THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2007, 60 (1), 32–42

Psychology Press
Taylor & Francis Group

Age of acquisition persists as the main factor in picture naming when cumulative word frequency and frequency trajectory are controlled

Miguel A. Pérez
University of Murcia, Murcia, Spain, and University of Wales Swansea, Swansea, UK

The aim of this study was to address the effect of *objective* age of acquisition (AoA) on picture-naming latencies when different measures of frequency (cumulative and adult word frequency) and frequency trajectory are taken into account. A total of 80 Spanish participants named a set of 178 pictures. Several multiple regression analyses assessed the influence of AoA, word frequency, frequency trajectory, object familiarity, name agreement, image agreement, image variability, name length, and orthographic neighbourhood density on naming times. The results revealed that AoA is the main predictor of picture-naming times. Cumulative frequency and adult word frequency (written or spoken) appeared as important factors in picture naming, but frequency trajectory and object familiarity did not. Other significant variables were image agreement, image variability, and neighbourhood density. These results (a) provide additional evidence of the predictive power of AoA in naming times independent of word-frequency and (b) suggest that image variability and neighbourhood density should also be taken into account in models of lexical production.

The *age of acquisition* (AoA) effect on lexical processing has been well established in several tasks over the last 30 years (see Juhasz, 2005, for review). The observed effect is that words learned earlier in life are processed faster and with greater accuracy than words acquired later, all else being equal. In picture naming, numerous studies have observed strong and independent effects of AoA and word frequency, with faster response times for high-frequency or early-acquired words (e.g., Cuetos, Ellis, & Alvarez, 1999; Ellis & Morrison, 1998) than for low-frequency or late-acquired words, respectively. Other studies, however, have reported significant AoA effects but no frequency effects (e.g., Bonin, Chalard, Méot, & Fayol, 2002), and only Barry, Morrison, and Ellis (1997) encountered an interaction between the two variables, with spoken word frequency affecting object naming times mainly for items with later acquired names.

Correspondence should be addressed to Miguel A. Pérez, Dpto. de Psicología Básica y Metodología, Universidad de Murcia, Campus de Espinardo, s/n–30080, Murcia, Spain. E-mail: maperez@um.es

This research was supported by Grant PL/17/FS/00 from the Séneca Foundation and by 2004 mobility grants from the Regional Ministry of the Education and Culture, Región de Murcia, Spain.

I would like to thank Rosario Zaragoza, María Padilla, Mila Muñoz, Begoña del Egido, Juani Belchí, Maribel Méndez, and Bárbara Peñalver for their help and collaboration, and a special acknowledgment to Hans Stadthagen-Gonzalez, Cristina Izura, and Eamon Fulcher for their suggestions and appropriate corrections on an earlier version of this paper. Also, I am thankful to A. W. Ellis, M. Perea, and an anonymous reviewer for their helpful comments.

http://www.psypress.com/qjep
DOI:10.1080/17470210600577423

A current issue open to discussion is whether the AoA is a reflection of cumulative or lifespan frequency (Brown & Watson, 1987; Carroll & White, 1973). Some authors have proposed that there could be a confound between AoA and *cumulative frequency* (e.g., Lewis, Gerhand, & Ellis, 2001), and some others have proposed that *frequency trajectory* should be used as the main measure of AoA (Zevin & Seidenberg, 2002). Cumulative frequency refers to the total number of times that a word is encountered in an individual's lifetime. Given that the frequencies with which words are encountered vary considerably over a lifetime, the frequencies obtained from adult texts should be adjusted according to the frequency of word occurrence throughout development. Frequency trajectory is used to denote the distribution of frequency over a lifetime.

Zevin and Seidenberg (2002) proposed that AoA of words does not affect word naming because the mapping between spelling and sounds in English is not completely arbitrary, but that cumulative and trajectory frequency do. In addition, they do not deny that AoA might affect other processes in which there are arbitrary mappings between input and output—for instance, picture naming (where there are arbitrary mappings between pictures and their names).

In this respect, recent evidence shows that controlling for cumulative frequency and frequency trajectory does not result in the removal of an effect of AoA in word recognition and word-naming tasks (Bonin, Barry, Méot, & Chalard, 2004, in French; Cuetos & Barbón, 2006, in Spanish; Ghyselinck, Lewis, & Brysbaert, 2004b; Stadthagen-Gonzalez, Bowers, & Damian, 2004, in English). Nevertheless, Zevin and Seidenberg (2004) failed to find AoA effects in reading aloud when cumulative frequency was controlled for, suggesting that cumulative frequency is a confounding variable on the usual AoA effect.

