Age of acquisition effects
in the semantic processing of pictures

ROBERT A. JOHNSTON
University of Birmingham, Birmingham, England

and

CHRISTOPHER BARRY
University of Essex, Colchester, England

In two experiments, we investigated the role of age of acquisition (AoA) in the categorizing of objects in semantic tasks that do not require access to the object names. In both a *found inside or outside the house* (Experiment 1A) and a *smaller or larger than a loaf of bread* (Experiment 2A) classification task, objects with earlier-acquired names were categorized more quickly than those with later-acquired names. Experiments 1B and 2B also showed AoA effects on object-naming times for the same pictures. We conclude that AoA operates within the identification process in a fashion not simply restricted to name retrieval. These effects may be better explained in terms of the connectionist model proposed by Ellis and Lambon Ralph (2000) or by accounts that locate AoA within the semantic system (e.g., Brysbaert, Van Wijnendaele, & De Deyne, 2000; van Loon-Vervoorn, 1989).

Age of acquisition (AoA) is increasingly being considered as a major variable that influences word and picture processing. It has been claimed that, when other relevant variables (e.g., frequency) are controlled, early-acquired words are recognized and named more quickly than late acquired words (e.g., Brown & Watson, 1987; Gerhand & Barry, 1998, 1999; Morrison & Ellis, 1995, 2000), although Zevin and Seidenberg (2002) have criticized some studies of word reading for not having controlled adequately for the cumulative frequency of the words used. Similarly, an increasing number of (as yet uncontested) studies have shown that pictures labeled by early-acquired words are named more quickly than objects with late-acquired labels (e.g., Barry, Hirsh, Johnston, & Williams, 2001; Barry, Morrison, & Ellis, 1997; Bonin, Fayol, & Chalard, 2001; Carroll & White, 1973; Gilhooly & Gilhooly, 1979; Lachman, Shaffer, & Hennrikus, 1974; Meschyan & Hernandez, 2002; Morrison, Ellis, & Quinlan, 1992; Snodgrass & Yuditsky, 1996). Zevin and Seidenberg (2002) have acknowledged that there should be AoA effects in picture naming even if the results for word naming are under contention. In order to better understand the independent contributions of AoA and frequency, it will be helpful to locate the operation of AoA in a more precise way than is currently available. The earliest explanations offered for AoA located its influence solely at the stage of name retrieval (e.g., Brown & Watson, 1987), whereas more recent accounts have proposed that it can operate at other locations as well. A robust demonstration of an AoA effect in a picture-processing task that did not require access to name representations would effectively rule out some explanations and would help to increase our understanding of others.

There exist several competing explanations of the operation of AoA effects. One of the earliest, proposed to account for AoA effects in word naming, was the *phonological completeness* hypothesis advanced by Brown and Watson (1987). They suggested that "phonological output representations are stored in a relatively complete form during the early stages of vocabulary acquisition" (p. 214), in contrast to later-acquired words, which are assumed to be stored in a more segmented fashion and for which "only minimal information is stored explicitly" (p. 215). This account assumes that later-acquired words, require their stored phonology to be assembled for spoken word production, which entails longer processing times, and so locates AoA effects at the stage of name retrieval, rather than at those of other aspects of lexical processing. However, Monaghan and Ellis (2002a) recently tested the phonological completeness hypothesis, using phonological awareness tasks, and found no evidence to support it. In addition, other recent research has begun to show an influence of AoA on several tasks that do not seem to require access to the target's name. These tasks include distinguishing between real and chimeric objects in the object decision task (Vitkovitch & Tyrrell, 1995), making face familiarity judgments (Moore & Valentine, 1999), categorizing celebrities according to the television show in which they appear (Lewis, 1999), and generating an associate to a presented word (Brysbaert, Van Wijnendaele, & De Deyne, 2000). AoA effects in these tasks are not easily accommodated within the phonological completeness hypothesis.

