely each picture resembles a participant’s mental image of the object—using a 5-point scale in which 1 indicates *very good match* and 5 indicates *very poor match*. Shu et al. also reported three measures of name agreement: (1) a name uncertainty variable ( $H$  value), (2) a percentage (%) name agreement, and (3) a concept agreement (CA).  $H$  reflects the number of alternative names given by participants and was computed using the following formula:

$$H = \sum_{i=1}^K P_i \log_2(1 / P_i),$$

where  $K$  is the number of names participants gave to a picture and  $P_i$  is the proportion of each name out of all possible names.  $H$  thus represents uncertainty about a picture name; this is reflected in participant responses. According to Snodgrass and Vanderwart,  $H$  captures more information about the distribution of names than the percentage name agreement. For example, if two concepts are both given their dominant names by 60% of participants, but one concept is given a single other name and the second is given four other names, both concepts will have equal percentage name agreement, but the first one will have a lower  $H$  value. Shu et al. used a strict criterion for counting different instances of names when computing  $H$  values. In many cases, the name given by a participant was similar but not identical to an established name category. Thus, an item was given the perfect score of  $H = 0$  only if a picture elicited the same name from every participant who was able to provide a name for the item. Increasing  $H$  values reflect decreasing certainty about a picture name and a decreasing number of participants who produce the same name. *Percentage name agreement* is the proportion of participants who produce the most common name when considering all possible names. *Concept agreement* is the proportion of participants who produce a name that reflects the concept instead of the target name. As the ratings for concept agreement reflect the actual concept, values for concept agreement are always more than or equal to the value percentage name agreement. Shu et al. also measured the length of an object name using the number of syllables in the most common (target) name of the object.

Shu et al. (1989) did not report printed-word frequency values for items in the Appendix (see archive) because of difficulty estimating printed-word frequency in Chinese. Morphemes are depicted with a character in Chinese; but most words are bisyllabic, and hence represent concepts made up of more than one morpheme. It is now possible to isolate the printed frequency of characters in a spoken name when they form the target morpheme, using data taken from a corpus of written texts comprising 1.22 million words published by Luo and Yuan (1996). We therefore calculated the printed-word frequency values for each object reported by Shu et al. where available, using this corpus. However,

this left us only 144 items (see archive). It is important to note that our measure of word frequency is a crude estimate of the cumulative frequency of exposure to spoken words in Putonghua; rated familiarity may be a better index.

Shu et al. (1989) also did not report AoA values for object names. We therefore collected AoA ratings from 20 Beijing Normal University students who were paid 10 yuan each for participation. We followed established procedures reported in studies of AoA ratings from speakers of English (Barry et al., 1997), French (Alario et al., 2004), and Spanish (Cuetos et al., 1999). Booklets with between 8 and 10 stimuli per page were prepared for all 232 items in the Appendix (see archive). All items were presented in a random order, and each participant received a different version of the pages in the booklet. Participants were given the following printed instructions: “Estimate the age at which you believe that you first learned the name of the object. Use the following scale: 1 - under 2 years old; 2 - 2 years old; 3 - 3 years old; and so on, up to 11 - 11 years old or older.” There was a rating scale placed at the top of each page to remind participants how they should score each item. All 20 respondents completed the AoA ratings for every item in the booklets.<sup>1</sup> Summary statistics for all predictor variables are given in Table 1.

## EXPERIMENT 1

### Method

**Participants and Procedure.** Participants were 30 undergraduate students at Beijing Normal University. All were Putonghua speakers and each was paid 10 yuan for participation. No participant was involved in the rating exercises described above. Our procedure was identical to that of Cuetos et al. (1999). The experiment was controlled using the DMDX experimental program (Forster & Forster, 2003). Pictures were presented at the center of a computer monitor. Each participant was seated at the front of the monitor, with a microphone held in place by headphones connected to a voice key in front of the mouth. When the picture appeared, a chronometer in the computer started, and when the participant responded, the chronometer stopped and the computer automatically recorded the time elapsed between stimulus presentation and response. The picture disappeared from the monitor as soon as a response was recorded, and in its place a cross serving as a fixation point appeared for 500 msec in the center of the monitor. After this, the fixation point disappeared and a new picture was presented on the monitor by the program. The maximum time that a picture stayed on the monitor was 2,000 msec. If after this time no reply had been given, an error was recorded and