In picture naming, only the Morrison, Hirsh, Chappell, and Ellis (2002) study has attempted to evaluate the cumulative frequency hypothesis by analysing whether or not the size of the AoA effect could be reduced with the age. They found an AoA effect in both elderly and young adults, despite the fact that the difference in cumulative frequency between early and late words should be less in the elderly group than in the other. However, cumulative frequency and frequency trajectory were not explicitly controlled for in that study. The present investigation extends Morrison et al.'s work by attempting to provide more evidence of the role of AoA on Spanish picture naming when cumulative or adult word frequency (written and spoken), frequency trajectory, and other psycholinguistic variables are taken into account.

Typically, the AoA values used in experiments are obtained from adults' estimations (e.g., Cuetos et al., 1999). These estimations have been shown to be valid, since high correlations between rated and objective AoA measures, calculated from oral production in children, have been reported (e.g., Morrison, Chappell, & Ellis, 1997). However, several studies have also found that familiarity has a higher correlation with estimated AoA than does objective AoA (e.g., Chalard, Bonin, Méot, & Fayol, 2003; Morrison et al., 1997; Pérez & Navalón, 2005), suggesting that AoA ratings may be biased by inferences based on various characteristics of the concept or name, such as its familiarity or frequency (see Bonin et al., 2004). Moreover, Chalard et al. and Bonin et al. have shown that objective AoA is a stronger predictor of spoken and written word-naming latencies than is estimated AoA. Hence, it may make better sense to use objective AoA measures because the differences between rated and objective AoA, although small, may be critical when any AoA influence on lexical processing or its possible relationships with other variables are assessed.

Another issue worth considering is that while in all the studies cited above written word frequency (adult) and/or object familiarity were taken into account, only two studies to date have also considered spoken word frequency measures separately (e.g., from CELEX). Interestingly, Barry et al. (1997) found spoken word frequency (CELEX) to be the stronger predictor (highest β-coefficient) on picture-naming times, when many other variables were included in the multiple regression analysis. This result is consistent with the assumption that spoken frequency, rather than written frequency, is a more precise estimate of the activation thresholds of phonological representations. This is important because phonological representations are necessarily activated in lexical production tasks while orthographic representations are not (see Levelt, Roelofs, & Meyer, 1999). However, Morrison et al. (1997) found negligible differences in the results whether written or spoken frequency was used on picture-naming latencies, although it is worth noting that they used spoken word frequencies from an outdated database (Howes, 1966; cited in Morrison et al., 1997). The role of spoken word frequency is also addressed in the present study.

Method

Participants

A total of 80 undergraduate students from the University of Murcia participated voluntarily in the experiment. In order to prevent fatigue effects, participants were randomly divided into two groups, and each group named half of the items. The proportion of men and women was approximately 2:3 in both groups, and the mean age was 20.7 years ($SD = 3.2$) in one group and 21.5 years ($SD = 3.1$) in the other. All participants were native speakers of Spanish, and none had speech disorders.

Materials

The stimuli were 178 pictures taken from the Pérez and Navalón (2003) battery. The pictures consisted of a black line drawing over a white background and were surrounded by a black frame. All items were objects with single-word names and fulfilled the following reliability criteria: (a) a *name agreement* (henceforth %NA) of at least 60%; (b) an *H-index*[1] of 1 or less; and (c) a mean *image agreement* score (henceforth IA) of at least 3 in a 5-point subjective scale. In addition, the following variables were collected and were added as factors in the subsequent regression analyses. The descriptive statistics of these variables are shown in Table 1. (An appendix with the norms used here and the naming times obtained is available for download at http://www.um.es/docencia/maperez/publications.html)

Table 1. *Descriptive statistics of items used in the experiment*

Variable	*Mean*	*SD*	*Min.*	*Max.*	*Skewness*
H-index	.34	.31	.00	1.00	.48
%NA	92.94	7.93	60.00	100.00	−1.57
IA	3.97	.47	3.00	4.91	−.22
FA	3.35	.99	1.58	4.85	−.05
VC	2.75	.79	1.28	4.61	.21
IV	2.24	.45	1.43	3.50	.41
Obj-AoA[a,b]	53.79	23.22	23.00	132.00	1.19
WF-AC[c]	49.40	181.67	.00	2,053.00	9.19
WF-LEXESP[c]	36.93	143.90	.00	1,633.80	9.35
WF-RAE[c]	37.65	178.26	.05	2,173.02	11.02
SF[c]	14.95	58.92	.00	595.67	8.07
CU-F[c]	113.00	316.44	1.25	3,081.93	7.30
FT[d]	.00	.67	−1.87	2.15	.03
NSyl	2.64	.83	1	5	.89
NLet	6.18	1.85	3	14	.85
NPhon	5.97	1.81	3	12	.65
ON	5.51	6.99	0	34	1.64
PN	5.58	7.82	0	35	1.89