An alternative—and indeed, a more general—account of AoA effects has been advanced by Ellis and Lambon Ralph (2000). They proposed that AoA effects are a natural property of connectionist networks that are trained with cumulative or interleaved presentation. Such cumulative training, whereby later patterns are added to the training set after the early- or first-entered patterns had been learned, but where the early patterns also continue to be presented, avoids "catastrophic interference" (e.g., McClelland, McNaughton, & O'Reilly, 1995) and could be seen as reflecting normal vocabulary development. Ellis and Lambon Ralph's simulations showed that early patterns are favored over later patterns, and they proposed that this arises from a gradual reduction of network plasticity, which causes later-learned items to be differentiated less effectively than earlier ones. On this account, one would expect that AoA effects will be found in all tasks in which learning is cumulative and, especially, in ones in which the mapping relations between input and output representations (e.g., between a picture and its name) are arbitrary or inconsistent (Ellis & Lambon Ralph, 2000; Monaghan & Ellis, 2002b; Zevin & Seidenberg, 2002). Monaghan and Ellis (2002b) reported that AoA had a reliable effect on reading latencies only for words with inconsistent spelling-to-sound relationships (such as *steak* and *pint*); there was no AoA effect for the reading of words with consistent correspondences (such as *grape* and *weep*). This result follows from connectionist models of AoA that show larger AoA effects when there are arbitrary or inconsistent mappings between input and outputs, as is the case for the relationship between objects and their names. However, it should also hold for the relationship between objects and their semantic representations.

An alternative explanation can be derived from the semantic hypothesis put forward by Brysbaert et al. (2000). Essentially, this position asserts that AoA or order of acquisition has a central role in the semantic system. The meanings of later-acquired concepts are constructed from those of earlier-acquired concepts (see also van Loon-Vervoorn, 1989).

A major absence in the current literature is a clear demonstration of any AoA effect in object processing that lies between object recognition (e.g., Vitkovitch & Tyrrell, 1995) and object naming (e.g., Morrison et al., 1992). Indeed, Morrison et al. found no effect of AoA in a categorization task that required objects to be judged as *natural* or *artificial*, even though there was an effect of AoA when the same objects were named. This failure to find an AoA effect may have arisen as an artifact of their picture stimuli and/or from the nature of the categorization task used, since it was possible that some participants were able to perform the task by assessing pictorial attributes of the presented objects that were confounded with the decision required (e.g., more straight lines could indicate that the object was more likely to be man-made than natural); such a response strategy would effectively bypass the need to process the objects semantically. However, a strong demonstration of the influence of AoA on a categorization task would further weaken any explanation that located its effect solely at the stage of the accessing of names (i.e., the phonological completeness hypothesis).

A difficulty for the selection of an appropriate categorization task is the large number of variables that need to be controlled. It is unlikely that this could be achieved in a way that allowed the comparison of the superordinate categories typically used in picture-processing studies (e.g., comparing fruit with vegetables or tools with musical instruments). Previous studies of object processing have employed a categorization task in which participants decide whether presented stimuli are more usually seen inside a house or outside of a house (e.g., Bruce, Carson, Burton, & Ellis, 2000). This task allows the use of a flexible range of picture stimuli and cannot be based reliably on visual or structural features (such as curves and straight lines) alone, since both natural and man-made objects can be found in both settings. In Experiment 1A, this conventionally found *inside the house versus outside the house* categorization task was used to examine the role of AoA in semantic processing. In Experiment 1B, the same objects were presented to a separate group of participants whose task was simply to name the items. It was predicted that AoA would have an effect on both categorization and naming, such that early-acquired items would be responded to more quickly than later-acquired items. There was no strong prediction for the effect of category decision (found inside or outside of the house). On the basis of ratings that we have collected, we would expect the difficulty of making both classifications to be equivalent. However, participants may decide to transform the inside-or-outside task into one of inside (or outside) or not. This may make it possible to perform one of the decisions more quickly than the other. However, this has no theoretical implications for our investigation of an AoA effect and will not be pursued further. In Experiment 2A, a different categorization task that also required access to stored semantic representations was used. Here, participants had to decide whether the object depicted would be smaller or larger than a loaf of bread in the real world. This is a judgment that can be made reliably only by accessing semantics and also, importantly, does not require that the name be accessed at all. In Experiment 2B, the same objects were presented for naming. As before, we predicted an advantage for the early-acquired over the late-acquired items. In both categorization tasks, the judgments that the participants had to make could be seen to be not objectively or completely determined; they relied on what was conventional, typical, or most likely, rather than on what was necessarily the case. For example, although a *tiger* is virtually never seen inside a house, a *caterpillar* may be more difficult to judge (and similarly, whereas it is relatively straightforward to judge the relative size of a *pin* or a *mountain*, it may be more difficult to decide whether a *rabbit* is larger or smaller than a *loaf of bread*). In order to attempt to overcome possible difficulties arising from such potential ambiguity, an initial rating study was conducted

to allow the construction of sets of early and late items that were matched in terms of both their likelihood of being seen inside or outside a house and their size relative to a loaf of bread.