**Table 1**  
Summary Statistics for Predictor Variables

Variable	$M$	$SD$	Min.	Max.	Skew
Name agreement ( $H$ )	0.97	0.50	0.00	1.95	-0.21
Percentage name agreement	0.65	0.24	0.04	1.00	-0.32
Concept agreement	0.82	0.21	0.04	1.75	-0.84
Visual complexity	2.97	0.81	1.19	4.76	-0.03
Image agreement	3.49	0.64	1.33	4.55	-0.72
Rated familiarity	3.83	0.80	1.87	5.00	-0.33
AoA	4.93	1.58	1.99	9.18	0.47
Log (1 + frequency)	1.52	0.73	0.51	3.54	0.54
Syllable number	1.98	0.55	1.00	4.00	0.29

Note—This summary is based on the 232 items reported in the Appendix (see archive). AoA, age of acquisition.

the picture replaced by the fixation point. For each participant the sequence of events was as follows: The voice key was adjusted to suit the voice sensitivity of each participant before the experiment commenced, and 12 practice items were then presented. A complete explanation of the experiment was given before the practice trial with instructions consisting of a request not to say anything except to speak the name of the object depicted in the picture, and a warning not to move heads, cough, or breathe loudly. We only presented items if they had a word-frequency value in Luo and Yuan (1996). The first set of stimuli consisted of 12 practice trials followed by 144 experimental trials. Stimulus presentation was randomized for each participant. All responses were recorded and transcribed for later analysis of responses.

## Results

Before calculating the means of the reaction times (RTs), we eliminated items for which the participants had given an incorrect response, or for which the voice key had been triggered by coughing or head movement. There were 4.9% spoiled responses and 2.1% voice key failures. Following the procedure recommended by Snodgrass and Yuditsky (1996), scores more than two standard deviations from the mean were removed in order to eliminate the influence of outliers. This resulted in removal of 6.9% of responses. The mean of all naming times was then calculated for each item. Mean naming latencies for each item are shown in the Appendix (see archive), along with predictor values. The global mean naming latency for all items ( $n = 144$ ) was 1,025 msec, which is compatible with results from Bates et al. (2003), who reported a mean picture naming latency of 1,200 msec for Mandarin Chinese speakers in Taiwan. This is slower than the RTs reported for British participants by Barry et al. (1997) (748 msec) and Ellis and Morrison (1998) (794 msec); for U.S. participants by Snodgrass and Yuditsky (1996) (791 msec); and for Spanish participants by Cuetos et al. (1999) (829 msec). Longer naming latencies may be due to the relatively large number of polysyllabic words in Putonghua compared with those in other languages (Bates et al., 2003).

Significant correlations between predictor variables and naming RTs are shown in Table 2.

The variables with the highest simple correlations with naming RT are name agreement ( $H$ ), followed by familiarity and AoA. Table 1 also reveals significant correlations between predictor variables. For example, rated familiarity correlated with AoA (familiar pictures have

names acquired at an early age) and frequency (familiar pictures have common names).

**Stepwise multiple regression.** To assess the unique impact of each predictor variable, a stepwise multiple regression was performed using all of the predictor variables shown in Table 1. Percentage name agreement was the first variable entered ( $b = -.38, t = -4.6, p < .001$ ), followed by familiarity ( $b = -.30, t = -3.6, p < .001$ ), AoA ( $b = .19, t = 2.3, p < .05$ ), and image agreement ( $b = -.14, t = 1.95, p = .05$ ). Adjusted  $R^2$  for these factors was equal to 0.35, which was significant [ $F(8,135) = 19.50, p < .001$ ]. None of the other variables reported in Table 1 were significant. When rated AoA was entered last in a hierarchical regression equation, it explained a significant 2.5% of additional variance ( $p < .05$ ). However, in an equivalent analysis with printed-word frequency, this variable explained less than 0.1% of the variance. The latter result is not due to a limited range of word frequency values, since all items entered in all analyses had a value from Luo and Yuan (1996).