Note: %NA, percentage of name agreement. IA, image agreement. FA, object familiarity. VC, visual complexity. IV, image variability. Obj-AoA, objective age of acquisition. WF-AC, written word frequency from Alameda and Cuetos (1995). WF-LEXESP, written word frequency from LEXESP (Sebastián, Marti, Carreiras, & Cuetos, 2000). WF-RAE, written word frequency from RAE (Pérez, Cuetos, & Alameda, 2003). SF, spoken word frequency from RAE (Pérez et al., 2003). CU-F, cumulative frequency. FT, frequency trajectory. NSyl, number of syllables. NLet, number of letters. NPhon, number of phonemes. ON, orthographical neighbourhood. PN, phonological neighbourhood.

[a]In months. [b]$N = 158$. [c]per million. [d]z-scores.

[1] The *H-index* is a logarithmic function that analyses the quantity of different names that an object receives. An H near to 0 indicates a very low dispersion—that is, high name agreement—while an H near to two indicates little unanimity in the object naming. H-index is defined as: $H = \sum_{i=1}^{k} p_i \log_2(1/p_i)$, where k is the number of different names produced to a picture, and p_i is the proportion of participants producing the ith name.

Object familiarity (henceforth FA), *visual complexity* (VC), and *image variability* (IV) were obtained from Pérez and Navalón's (2003) norms. Object familiarity corresponds to the participants' ratings of the frequency or degree to which they daily enter in contact with, or think about, the object drawn (scale from 1 = *very unfamiliar object* to 5 = *very familiar object*). Visual complexity was rated as the detail sum or intricacy of lines of the drawing (scale from 1 = *low complexity* to 5 = *high complexity*). Image variability refers to the participants' estimation of how many different images could be formed from the given name (scale from 1 = *low variability* to 5= *high variability*).

Objective age of acquisition (henceforth Obj-AoA) was collected from Pérez and Navalón's (2005) child-production norms. They asked children, divided in 10 age groups, to name objects aloud. The AoA assigned to a name was (a) the mean age of the group in which at least 75% of participants answered that name, provided that the upper following age groups also kept at least that percentage; and (b) the age predicted by the logistic-regression curve at $p = .75$, when the data fitted to it. The measure of AoA selected for this study was the so-called *combined-AoA* by Pérez and Navalón, which was set up adding those significant AoA data from the logistic-regression method to the 75%-rule data set.

Three different objective measures of *adult written frequency* were obtained: (a) frequencies from the *Linguistic-Units Dictionary of Spanish* (Alameda & Cuetos, 1995; henceforth WF-AC), which is based in a corpus of 2 million words; (b) frequencies from the *Computerized Database of the Spanish Language* (Sebastián, Martí, Carreiras, & Cuetos, 2000; henceforth, WF-LEXESP), which contains 5 million words; and (c) frequencies from the *Reference Corpus of the Current Spanish Language* (Real Academia Española, RAE, 2003; henceforth, WF-RAE), composed from approximately 125 million words,[2] (available at http://www.rae.es). The spoken word frequency (henceforth, SF) was also collected from the RAE's database, which compiles spoken frequency from a variety sources in Spanish, including TV and radio transmissions and formal and informal face-to-face recordings. It includes around 6 million words.

Cumulative frequency (CU-F) was calculated as the sum (per million) of adult frequency (WF-LEXESP) plus a child frequency measure. Child frequency was taken from the Martínez and García (2005) database, which samples 2,600,000 tokens from reading books as well as textbooks from Grades 1 to 6 (6 to 12 years of age). The sum of the frequencies from all six levels was used to determine child frequency. *Frequency trajectory* (FT) was computed with the following procedure: First, in order to reduce skewness, the logarithm of the adult and child frequency scores was computed; second, log-scores were transformed into z-scores; and third, the z-scores of child frequency were subtracted from the z-scores of adult frequency (see Bonin et al., 2004; Cuetos & Barbón, 2006, for a similar procedure).

The *orthographic neighbourhood* (defined as the number of words computed by substituting one letter at a time within the target word) of each picture name was collected from the Pérez, Cuetos, and Alameda's (2003) database (henceforth, ON). Scores of *phonological neighbourhood* (defined as the orthographical neighbourhood but substituting phonemes instead letters) were also collected from the "BuscaPalabras" database[3] (Davis & Perea, in press; henceforth, PN). Finally, word *length* was also calculated by counting the number of letters (NLet), syllables (NSyl), and phonemes (NPhon) of each picture name.