EXPERIMENT 1A

Method

Participants. The participants were 20 undergraduate students (between 18 and 21 years of age) at the University of Birmingham, who either were paid or received course credit for their participation. All were monolingual English speakers with normal or corrected-to-normal vision.

Materials. The critical stimuli were 48 pictures taken from Snodgrass and Vanderwart (1980). The stimuli were constructed on the basis of a rating study in which a subset of the Snodgrass and Vanderwart items were shown to 18 judges who were asked to rate (on a 7-point scale) whether they were more likely to be seen inside a house or outside the house (1 = *most likely seen inside*, 7 = *most likely seen outside*). On the basis of these ratings and the AoA ratings provided by Barry et al. (1997), four groups of items were created: 12 objects rated as being likely to be seen inside the house with early-acquired names (e.g., *comb*, *spider*), 12 objects likely to be seen inside the house with later-acquired names (e.g., *lamp*, *pepper*), 12 objects rated as being likely to be seen outside the house with early-acquired names (e.g., *tiger*, *snowman*), and 12 items likely to be seen outside the house with later-acquired names (e.g., *flag*, *deer*). Table 1 shows the properties of the early and late items, and Appendix A lists the names of the objects used.

The early and the late sets of items were matched across the lists for the word frequency of their names, using frequency counts taken from both Celex (Centre for Lexical Information, 1993) and Kučera and Francis (1967), the rated familiarity and percentage of name agreement of the objects (Barry et al., 1997), image agreement and visual complexity ratings (Snodgrass & Vanderwart, 1980), and word length (in terms of the number of phonemes in the names). The name frequency, agreement, and length measures were established for the names of the objects that were most commonly produced by British speakers, using the normative data provided by Barry et al. (and so, for example, the measures were for *spanner* rather than *wrench* and for *tortoise* rather than *turtle*). The early and the late sets were also matched on the ratings of where they were likely to be seen, so that the ease of making the inside–outside decision should be equivalent for the early and the late items.

Design. We employed a 2 (AoA: early vs. late) × 2 (category type: inside house vs. outside house) design. The dependent variable was the time required to make a correct classification decision.

Procedure. All the participants were tested individually. Each participant was seated at a comfortable distance in front of a Mac-

intosh computer. The pictures were high-contrast line drawings that were centered on the computer screen. Presentation of a picture started a millisecond timer that was terminated by the pressing of a computer key labeled "I" for inside and "O" for outside. These keys were pressed by the participant's index and middle fingers, and the allocation of finger to response was counterbalanced across participants. The participants were asked to respond to each item as quickly and accurately as possible. Responses stopped the timer and terminated the presentation of the picture. Response latencies were recorded to the nearest millisecond. There was a 1-sec intertrial interval. The participants were presented with a randomized order of the 48 experimental pictures, which was preceded by a short practice session involving 8 pictures that did not appear in the main experiment.

Results

Response times (RTs) were excluded if they were further than three standard deviations from the mean. Responses were treated as errors if they did not agree with the majority decision provided by the initial judges in the rating exercise. The mean times to decide correctly that an item was most likely to be seen either inside or outside of a house are shown in Table 2 (with percentages of errors). Overall, 84 (out of 960) responses were excluded. On 73 occasions (7.6%), these arose from an incorrect keypress, and on 11 occasions (1.1%), they were outliers.

Mean times were calculated both by subjects, in each condition, and by items, and separate two-factor ANOVAs were conducted by subjects (F_s) and by items (F_i) on the mean correct categorization times. Each ANOVA examined the variables of AoA (early vs. late) and category type (inside house vs. outside house). In the analyses by subjects, both variables were within-subjects factors (F_s), whereas in the analysis by items, both were between-items factors (F_i). We adopted an alpha level of .05 for all our tests of significance.

There was a significant main effect of AoA [$F_s(1,19) = 25.8$, $MS_e = 5,799$; $F_i(1,44) = 9.4$, $MS_e = 9,720$]. More time was required to make the inside–outside judgment about pictures that were labeled with a late-acquired word, as compared with pictures labeled with an early acquired word. The main effect of category type was also significant [$F_s(1,19) = 13.5$, $MS_e = 9,623$; $F_i(1,44) = 7.3$, $MS_e = 9,720$]. The interaction between AoA and category type was not significant.