## Discussion

The results from the preparatory study and the picture naming experiment show that the variables that predict timed picture naming latency in Chinese are similar to those that predict performance in a variety of Indo-European languages: name agreement, rated familiarity, and AoA. However, the results contrast with Bates et al. (2003), as we found no effect of printed-word frequency on performance. We also identified an effect of image agreement as in other studies of timed picture naming (see Alario et al., 2004); this effect may be similar to the GoD effect reported by Bates et al. Image agreement refers to how closely a picture resembled the mental image of an object using a 5-point scale in which 1 indicated *very good match* and 5 *very poor match*, whereas GoD estimates were subjective ratings of object pictures presented to a group of U.S. college students asked to rate, also on a 5-point scale, how well a picture fit its dominant name. Our ratings were obtained from native Putonghua speakers, and thus may be more reliable than estimates from U.S. students. It is clear from both studies that image quality has an impact on timed naming performance. These results may have theoretical value to the modeling of picture naming, given that both image agreement and GoD probably reflect early

Table 2  
Correlation Matrix for Predictor Variables and Naming Latency in Experiment 1

Variable	1	2	3	4	5	6	7	8	9
1. Naming RT									
2. Visual complexity	.12								
3. Image agreement	-.20 <sup>‡</sup>	.00							
4. Name agreement ( $H$ )	-.42*	.00	.13						
5. Percentage name agreement	-.39*	.08	.13	-.83*					
6. Concept agreement	-.23*	.00	.14	-.49*	.55*				
7. AoA	.27*	.13	.07	.04	.09	.09			
8. Rated familiarity	-.40*	-.41*	.08	.05	.04	.26*	-.25*		
9. Log (1 + frequency)	-.25*	.04	.08	.02	.05	.04	-.50*	.23*	
10. Syllable number	.07	.10	.06	-.11	.12	.05	.44*	-.15	-.48*

\* $p < .01$ . <sup>‡</sup> $p < .05$ .

visual processing in naming (Alario et al., 2004; Bates et al., 2003). However, these findings also suggest that the quality of line drawings may be an impediment to identifying the key variables that predict naming latency. Black-and-white line drawings are the standard used in all studies of picture naming, but colored pictures of objects are now available (Rossion & Pourtois, 2004). No study of picture naming has yet used these stimuli. We hypothesized that image agreement might have a reduced impact on timed picture naming when colored stimuli are presented, and we tested this prediction in Experiment 2.

**EXPERIMENT 2**

**Method**

Participants were 100 undergraduate students from Beijing Normal University. All were Putonghua speakers, and each was paid 10 yuan for participation. No participant was involved in the rating exercises described above. The procedure was similar to that in Experiment 1, except that all 232 items reported in the Appendix (see archive) were presented, and the stimuli were color drawings of the Snodgrass and Vanderwart (1980) objects produced by Rossion and Pourtois (2004) and accessed from the Web site [www.cog.brown.edu/~tarr/stimuli.html](http://www.cog.brown.edu/~tarr/stimuli.html). The first block of stimuli consisted of 12 practice trials and 116 experimental trials. There was then a brief rest period followed by the second set of stimuli, which consisted of 12 practice and 116 experimental trials. Stimulus presentation was randomized for each participant.

**Results**

There were 5.6% spoiled responses and 2.6% voice key failures. Scores with more than two standard deviations from the mean were removed (3.5% of responses). The mean of all naming times was then calculated for each item. Mean naming latencies for each item are shown in the Appendix (see archive), along with the proportion of nontarget responses to each item. The global mean naming latency for the 232 items was 1,121 msec, which was not significantly different from the global mean in Experiment 1 ( $t < 1$ ). The correlation between naming latency in Experiments 1 and 2 was highly significant ( $r = .85, p < .0001$ ). Significant correlations between predictor variables and naming RTs are shown in Table 3. As in Experiment 1, the predictor variables with the highest correlations with naming latency are familiarity, followed by measures of name agreement ( $H$ , percentage name agreement, and concept agreement) and rated AoA.

**Stepwise multiple regression.** Stepwise multiple regression was performed using the predictor variables in Table 2. Rated familiarity was the first variable entered ( $b = -.40, t = 6.2, p < .001$ ) followed by  $H$  ( $b = -.22, t = 2.6, p < .05$ ), then by AoA ( $b = .24, t = 3.8, p < .001$ ). Adjusted  $R^2$  for these factors was equal to 0.46, which was significant [ $F(8,223) = 18.85, p < .001$ ]. No other variable in Table 3 was a significant source of variance.