[2] Of the 125 million words (of which, 10% are from oral transcriptions), 50% come from sources published in Spain and the remainder from sources in Latin America. The data used in the current study were obtained by filtering out Latin American sources.

[3] There were six names that do not have phonological neighbourhood scores in the BuscaPalabras. One of them, "destonillador", was not found because the database limits the search for words that must not exceed 12 letters. The scores for the other words were not available because those words are not included in the LEXESP database, on which BuscaPalabras calculates the phonological neighbourhood. These missing data were replaced by the orthographical neighbourhood scores from Pérez et al. (2003).

Procedure
Stimuli were presented in the centre of a 2-GHz PC screen as black outline drawings over a white background. Each picture was preceded by a fixation cross on the centre of the screen for 1,000 ms. Participants were asked to name each picture as soon as possible, and the pictures disappeared after a response was registered or after a timeout of 3,000 ms. Participants' responses were recorded via a microphone connected to a voice-key that measured pronunciation onset. An E-Prime 1.1 routine controlled the stimuli presentation and the response registration (Schneider, Eschman, & Zuccolotto, 2002). Stimuli were presented in random order to each participant.

Participants were encouraged to respond clearly, by giving just one name, by avoiding articles (which are common in Spanish) and hesitations, and by preventing strong exhalations. When a participant could not give a response, he or she was asked to indicate whether this was because they did not recognize the object or because they did not know the name of the object. There were 30 practice trials with feedback that preceded the experimental trials. Responses were classified online by the experimenter as follows: (a) correct response, when a participant gave the target name of the object; (b) incorrect response, when a participant gave a name different from the target name; and (c) error, when the voice-key failed or participants inadvertently activated the voice-key (e.g., through a hesitation or exhalation).

Results

A total of 17 items (*donkey*, *tricycle*, *lemon*, *fox*, *apple*, *artichoke*, *skateboard*, *tureen*, *stomach*, *pliers*, *speedboat*, *photocopier*, *humming bird*, *nail*, *staple*, *apricot*, and *eel*) showed less than 60% of spoken naming accuracy and were therefore removed from further analyses. Once these items were removed, the percentage of correct answers was 92.9% (SD = 7.9). Incorrect or error responses and outliers (those exceeding twice the amplitude between Q1 and Q3 from mean latency, 3.8%) were also removed from further analysis. Overall, 7.8% of 6,440 observations were eliminated. The global mean naming latency was 855 ms.

In order to check that no statistical differences existed between the two sets of words used in the experiment, two independent sample t tests were carried out, one on accuracy and one on reaction time. These revealed no significant differences between the sets either on mean accuracy (93.2 and 92.6%), $t(159) < 1$, or on reaction time (RT; 842 and 867 ms), $t(159) = -1.228$, $p = .221$. This allowed for data from both sets to be merged and analysed together in the subsequent analyses.

Word initial-phoneme effect. To rule out the possibility of these results being caused by a word initial-phoneme effect (see Kessler, Treiman, & Mullennix, 2002), items were divided according to the initial-phoneme characteristics, and their mean latencies were then compared. There were five initial-phoneme categories: (a) voiceless stop consonants ($n = 64$); (b) voiced stop consonants ($n = 20$); (c) vowels or glides ($n = 23$); (d) liquid, nasal, labiodental, and interdental fricatives ($n = 40$); (e) palatal, alveolar, and postalveolar fricatives and affricates ($n = 14$). A one-way analysis of variance (ANOVA) on the mean RTs for these five categories showed no significant effect between the groups, $F(4, 156) = 1.085$, $MSE = 16{,}875$, $p = .366$.

Correlations among the variables. Table 2 shows the correlation matrix for all the independent variables and the naming latencies. In all subsequent analyses, frequency measures were transformed using the formula $log\ (1 + x)$ to reduce skew. All variables but three (%NA, H-index, and frequency trajectory) correlated significantly with RT. Many predictors also correlated with each other (e.g., Obj-AoA was highly correlated with frequency measures), meaning that the single correlations with the RT must be interpreted with caution.