Table 1
Characteristics of the Stimulus Items Used in Experiments 1A and 1B

| | Early Acquired | | | | Late Acquired | | | |
| | Inside | | Outside | | Inside | | Outside | |
Characteristic	M	SD	M	SD	M	SD	M	SD
AoA ratings	2.15	0.22	2.12	0.21	3.21	0.46	3.24	0.50
Celex total frequency	26.2	37.9	18.6	24.4	16.3	24.7	8.3	12.5
Kučera and Francis frequency	17.5	22.5	22.2	22.4	14.7	18.2	8.8	12.4
Object familiarity	3.55	0.68	2.66	0.91	3.21	0.69	2.31	0.52
Percentage of name agreement	91.7	14.6	90.8	12.90	87.3	15.5	84.0	16.9
Visual complexity	2.48	0.56	3.6	0.80	2.53	0.83	3.66	0.68
Image agreement	3.73	0.73	3.8	0.30	3.93	0.67	3.59	0.56
No. of phonemes	5.42	2.00	4.25	1.40	5.75	2.00	5.50	1.60
Inside–outside house ratings	2.12	0.93	6.7	0.33	2.27	0.64	6.52	0.63

The analysis of the errors showed only a main effect of category type [$F_s(1,19) = 13.4$, $MS_e = 1.81$; $F_i(1,44) = 12.7$, $MS_e = 3.17$]; more errors were made when it was decided that items were more likely to be encountered inside a house than when the *outside* decision was made.

EXPERIMENT 1B

Experiment 1A showed an advantage for objects with early-acquired names in the inside–outside categorization task. It was also found that participants were able to make an *inside* decision more rapidly than an *outside* one. However, since this had no theoretical interest and did not interact with the effect of AoA, it was not pursued further. Experiment 1B was conducted in order to establish that the items selected for that task also showed an AoA effect in object naming.

Method

Participants. Twenty undergraduate students at the University of Birmingham (between 18 and 21 years of age) received course credit for their participation. All were monolingual English speakers, with normal or corrected-to-normal vision, and none had taken part in Experiment 1A.

Design. We employed a 2 (AoA: early vs. late) × 2 (category type: inside house vs. outside house) design. The dependent variable was the time required to name each item correctly.

Procedure. The same 48 objects as those used in Experiment 1A were presented for object naming. The participants were asked to respond to each picture by saying its name and were instructed to say the name clearly and quickly and to avoid saying "um" or "er." Presentation of a picture started a millisecond timer that was terminated by the initiation of the oral naming response made into a high-sensitivity voice key microphone attached to the computer. Prior to the main experiment, the participants completed a practice session by naming the same practice items as those used in Experiment 1A.

Results

Naming times were excluded if they were further than three standard deviations from the mean. Responses were treated as errors if they did not coincide with the anticipated name or if they involved some verbal dysfluency or other noise that stopped the timer. Table 2 shows the mean correct naming times in each condition (with percentages of errors). Overall, 115 responses were excluded. On 107 occasions (11.1%), these arose from an incorrect response, and on 8 occasions (0.8%), they were outliers.

Separate two-factor ANOVAs were conducted by participants and by items on the mean correct naming times. Each ANOVA examined the variables of AoA (early vs. late) and house (inside vs. outside). In the analyses by participants, both variables were within-subjects factors, whereas in the analysis by items, both were between-items factors. There was a significant main effect of AoA [$F_s(1,19) = 27.8$, $MS_e = 4,695$; $F_i(1,44) = 7.02$, $MS_e = 12,650$] but no significant main effect of category type [$F_s < 1$; $F_i < 1$]. The AoA × category type interaction was not significant.

The analysis of the errors revealed only a significant main effect of AoA [$F_s(1,19) = 26.8$, $MS_e = 0.64$;

Table 2
Mean Correct Response Times (RTs, in Milliseconds) to Categorize Where Items Are Most Likely Encountered (Experiment 1A) and to Name the Same Items (Experiment 1B), and Percentages of Errors

Item Condition	Experiment 1A (Categorization)		Experiment 1B (Naming)	
	RT	E%	RT	E%
Early inside	821	12.5	919	7.5
Early outside	748	3.3	873	8.8
Late inside	915	14.2	976	13.8
Late outside	827	5.0	978	17.9

$F_i(1,44) = 5.9$, $MS_e = 4.9$]. Fewer errors were made to the early items than to the late ones.

Discussion

Experiment 1A showed that AoA had an effect on the time taken to categorize objects as being conventionally found either inside or outside a house. Experiment 1B showed that AoA also affected the time taken to name the objects, a result that replicated a number of other studies (e.g., Barry et al., 2001; Carroll & White, 1973; Morrison et al., 1992).