Inspection of the Appendix (see archive) shows that values for word frequency are missing for many items. This could prejudice the results against an effect of printed-word frequency. We therefore performed an analysis including only items for which frequency data were available ( $n = 144$ ). Rated familiarity was the first variable entered ( $b = -.45, t = 5.3, p < .001$ ), followed by percentage name agreement ( $b = -.24, t = 1.9, p < .05$ ). Adjusted  $R^2$  for these factors was .38, which was significant [ $F(9,134) = 10.01, p < .001$ ]. No other variable in Table 3 was a significant source of variance. The latter result cannot be due to a limited range of word frequency values, because all items entered in all analyses had a word frequency value from Luo and Yuan (1996).

**Simultaneous multiple regressions.** Multiple regressions can be problematic when the predictor variables are intercorrelated (Morris, 1981). This may be a problem for our regression analyses, because we used three measures of name agreement (all correlated), and each was entered into the regression analyses. To confirm the results, simultaneous multiple regression analyses were also performed. Analyses had rated familiarity, AoA, image agreement, visual complexity, number of syllables, and written frequency as predictor variables, as well as one measure of name agreement. Table 4 shows the results of analyses with  $H$  as the measure of name agreement. The overall regression equation was highly significant [ $F(7,225) = 29.26, p < .001, R^2 = .69$ ]. Taken together, the predictor variables accounted for 48% of the variance. These results show a similar pattern to the results from the stepwise multiple regression analyses because name agreement, familiarity, and AoA all made unique independent contributions to naming speed. All other measures of name agreement were significant in every other simultaneous regression that was performed (as were rated familiarity and AoA). However, written word frequency made no sig-

**Table 3**  
**Correlation Matrix for Predictor Variables and Naming Latency in Experiment 2**

Variable	1	2	3	4	5	6	7	8	9
1. Naming RT									
2. Visual complexity	.09								
3. Image agreement	-.17	.05							
4. Name agreement ( $H$ )	-.47*	.03	-.16						
5. Percentage name agreement	-.44*	.07	.16	-.80*					
6. Concept agreement	-.41*	.01	.17	-.52*	.59*				
7. AoA	.41*	.20*	.05	.14	.10	-.13			
8. Rated familiarity	-.53*	-.40*	.13	-.19	.18	.37*	-.56*		
9. Log (1 + frequency)	.03	.14	-.03	.08	.07	.18‡	-.18‡	.21‡	
10. Syllable number	.06	.08	.02	-.01	.03	.04	.40*	-.15	-.48*

\* $p < .01$ . ‡ $p < .05$ .

**Table 4**  
**Simultaneous Multiple Regression Analysis on**  
**Naming Latency in Experiment 2**

Variable	<i>M</i>	<i>SD</i>	<i>b</i> Coefficient	<i>SE</i>	<i>t</i> Value
Name agreement ( <i>H</i> )	0.97	0.50	-0.39	0.08	5.51*
Rated familiarity	3.80	0.80	-0.32	0.08	5.37*
AoA	4.92	1.57	0.26	0.07	3.91*
Visual complexity	2.88	0.88	-0.09	0.01	1.63
Image agreement	3.49	0.64	-0.07	0.06	1.37
Concept agreement	0.82	0.21	-0.05	0.03	0.94
Log (1 + frequency)	1.51	0.72	-0.001	0.01	0.02
Syllable number	1.98	0.55	-0.001	0.01	0.01

\* $p < .01$ .

nificant independent contribution to naming time nor did syllable number, visual complexity or image agreement (see note 1).

## GENERAL DISCUSSION

The results from Experiment 2 replicate the main findings from Experiment 1, but with more participants and stimuli. Results from both experiments show clearly that name agreement, rated familiarity, and AoA have the greatest impact on timed picture naming in Putonghua. Our results are compatible with studies in American English (Snodgrass & Yuditsky, 1996), British English (Ellis & Morrison, 1998), Dutch (Severens et al., 2005), French (Alario et al., 2004), Icelandic (Pind & Tryggvadottir, 2002), Italian (Dell'Acqua et al., 2000), and Spanish (Cuetos et al., 1999). Indeed, AoA and name agreement have significant independent effects on picture naming latency in every language studied, and all studies that include these variables report significant effects on picture naming latency (this was noted by Alario et al., 2004). The present study is the first to extend these observations to picture naming in Chinese speakers.