Multiple regression analyses. The strong intercorrelation between some variables (see marked correlations in Table 2) could be problematic in a stepwise regression (see Morris, 1981). In order to avoid the problem of multicollinearity

Table 2. *Correlation matrix among all independent variables and picture-naming times*

	Variable	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
1.	Naming RT	1.000																	
2.	*H*-index	.150	1.000																
3.	%NA	−.040	−.929***	1.000															
4.	IA	−.121	−.064	.085	1.000														
5.	FA	−.325***	.025	−.039	.149	1.000													
6.	VC	.171**	.103	−.108	−.310***	−.422***	1.000												
7.	IV	−.337***	.194*	−.238**	−.162*	.344***	−.112	1.000											
8.	Obj-AoA[a]	.578***	.170*	−.081	.015	−.197*	.174*	−.202*	1.000										
9.	WF-AC	−.471***	−.053	.004	−.153	.406***	−.182*	.469***	−.411***	1.000									
10.	WF-LEXESP	−.413***	−.058	.021	−.141	.315***	−.132	.393***	−.343***	.901***	1.000								
11.	WF-RAE	−.465***	−.010	−.029	−.079	.391***	−.179*	.451***	−.374***	.904***	.864***	1.000							
12.	SF	−.423***	−.103	.061	−.129	.293***	−.037	.373***	−.358***	.820***	.778***	.852***	1.000						
13.	CU-F	−.513***	−.087	.042	−.141	.290***	−.157*	.338***	−.544***	.905***	.881***	.863***	.792***	1.000					
14.	FT	.082	.125	−.134	−.094	.204**	.000	.352***	.254***	.241**	.255***	.209**	.132	−.094	1.000				
15.	NSyl	.182**	−.029	−.045	.038	−.069	.208***	−.139	.218**	−.315***	−.277***	−.311***	−.294***	−.302***	−.032	1.000			
16.	NLet	.255***	−.046	−.013	.061	−.100	.160*	−.134	.222**	−.336***	−.300***	−.317***	−.316***	−.348***	.022	.888***	1.000		
17.	NPhon	.268***	−.036	−.025	.042	−.054	.140	−.136	.225**	−.334***	−.302***	−.306***	−.314***	−.346***	.023	.888***	.966***	1.000	
18.	ON	−.255***	.049	−.043	−.106	.085	−.085	.205**	−.134	.223**	.194*	.233**	.266***	.239**	−.012	−.463***	−.598***	−.599***	1.000
19.	PN	−.272***	.006	.003	−.121	.037	−.023	.174*	−.152	.242**	.241**	.243**	.310***	.278***	−.041***	−.456***	−.604***	−.627***	.920***

Note: RT, reaction time. %NA, percentage of name agreement. IA, image agreement. FA, object familiarity. VC, visual complexity. IV, image variability. Obj-AoA, objective age of acquisition. WF-AC, written word frequency from Alameda and Cuetos (1995). WF-LEXESP, written word frequency from LEXESP (Sebastián, Marti, Carreiras, & Cuetos, 2000). WF-RAE, written word frequency from RAE (Pérez, Cuetos, & Alameda, 2003). SF, spoken word frequency from RAE (Pérez et al., 2003). CU-F, cumulative frequency. FT, frequency trajectory. NSyl, number of syllables. NLet, number of letters. NPhon, number of phonemes. ON, orthographical neighbourhood. PN, phonological neighbourhood.

[a]$N = 158$.

*$p < .05$. **$p < .01$. ***$p < .001$.

between factors, only one variable from the pair of %NA and *H*-index and one from the quintet of NSyl, NLet, NPhon, ON, and PN[4] were introduce into the analysis (those that had the highest simple correlation with the RT, i.e., *H*-index and PN; see Pedhazur, 1997). Moreover, the potential multicollinearity of predictors was monitored by means of two statistical indices, *tolerance* and *variance inflation factor* (VIF),[5] guaranteeing the reliability of the analyses. In order to assess the importance of AoA on picture-naming latencies when both frequency (adult or cumulative) and frequency trajectory are taken into account, several hierarchical multiple regression analyses were carried out (see Pedhazur, 1997). Each analysis first included FT and one measure of word frequency (i.e., WF-AC, WF-LEXESP, WF-RAE, SF, or CU-F) by the *enter* method, and then *H*-index, PN, IA, FA, VC, IV, and Obj-AoA by the *stepwise* method.

The best model thus obtained (see Table 3) accounts for 47.6% of the variance, including effects of Obj-AoA, WF-AC, IA, IV, and PN, $F(6, 151) = 22.89$, $p < .001$; $R = .690$, with no significant contribution from *H*-index, FA, CV, or FT. Adequate indexes of independence of the predictors within the model were obtained. The independent variable that best accounted for the RT was Obj-AoA, followed by the WF-AC (see β-weights). In addition, there was a small but significant improvement in the global R^2 when IA, IV, and PN were included in the model.

Other models that emerged from the measures of cumulative or adult word frequency (written or spoken) were very similar to that shown in Table 3 using WF-AC. Table 4 shows the details of the β weights for Obj-AoA and word frequency, and the multiple *R* of the global model when the frequency measures were used in the regression analyses.