The finding of principal interest was that objects whose names were early-acquired words were categorized more quickly than objects labeled with later-acquired words. We believe that this is the first time an AoA effect has been demonstrated for the semantic processing of pictures of objects. However, it is completely in accordance with the explanation of AoA suggested by Ellis and Lambon Ralph (2000). In order to decide whether an item is more likely to be seen either inside or outside of a house, participants must access (and process) their stored semantic representations of the pictured objects. Furthermore, this categorization task is one that does not require participants to retrieve the names of the objects, and the fact that the naming latencies observed in Experiment 1B were substantially longer (by an average of 109 msec) than the categorization times found in Experiment 1A suggests that the participants did not generate the names of objects before categorizing them. Indeed, an analysis of the RTs in Experiments 1A and 1B revealed a significant main effect of experimental task [$F_s(1,38) = 5.27$; $MS_e = 90,346$].

The semantic categorization task used in Experiment 1A differs from typical superordinate classification tasks (such as categorizing items as being *animals* or *items of furniture*). Indeed, it is more similar to what Barsalou (1983) called *ad hoc* (or *goal-directed*) categories, such as *items of food to avoid when on a slimming diet*. However, we maintain that, in order to perform categorization, it is necessary for people to activate and consult their detailed semantic (and conceptual) knowledge of the objects presented. The categorization task is one that has been employed before (e.g., Bruce et al., 2000), and it is also one that participants can perform readily (with mean RTs of 828 msec and with high levels of accuracy) and more quickly (by 109 msec) than the time it takes to

overtly name the objects. This categorization task is a very useful for the purpose of examining AoA effects. If the AoA effect is not restricted to the stage of name retrieval, as is proposed by the phonological completeness hypothesis, it would be expected to appear in any semantic categorization task. However, this has not always been the case. Morrison et al. (1992) failed to show an AoA effect on the categorizing of pictures as being *natural* or *artificial*, even though they observed a reliable AoA effect on naming the same items in the same laboratory by the same participant population. In our laboratory, we have attempted to replicate Morrison et al.'s study (using their items), and we too failed to find any AoA effect in the *natural versus artificial* classification task. In order to test the generality of the AoA effect in semantic processing found in Experiment 1, a second experiment was designed, employing another categorization task, to attempt to show that this effect is not restricted to a single task. In Experiment 2A, the participants were asked to decide whether a presented object was, in real life, smaller or larger than a loaf of bread, a task that also required access to stored world knowledge. In Experiment 2B, the same pictures were presented to the participants for object naming.

EXPERIMENT 2A

Method

Participants. The participants were 20 undergraduate students at the University of Birmingham (between 18 and 21 years of age), none of whom had participated in Experiments 1A or 1B. All were monolingual English speakers, with normal or corrected-to-normal vision, and either were paid or received course credit for their participation.

Materials. The critical stimuli were 56 pictures taken from Snodgrass and Vanderwart (1980). As in Experiment 1A, the stimuli were constructed from the results of an initial rating study in which a subset of the Snodgrass and Vanderwart items were shown to 18 judges who were asked to judge how the size of the object compared with the size of a typical loaf of bread (1 = *much bigger*, 7 = *much smaller*). On the basis of these ratings and the AoA ratings provided by Barry et al. (1997), four groups of items were created: 14 objects considered smaller than a loaf with early-acquired names (e.g., *tomato*, *pear*), 14 objects smaller than a loaf with later-acquired names (e.g., *ruler*, *screw*), 14 objects considered larger than a loaf with early-acquired names (e.g., *fox*, *pig*), and 14 objects larger than a loaf with later-acquired names (e.g., *deer*, *harp*). Table 3 shows

the properties of the early and late items, and Appendix B lists the names of the objects used.

The early and late sets of items were matched on the frequency, length, and name agreement of their names and on the familiarity, image agreement, and visual complexity ratings of the objects (using the same sources as those for the items selected for Experiment 1A). The early and late items were also matched on ratings of how they differed in size from a typical loaf of bread, so that the ease of making the *smaller–bigger* categorization should be equivalent for both sets.

Design. We employed a 2 (AoA: early vs. late) × 2 (category type: bigger than a loaf vs. smaller than a loaf) design. The dependent variable was the time required to make a correct classification decision.

Procedure. The procedure was identical to that employed in Experiment 1A, except that the keys were now labeled "B" and "S."