Why do these variables influence picture naming latency? According to Alario et al. (2004), cognitive models of picture naming assume that lexical name retrieval is accomplished by a sequence of at least four separate processes: (1) activation of stored structural knowledge about an object's appearance; (2) activation of semantic information; (3) name retrieval; and (4) articulation. Each level of cognitive processing can be investigated by examining independently the effects of specific predictor variables on naming speed. For example, visual complexity and image agreement probably reflect recognizing an image as a representation of a familiar object. We found some evidence that image agreement was a significant constraint on picture naming in Chinese when participants named line drawings. However, this effect was eliminated when we presented color images; this suggests that effects of visual variables in previous studies of picture naming might be due to the use of images of poor quality. Rated familiarity can be thought of as the perceptual equivalent of word frequency: The more often an object is encountered, the more familiar—and, therefore, the more easily recognized—it will be (see Cuetos et al., 1999; Ellis & Morrison, 1998). Another possibility is that the effects of

familiarity operate at the stage of semantic processing, reflecting processing of the meaning of an object (Bates et al., 2003).

Measures of name agreement probably reflect processes of lexical retrieval—specifically, inhibition of competitor names for output—whereas the measure of the number of syllables most likely operates at the level of phonological output. According to Alario et al. (2004), name agreement reflects a stage of processing which involves accessing the spoken name. This interpretation of the effects of name agreement is consistent with proposals in other studies of picture naming (e.g., Lachman, Shaffer, & Hennrikus, 1974; Snodgrass & Yuditsky, 1996; Vitkovitch & Tyrrell, 1995). In those studies, it is usually assumed that if an object can be given more than one plausible name, picture naming is slowed by the need to select between competitors for response. However, name agreement may have a relatively strong influence on picture naming in Chinese because of the small number of syllables in Chinese (around 400). Each syllable can be expressed with a different tone and in a unique order to convey the spoken name of an object. This requires selection as well as inhibition of competitor syllables (with a different tone). The requirement to select and inhibit competitors suggests that a lemma level of processing is required in Chinese spoken word production; if so, the effects of name agreement could also be located at this stage.

One possible locus for the effect of AoA is at the lemma level (Belke, Brysbaert, Meyer, & Ghyselinck, 2005). A different view is that AoA has an impact via the mappings between semantic memory and spoken name representations (Barry et al., 1997). There is little in our data to distinguish between these possibilities. However, our data are compatible with the modest claim that AoA influences the retrieval of phonological word forms in Putonghua.

We found that printed-word frequency had no impact on picture naming latency. Other studies of picture naming have also failed to observe an effect of printed-word frequency on naming speed, when correlated variables such as AoA are controlled statistically (e.g., Bonin et al., 2002; Carroll & White, 1973; Dell'Acqua et al., 2000; Gilhooly & Gilhooly, 1979; Morrison et al., 1992). Despite this, recent studies using a relatively large number of items have found an influence of printed-word frequency over and above effects of AoA (Barry et al., 1997; Cuetos et al., 1999; Ellis & Morrison, 1998; Snodgrass & Yuditsky, 1996). Moreover, Bates et al. (2003) reported an effect of word frequency on picture naming in Chinese (see also Zhang & Yang, 2003).

One reason for the discrepant results may be selection of printed-word frequency values. Our estimates were based on printed samples from a large corpus of words but these may not be the best estimate of frequency of exposure to words in Putonghua. We acknowledge that printed-word frequency could have an effect on picture naming latency, using larger corpora. However, we question whether estimates of printed-word frequency are meaningful predictors of picture naming latency. Ratings of familiarity are correlated with objective measures of printed-word frequency (Peereman, Content, & Bonin, 1998). We con-

tend that familiarity ratings may be preferable estimates of the effects of word frequency on spoken word processing in Putonghua. Our findings can thereby be interpreted as supporting the view that frequency has a significant impact on picture naming in Putonghua when the effects of AoA and name agreement are controlled. Previous studies report nonsignificant effects of frequency once AoA is controlled (Morrison et al., 1992). However, in studies using a relatively large number of stimuli, frequency remains significant, even when the effects of AoA are controlled (Barry et al., 1997; Snodgrass & Yuditsky, 1996). We did not use the full set of Snodgrass and Vanderwart (1980) pictures in Experiments 1 or 2, and Bates et al. (2003) tested over 500 objects. This could explain our failure to find a significant effect of printed-word frequency (although Cuetos et al., 1999, and Pind & Tryggvadottir, 2002, found a significant effect with 140 stimuli and 175 items, respectively). It is also important to note that we used estimates of frequency based on texts read by adults. Zevin and Seidenberg (2002, 2004) argued that estimating frequency by using measures that incorporate exposure to written words in childhood is a more reliable method. This is not yet possible in Chinese, but is a part of our ongoing research program.