Table 3. *Better multiple regression model on the picture-naming latencies*

Variable	*Standardized β-coefficient*	*t-student*	*p*	*Tolerance*	*VIF*
FT	.072	1.031	.304	.707	1.415
WF-AC	−.233	−3.092	.002	.609	1.643
Obj-AoA[a]	.409	5.749	.000	.684	1.462
IA	−.202	−3.358	.001	.956	1.046
IV	−.178	−2.513	.013	.692	1.444
PN	−.144	−2.337	.021	.914	1.094
H-index	.089	1.447	.150	.907	1.103
FA	−.091	−1.329	.186	.740	1.351
VC	−.034	−.524	.601	.838	1.194

Note: FT, frequency trajectory. WF-AC, written word frequency from Alameda and Cuetos (1995). Obj-AoA, objective age of acquisition. IA, image agreement. IV, image variability. PN, phonological neighbourhood. FA, object familiarity. VC, visual complexity.

[a]Three items had missing values at Obj-AoA, but they were excluded pairwise from the analysis—that is, only when the Obj-AoA and one of the other variables were correlated for calculating a specific parameter.

Entry criterion at the stepwise method $p = .05$, removal criterion $p = .10$.

The same significant predictors emerged in each model, and the β-weights for Obj-AoA and word frequency did not changed significantly depending on whether cumulative or adult frequency was used.

Why did the H-index not affect naming times?. Name agreement and *H*-index have been systematically found to be strong predictors of picture-naming times. However, in the present experiment *H*-index did not emerge as a significant predictor. One possible explanation for this is the low variability of the *H*-index distribution (30% were scores of 0.0, and 50% were scores less than 0.24). In order to address this question, a *t* test on the

[4] It is worth noting that when the author carried out several multiple regression analyses including NSyl, NLet, or NPhon, neither appeared as significant predictor of RT, whether or not ON or PN was also included. When ON or PN and any of the length measures were simultaneously included, multicollinearity indexes were highly inadequate.

[5] Tolerance is the percentage of the variance in a given predictor that cannot be explained by the other predictors. When tolerance is close to 0, there is a high multicollinearity, and the standard error of the regression coefficients would be inflated. VIF is the inverse of tolerance, and it is usually considered that when it is very near or greater than 2 the multiple regression analysis might be problematic.

Table 4. *Beta-coefficients of the Obj-AoA and the frequency measures used in the five multiple regression analyses*

Frequency measure	*Std. β-coefficient*		*Mult. regression of model*
	Frequency	*Obj-AoA*	
WF-AC	−.233	.409	.690
WF-LEXESP	−.194	.434	.685
WF-RAE	−.217	.427	.689
SF	−.176	.447	.682
CU-F	−.220	.402	.688

Note: WF-AC, written word frequency from Alameda and Cuetos (1995). WF-LEXESP, written word frequency from LEXESP (Sebastián, Marti, Carreiras, & Cuetos, 2000). WF-RAE, written word frequency from RAE (Pérez, Cuetos, & Alameda, 2003). SF, spoken word frequency from RAE (Pérez et al., 2003). CU-F, cumulative frequency. All standardized β-coefficients were significant at $p < .05$.

mean RTs for two item groups divided at the 50th percentile in the *H*-index distribution was carried out. A significant difference between the mean RTs of the groups was observed (827 ms for the group formed with *H*-scores under percentile 50 and 881 ms for the group formed with the rest of *H*-scores), $t(159) = 2.673$, $p = .008$. The effect is such that pictures with low name dispersion were named faster than pictures with high name dispersion. These analyses seem to confirm that the narrow range of the *H*-index is the reason why this variable did not reach statistical significance in the regression model (see stimuli selection by reliability criteria in Materials section).

Discussion

The results of the present experiment show that objective AoA is a relevant predictor of picture-naming times in the Spanish language, even when cumulative frequency and frequency trajectory were controlled for. These results support the findings of Morrison et al. (2002, Exp. 2) in picture naming and also those observed by Cuetos and Barbón (2006) in a Spanish word-naming task.