Results

RTs were excluded if they were further than three standard deviations from the mean. Responses were treated as errors if they did not coincide with the decision given by the initial judges in the rating exercise. The mean times to decide correctly that an item was smaller or bigger than a typical loaf of bread are shown in Table 4 (with percentages of errors). Overall, 86 (out of 1,120) responses were excluded. On 67 occasions (6.0%); these arose from an incorrect keypress, and on 19 occasions (1.7%), they were outliers.

Separate two-factor AoA (early vs. late) × category type (smaller vs. bigger) ANOVAs were conducted on the mean correct categorization times. In the analyses by participants, both variables were within-subjects factors, whereas in the analysis by items, both were between-items factors. There was a significant main effect of AoA $[F_s(1,19) = 10.12, MS_e = 6,467; F_i(1,52) = 8.2, MS_e = 7,354]$. The participants made the smaller–bigger judgment to pictures with early-acquired labels more quickly than to pictures labeled with late-acquired words. The main effect of category type was also significant $[F_s(1,19) = 3.25, MS_e = 14,631, p = .09; F_i(1,52) = 6.4, MS_e = 7,354]$. These main effects were modified by an AoA × category type interaction that was significant only by participants $[F_s(1,19) = 6.18, MS_e = 2,933]$. An analysis of the simple main effects showed that the AoA effect was significant only for the *small* decisions $[F_s(1,19) = 11.8, MS_e = 6,467]$, and not for the *big* decisions.

Table 3
Characteristics of the Stimulus Items Used in Experiments 2A and 2B

| | Early Acquired | | | | Late Acquired | | | |
| | Smaller | | Bigger | | Smaller | | Bigger | |
Characteristic	M	SD	M	SD	M	SD	M	SD
AoA ratings	2.33	0.28	2.22	0.24	3.19	0.46	3.22	0.41
Celex total frequency	14.3	26.20	25.30	35.00	8.57	9.30	14.10	22.6
Kučera and Francis frequency	8.29	8.90	26.90	6.50	10.57	10.70	17.20	18.40
Object familiarity	2.89	0.44	2.92	1.10	2.84	0.73	2.31	0.69
Percentage of name agreement	91.90	12.50	86.90	17.40	86.10	15.10	90.2	10.00
Visual complexity	2.62	0.91	3.30	0.96	2.96	0.73	3.49	0.71
Image agreement	3.87	0.59	3.67	0.53	3.76	0.64	3.84	0.47
No. of phonemes	4.36	2.00	4.43	1.30	4.86	1.40	4.64	1.60
Bigger–smaller than loaf ratings	6.18	0.40	1.68	0.44	5.86	0.95	1.84	0.63

Table 4
Mean Correct Response Times (RTs, in Milliseconds) to
Categorize the Size of Items (Experiment 2A) and to Name the
Same Items (Experiment 2B), and Percentages of Errors

Item	Experiment 2A (Categorization)		Experiment 2B (Naming)	
Condition	RT	%E	RT	%E
Early small	681	5.7	830	6.8
Early big	662	3.6	825	4.6
Late small	768	14.6	946	13.6
Late big	689	6.8	908	7.1

The analysis of the errors showed a significant main effect of AoA [$F_s(1,19) = 16.59$, $MS_e = 0.87$; $F_i(1,52) = 5.46$, $MS_e = 3.78$]. The participants tended to make fewer errors to early than to late items (0.65 and 1.5, respectively). The main effect of category type was also significant [$F_s(1,19) = 11.49$, $MS_e = 0.85$; $F_i(1,52) = 3.7$, $MS_e = 3.78$], showing that the participants made more errors to the smaller than to the bigger items (1.43 and 0.73, respectively).

EXPERIMENT 2B

Method

Participants. Twenty undergraduate students at the University of Birmingham (between 18 and 21 years of age), none of whom had participated in Experiment 1A, 1B, or 2A, either were paid or received course credit for their participation. All were monolingual English speakers with normal or corrected-to-normal vision.

Design. We employed a 2 (AoA: early vs. late) × 2 (category type: bigger than a loaf vs. smaller than a loaf) design. The dependent variable was the time required to name each item correctly.

Procedure. The same 56 pictures as those used in Experiment 2A were presented for object naming, using the same procedure as that in Experiment 1B.

Results

Naming times were excluded if they were further than three standard deviations from the mean. Responses were treated as errors if they did not coincide with the anticipated name or if they involved some verbal dysfluency or other noise that stopped the timer. Table 4 shows the mean correct naming times in each condition (with percentages of errors). Overall, 90 responses were excluded. On 76 occasions (6.8%), these arose from an incorrect response, and on 14 occasions (1.25%), they were outliers.