In sum, the results of the present study show that the predictors of picture naming in Chinese are remarkably similar to those that predict naming latency in Indo-European languages. The other novel finding is that AoA has an effect on picture naming in Putonghua; this adds to the growing literature on the robust effects of AoA on lexical processing across languages (Juhasz, 2005). These data are of value to theorists engaged in modeling lexical processing that are intended to be universal in application.

#### AUTHOR NOTE

This research was supported by Grant HKU7275/03H from the Research Grants Council of the Hong Kong government. Correspondence should be addressed to B. S. Weekes, Department of Psychology, School of Life Sciences, University of Sussex, Falmer, Brighton BN1 9QG, England (e-mail: bsw@biols.susx.ac.uk).

#### REFERENCES

- ALARIO, F.-X., FERRAND, L., LAGANARO, M., NEW, B., FRAUENFELDER, U., & SEGUI, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, **36**, 140-155.
- BARRY, C., MORRISON, C., & ELLIS, A. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency and name agreement. *Quarterly Journal of Experimental Psychology*, **50A**, 560-585.
- BATES, E., D'AMICO, S., JACOBSEN, T., SZÉKELY, A., ANDONOVA, E., DEVESCOVI, A., ET AL. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, **10**, 344-380.
- BELKE, E., BRYLSBAERT, M., MEYER, A. S., & GHYSELINCK, M. (2005). Age of acquisition effects in picture naming: Evidence for a lexical-semantic competition hypothesis. *Cognition*, **96**, B45-B54.
- BONIN, P., CHALARD, M., MÉOT, A., & FAYOL, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, **93**, 89-114.
- BONIN, P., PEEREMAN, R., MALARDIER, N., MÉOT, A., & CHALARD, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments, & Computers*, **35**, 158-167.
- CARROLL, J., & WHITE, M. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology*, **25**, 85-95.
- CHALARD, M., BONIN, P., MÉOT, A., BOYER, B., & FAYOL, M. (2003). Objective age-of-acquisition (AOA) norms for a set of 230 object names in French: Relationships with psycholinguistic variables, the English data from Morrison et al. (1997), and naming latencies. *European Journal of Cognitive Psychology*, **15**, 209-245.
- CUETOS, F., ELLIS, A., & ALVAREZ, B. (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments, & Computers*, **31**, 650-658.
- DELL'ACQUA, R., LOTTO, L., & JOB, R. (2000). Naming times and standardized norms for the Italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods, Instruments, & Computers*, **32**, 588-615.
- ELLIS, A. W., & MORRISON, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 515-523.
- FORSTER, K. I., & FORSTER, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, **35**, 116-124.
- GERNSBACHER, M. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness and polysemy. *Journal of Experimental Psychology*, **113**, 256-281.
- GILHOOLY, K. J., & GILHOOLY, M. L. (1979). Age-of-acquisition effects in lexical and episodic memory tasks. *Memory & Cognition*, **7**, 214-223.
- JESCHENIAK, J., & LEVELT, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 824-843.
- JUHASZ, B. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin*, **131**, 684-712.
- KREMIN, H., HAMEREL, M., DORDAIN, M., DE WILDE, M., & PERRIER, D. (2000). Age of acquisition and name agreement as predictors of mean response latencies in picture naming of French adults. *Brain & Cognition*, **43**, 286-291.
- LACHMAN, R., SHAFFER, J., & HENNRICKUS, D. (1974). Language and cognition: Effects of stimulus codability, name-word frequency, and age of acquisition on lexical reaction time. *Journal of Verbal Learning & Verbal Behavior*, **13**, 613-625.
- LEVELT, W. J. M. (2002). Picture naming and word frequency: Comments on Alario, Costa and Caramazza. *Language & Cognitive Processes*, **17**, 663-671.
- LUO, Z., & YUAN, Y. (Eds.) (1996). *The research of Chinese and characters in the computer era*. Beijing: Tsinghua University Press.
- MORRIS, P. E. (1981). Age of acquisition, imagery, recall, and the limitations of multiple-regression analysis. *Memory & Cognition*, **9**, 277-282.
- MORRISON, C. M., CHAPPELL, T. D., & ELLIS, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology*, **50A**, 528-559.
- MORRISON, C. M., & ELLIS, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology*, **91**, 167-180.
- MORRISON, C. M., ELLIS, A. W., & QUINLAN, P. T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory & Cognition*, **20**, 705-714.
- NISHIMOTO, T., MIYAWAKI, K., UEDA, T., UNE, Y., & TAKAHASHI, M. (2005). Japanese normative set of 359 pictures. *Behavior Research Methods*, **37**, 398-416.
- PEEREMAN, R., CONTENT, A., & BONIN, P. (1998). Is perception a two-way street? The case of feedback consistency in visual word recognition. *Journal of Memory & Language*, **39**, 151-174.
- PÉREZ, M. A., & NAVALÓN, C. (2005). Objective-AoA norms for 175 names in Spanish: Relationships with other psycholinguistic variables, estimated AoA, and data from other languages. *European Journal of Cognitive Psychology*, **17**, 179-206.