Zevin and Seidenberg (2002) proposed that frequency trajectory is the main measure of AoA, rather than adult estimations or child picture naming (the so-called "objective" AoA). The correlation observed here between AoA and frequency trajectory was significant but rather low ($r = .254$), indicating that the frequency trajectory calculated here is not a good predictor of objective AoA. Similar results have been observed by Cuetos and Barbón (2006), who did not find a significant correlation between frequency trajectory and objective AoA and found a significant but small correlation with rated AoA ($r = .200$; $p < .01$). Bonin et al. (2004) also encountered weak relationships between those variables ($r = .215$ between frequency trajectory and rated AoA, and $r = .325$ with objective AoA, both at $p < .05$). On the other hand, frequency trajectory does not appear to be a relevant predictor of picture-naming times but objective AoA is, suggesting that the former is not relevant in picture naming. No effect of frequency trajectory was observed by Bonin et al. either (when they took into account objective AoA), and Cuetos and Barbón report a similar finding in word naming.

These results are broadly compatible with the idea that the locus of AoA is at the semantic level (e.g., Ghyselinck, Custers, & Brysbaert, 2004a; see also Steyvers & Tenenbaum, 2005, for a connectionist account) and also with the network plasticity hypothesis proposed by Ellis and Lambon Ralph (2000). The semantic account proposes that the AoA may be the most important factor in how concepts in the semantic system are organized, with the later learned concepts being built onto those learned earlier. Ellis and Lambon Ralph claim that the AoA effect is a consequence of the loss of plasticity in neural systems in which the mapping from input to output is arbitrary (see also Lambon Ralph & Ehsan, in press). Since, in picture naming, it is assumed that there is a semantic-to-lexical processing, the AoA effect could be due to the organization of concepts inside the semantic system (e.g., Ghyselinck et al., 2004a) or due to the arbitrary relationships between the concepts and the phonological forms of words at a connection level (Ellis & Lambon Ralph, 2000; see also Izura & Ellis, 2004).

Recently, Brysbaert and Ghyselinck (in press; see also Belke, Brysbaert, Meyer, & Ghyselinck, 2005) have indicated that in picture naming the size of the AoA effect is often larger than that of the frequency effect, while in word naming or lexical decision both factors have a similar impact. Thus, in picture naming there are one or more processes in which AoA exerts an influence but word frequency does not. The authors refer to this as the *frequency-independent AoA effect*. Since, in picture naming, only one concept can be selected to respond, Brysbaert and Ghyselinck suggest that the frequency-independent AoA effect emerges from the competition in the conceptual system or, in the terms of the Levelt et al.'s (1999) word production model, from the competition at the lemma level. The results presented here provide evidence that the size of the AoA effect is larger than the size of any frequency measure. This supports Brysbaert and Ghyselinck's suggestion of incorporating AoA as a factor independent of word frequency in models of lexical production and to not consider that both variables can be modelled in the same way (see Levelt et al.).

Regarding measures of word frequency, both adult and cumulative frequency were important predictors, with slight differences between them in the amount of explained variance. Written and spoken adult word frequencies make similar contributions to the RTs (see Morrison et al., 1997, for a similar result, but see also Barry et al., 1997). Note that this is counterintuitive if it is assumed that spoken, but not written, frequency has a stronger relation with the activation thresholds of *phonological* representations needed in a picture-naming task (see, e.g., Levelt et al., 1999).

Image agreement and image variability were also other important predictors of picture-naming times. The importance of image agreement in picture naming has been well established in numerous other studies (e.g., Bonin et al., 2004; Bonin et al., 2002; Chalard et al., 2003; Cuetos et al., 1999). However, image variability—that is, the number of different images evoked by the name of an object—has not been studied so thoroughly. Recently, Chalard et al. (also Bonin et al., 2002) found similar effects to those obtained here: The greater the number of evoked images, the shorter the naming time is. According to one idea pointed out by Bonin et al. (2002), the "pictured objects having high image variability scores possibly possess richer structural/semantic representations than those having low image variability scores" (pp. 104−105). Hence a speculative explanation of these results would be that the mental images evoked by a given word are different examples of the same object, distinguished from each other by some few, minor, but very familiar characteristics (since image variability and object familiarity are positively correlated). For example, *dog*, which is very familiar (mean score of 4.21 in a scale of 1 to 5, from data used in the current study), presents more evocations of different images (mean score of 2.36) than does a wild animal such as *hyena*, which is very unfamiliar (1.67) and will probably evoke few (1.64 out 5) images. Thus, if each of those images sends activation to the object name, a concept with a high number of associated images would be recognized quicker than others with less associated images.

Finally, the number of phonological neighbours was found to be significantly inversely correlated with latencies. This facilitative neighbourhood effect has been commonly found in reading-aloud tasks but not so often in picture naming. Nevertheless, the current result is not completely unexpected since Vitevitch and Sommers (2003, Exp. 3) showed that the number of neighbours and their frequencies facilitated picture-naming latencies, challenging certain models of speech production (e.g., Levelt et al., 1999).