Separate two-factor AoA (early vs. late) × category type (smaller vs. bigger) ANOVAs were conducted on the mean correct naming times. In the analyses by participants, both variables were within-subjects factors, whereas in the analysis by items, both were between-items factors. There was a significant main effect of AoA [$F_s(1,19) = 36.9$, $MS_e = 5,361$; $F_i(1,52) = 15.66$, $MS_e = 11,110$], but the main effect of category type was not significant.

The analysis of the errors showed a significant main effect of AoA [$F_s(1,19) = 8.9$, $MS_e = 0.95$; $F_i(1,52) = 4.3$, $MS_e = 2.8$]. The participants tended to make fewer errors to early than to late items (0.8 and 1.45, respec-

tively). The main effect of category type was also significant [$F_s(1,19) = 9.57$, $MS_e = 0.75$; $F_i(1,52) = 3.7$, $MS_e = 2.8$; $p = .06$], showing that the participants made more errors to the smaller than to the bigger items (1.43 and 0.83, respectively).

Discussion

The results of Experiments 2A and 2B confirm those of Experiments 1A and 1B: AoA has a reliable effect on the time required to semantically categorize and name objects, with a clear advantage found for objects labeled with early-acquired words. The semantic categorization task used in Experiment 2A—namely, to decide whether an object was smaller or larger than a loaf of bread—was one that required the participants to access their stored semantic representations corresponding to the pictured objects, but it did not require them to retrieve the names of the objects. As can be seen from the results presented in Table 4, the naming latencies for these objects were substantially longer than the categorization times. An analysis of the RTs in Experiments 2A and 2B confirmed a significant main effect of experimental task [$F_s(1,38) = 15.3$, $MS_e = 82,214$], with naming latencies being 177 msec longer than categorization times.

In Experiments 1A and 2A, we have shown that AoA influences the time to categorize pictures of objects. In Experiment 1A, this was significant for both parts of the inside or outside of a house decision. In Experiment 2, AoA was significant only when the participants were making a *smaller* decision, however, the pattern of the RTs was in a direction that favored decisions to early items. The mean time required to make a *bigger* decision to late items was 27 msec longer than that required to respond to early ones.

GENERAL DISCUSSION

Two experiments have shown that the rated age at which the name of a picture is first learned has a reliable impact on the time to categorize the picture. In Experiment 1A, the participants made a judgment about whether objects would be more likely to be found inside or outside a house, and in Experiment 2A, the participants judged whether objects were smaller or larger than a typical loaf of bread. The sets of stimuli making up the conditions in both experiments were matched on other variables likely to affect picture identification (see Tables 1 and 3). Furthermore, and importantly, the objects used had been matched on ratings from an independent exercise in which judges (who did not participate in the experimental part of the study) rated how likely items were to be found either inside or outside of a house or how their size compared with that of a typical loaf of bread. Clear and consistent AoA effects were found in all these semantic categorization tasks. Two further experiments (1B and 2B) also showed that objects labeled with early-acquired words were named reliably more quickly than those with late-acquired names, which confirms the results of a growing body of experimental investigations

(e.g., Barry et al., 2001; Carroll & White, 1973; Gilhooly & Gilhooly, 1979; Morrison et al., 1992; Snodgrass & Yuditsky, 1996).

Our demonstration of an AoA effect in the semantic categorization of pictured objects joins an increasing number of studies that have shown an influence of AoA in picture-, word-, and face-processing tasks that do not require any access to the phonology of the target item (e.g., Brysbaert et al., 2000; Lewis, 1999; Moore & Valentine, 1999; Vitkovitch & Tyrrell, 1995), and even in studies in which word recognition memory has been examined (Dewhurst, Hitch, & Barry, 1998). Thus, it becomes increasingly difficult to explain the impact of AoA entirely in terms of a mechanism involving access to names, as has been suggested by Brown and Watson's (1987) phonological completeness hypothesis.

A number of connectionist accounts of AoA have been offered recently. For word reading, a version of the cumulative frequency account has been advanced by Zevin and Seidenberg (2002). In their connectionist-modeling work with word reading, Zevin and Seidenberg's simulations showed that cumulative frequency had long-lasting effects on the network performance, so that high-frequency items were better encoded than low-frequency items. Zevin and Seidenberg found that order of training had an effect in their simulations only when what was learned about early patterns did not carry over to later ones, which, they argued, is not a characteristic of learning spelling–sound mappings but is relevant in learning the names of objects and faces. In their modeling work, Ellis and Lambon Ralph (2000) proposed that AoA effects arise from a gradual reduction of network plasticity, which causes later items to be differentiated less effectively than earlier ones. If this is correct, we might easily expect to uncover AoA effects in any task in which learning has been cumulative and in which the input-to-output mappings are arbitrary. In the categorization tasks employed in this study, no access to phonology was required. Our findings from Experiments 1A, and 2A are not easy to reconcile with a phonological completeness hypothesis, but they are entirely compatible with these connectionist accounts of AoA.