- PIND, J., & TRYGGVADOTTIR, H. B. J. (2002). Determinants of picture naming times in Icelandic. *Scandinavian Journal of Psychology*, **43**, 221-226.
- ROSSION, B., & POURTOIS, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface feature detail in basic-level object recognition. *Perception*, **33**, 217-236.
- SANFELIU, M. C., & FERNANDEZ, A. (1996). A set of 254 Snodgrass-Vanderwart pictures standardized for Spanish: Norms for name agreement, image agreement, familiarity, and visual complexity. *Behavior Research Methods, Instruments, & Computers*, **28**, 537-555.
- SCHWITTER, V., BOYER, B., MÉOT, A., BONIN, P., & LAGANARO, M. (2004). French normative data and naming times for action pictures. *Behavior Research Methods, Instruments, & Computers*, **36**, 564-576.
- SEVERENS, E., VAN LOMMEL, S., RATINCKX, E., & HARTSUIKER, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, **119**, 159-187.
- SHU, H., CHENG, Y., & ZHANG, H. (1989). Name agreement, familiarity, image agreement and visual complexity for 235 pictures. *Acta Psychologica Sinica*, **21**, 389-396.
- SNODGRASS, J., & VANDERWART, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, **6**, 174-215.
- SNODGRASS, J. G., & YUDITSKY, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, & Computers*, **28**, 516-536.
- SZÉKELY, A., D'AMICO, S., DEVESCOVI, A., FEDERMEIER, K., HERRON, D., IYER, G., ET AL. (2003). Timed picture naming: Extended norms and validation against previous studies. *Behavior Research Methods, Instruments, & Computers*, **35**, 621-633.
- SZÉKELY, A., JACOBSEN, T., D'AMICO, S., DEVESCOVI, A., ANDONOVA, E., HERRON, D., ET AL. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory & Language*, **51**, 247-250.
- VITKOVITCH, M., & TYRRELL, L. (1995). Sources of disagreement in object naming. *Quarterly Journal of Experimental Psychology*, **48A**, 822-848.
- ZEVIN, J. D., & SEIDENBERG, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory & Language*, **47**, 1-29.
- ZEVIN, J. D., & SEIDENBERG, M. S. (2004). Age-of-acquisition effects in reading aloud: Tests of cumulative frequency and frequency trajectory. *Memory & Cognition*, **32**, 31-38.
- ZHANG, Q. F., & YANG, Y. F. (2003). The determiners of picture naming latency [in Chinese]. *Acta Psychologica Sinica*, **35**, 447-454.

#### NOTE

1. Note that the effects of image agreement and visual complexity may be underestimated in this analysis because we presented colored pictures.

#### ARCHIVED MATERIALS

The following materials associated with this article may be accessed through the Psychonomic Society's Norms, Stimuli, and Data archive, [www.psychonomic.org/archive](http://www.psychonomic.org/archive).

To access these files, search the archive for this article using the first author's name (Weekes).

FILE: Weekes-BRM-2007.zip

DESCRIPTION: The compressed archive file contains two files:

WeekesEtAl(2007).txt, containing normative data for 232 Chinese nouns.

WeekesEtAl(2007).xls, containing the information above in Excel spreadsheet format.

AUTHOR'S E-MAIL ADDRESS: [bsw@biols.susx.ac.uk](mailto:bsw@biols.susx.ac.uk).

(Manuscript received November 9, 2005;  
revision accepted for publication February 28, 2006.)