Original manuscript received 21 June 2005
Accepted revision received 13 December 2005
First published online 30 March 2006

REFERENCES

Alameda, J. R., & Cuetos, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano* [Linguistic-units dictionary of Castillian]. Oviedo, Spain: Servicio de Publicaciones de la Universidad de Oviedo.

Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency and name agreement. *Quarterly Journal of Experimental Psychology, 50A*, 560–585.

Belke, E., Brysbaert, 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., Barry, C., Méot, A., & Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? *Journal of Memory and Language, 50*, 456–476.

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.

Brown, G. D., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory and Cognition, 15*, 208–216.

Brysbaert, M., & Ghyselinck, M. (in press). The effect of age-of-acquisition: Partly frequency-related partly frequency-independent. *Visual Cognition*.

Carroll, J. B., & White, M. N. (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., & 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*, 209–246.

Cuetos, F., & Barbón, A. (2006). Word naming in Spanish. *European Journal of Cognitive Psychology, 18*, 415–436.

Cuetos, F., Ellis, A. W., & Alvarez, B. (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments, and Computers, 31*, 650–658.

Davis, C. J., & Perea, M. (in press). BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods*.

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1103–1123.

Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 515–523.

Ghyselinck, M., Custers, R., & Brysbaert, M. (2004a). The effect of age of acquisition in visual word processing: Further evidence for the semantic hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 550–554.

Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004b). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica, 115*, 43–67.

Izura, C., & Ellis, A. W. (2004). Age of acquisition effects in translation judgement tasks. *Journal of Memory and Language, 50*, 165–181.

Juhasz, B. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin, 131*, 684–712.

Kessler, B., Treiman, R., & Mullennix, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language, 47*, 145–171.

Lambon Ralph, M. A., & Ehsan, S. (in press). Age of acquisition effects depend on the mapping between representations and the frequency of occurence: Empirical and computational evidence. *Visual Cognition*.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1–75.

Lewis, M. B., Gerhand, S., & Ellis, H. D. (2001). Re-evaluating age-of-acquisition effects: Are they simply cumulative-frequency effects? *Cognition*, *78*, 189–205.

Martínez, J. A., & García, E. (2005). *Diccionario de frecuencias del castellano escrito en niños de 6 a 12 años* [Written Castillian frequencies dictionary in 6-to-12-year-old children]. Salamanca, Spain: Servicio de Publicaciones de la Universidad Pontificia de Salamanca.

Morris, P. E. (1981). Age of acquisition, imagery, recall, and the limitations of multiple-regression analysis. *Memory and 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., Hirsh, K. W., Chappell, T. D., & Ellis, A. W. (2002). Age and age of acquisition: An evaluation of the cumulative frequency hypothesis. *European Journal of Cognitive Psychology*, *14*, 435–459.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research* (3rd ed.). New York: Harcourt Brace College Publishers.

Pérez, M. A., Cuetos, F., & Alameda, J. R. (2003). Frecuencia, longitud y vecindad ortográfica de las palabras de 3 a 16 letras del Diccionario de la Lengua Española (RAE, 1992) [Frequency, length, and orthographic neighbourhood of the 3-to-16-letter words from the Spanish Language Dictionary (RAE, 1992)]. *Revista Electrónica de Metodología Aplicada*, *8*, 1–10.

Pérez, M. A., & Navalón, C. (2003). Normas españolas de 290 nuevos dibujos: Acuerdo en la denominación, concordancia de la imagen, familiaridad, complejidad visual y variabilidad de la imagen [Spanish norms for 290 new drawings: Name-agreement, image-agreements, familiarity, visual-complexity, and image-variability]. *Psicológica, 24*, 215–241.

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.

Real Academia Española. (2003). *Corpus de Referencia del Español Actual (CREA)* [Reference Corpus of the Current Spanish Language]. Retrieved March 15, 2003, from http://www.rae.es

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime user's guide*. Pittsburgh, PA: Psychology Software Tools Inc.

Sebastián, N., Marti, M. A., Carreiras, M. F., & Cuetos, F. (2000). *LEXESP, léxico informatizado del Español* [LEXESP, Computerized database of the Spanish language]. Barcelona, Spain: Ediciones de la Universitat de Barcelona.

Stadthagen-Gonzalez, H., Bowers, J. S., & Damian, M. F. (2004). Age-of-acquisition effects in visual word recognition: Evidence from expert vocabularies. *Cognition, 93*, B11–B26.

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, *29*, 41–78.

Vitevitch, M. S., & Sommers, M. S. (2003). The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Memory and Cognition, 31*, 491–504.

Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and 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 and Cognition*, *32*, 31–38.