Similarly, an explanation of AoA effects that is based on the structure of the semantic system could also accommodate our present findings. Brysbaert, Ghyselinck, and Storms (2002) have suggested that semantic categories may be defined to a greater extent by earlier-acquired concepts than by later-acquired ones. Steyvers and Tenenbaum (2005) have presented a statistical analysis of semantic networks that describes their development. Nodes that are formed in a network at an early stage in its development typically possess larger numbers of connections than do nodes that are formed later. This account differs from the model proposed by Ellis and Lambon Ralph (2000), because it does not suggest that later-acquired words are encoded any less effectively than early-acquired words; instead, the processing advantages seen for early-acquired words arise because they are more richly connected to the network.

Throughout this article, we have referred to our object stimuli as those labeled by early- or late-acquired *words*. This is because we have reliable (and valid) norms only for the AoA of words, and not for the objects themselves (or for the concepts represented by the pictures we have used). Although it is eminently reasonable to assume that the age at which object names are acquired also reflects the age at which the objects are encountered and assimilated into secure semantic representations, it will be a matter for future research to determine whether the effects we have reported result from the AoA of the object names or from the objects themselves. However, some studies that have employed multiple measures of AoA have shown both to be significant (e.g., Yamazaki, Ellis, Morrison, & Lambon Ralph, 1997).

Finally, it needs to be noted that all the categorization tasks employed in this study were of a particular type. Bruce et al. (2000) employed a similar *inside–outside of the house* task and remarked on how this is not a "natural" category but, rather, one that must be computed "on line." Indeed, they go on to say that "whether an object is normally found inside the house is a question which can be readily answered, but probably not one that structures semantic space" (p. 312). We would suggest that the decisions concerning the real-world size of an object (in relation to that of a typical loaf of bread) must also fall squarely within this class of tasks. This raises an interesting question about the failure to demonstrate an AoA effect when the task was to distinguish between natural and man-made objects (Morrison et al., 1992). On first consideration, this may also appear to be an on-line category, but this is not straightforward, since the distinction between living things and artifacts has been observed in cognitive neuropsychological double dissociations (e.g., Saffran & Schwartz, 1994). At present, the perceptual discontinuities between natural (typically, living things) and man-made object categories, which may allow categorization at a very superficial level, remain our preferred explanation for the lack of an AoA effect in Morrison et al.'s study. The challenge for future work will be to show that AoA effects can be found using tasks that require access to the sort of categories that are unequivocally considered to structure our semantic memory.

REFERENCES

Barry, C., Hirsh, K. W., Johnston, R. A., & Williams, C. L. (2001). Age of acquisition, word frequency, and the locus of repetition priming of picture naming. *Journal of Memory & Language*, **44**, 350-375.

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.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, **11**, 211-227.

Bonin, P., Fayol, M., & Chalard, M. (2001). Age of acquisition and word frequency in written picture naming. *Quarterly Journal of Experimental Psychology*, **54A**, 469-489.

Brown, G. D. A., & 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 & Cognition*, **15**, 208-216.

Bruce, V., Carson, D., Burton, A. M., & Ellis, A. W. (2000). Per-

ceptual priming is not a necessary consequence of semantic classification of pictures. *Quarterly Journal of Experimental Psychology*, **53A**, 289-324.

BRYSBAERT, M., GHYSELINCK, M., & STORMS, G. (2002, April). *Effects of age of acquisition in semantic categorization tasks*. Paper presented at the joint meeting of EPS and BVP/SBP, Leuven, Belgium.

BRYSBAERT, M., VAN WIJNENDAELE, I., & DE DEYNE, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica*, **104**, 215-226.

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.

CENTRE FOR LEXICAL INFORMATION (1993). *The Celex lexical database*. Nijmegen, The Netherlands: Max Planck Institute for Psycholinguistics.

DEWHURST, S. A., HITCH, G. J., & BARRY, C. (1998). Separate effects of word frequency and age of acquisition in recognition and recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 284-298.

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, & Cognition*, **26**, 1103-1123.

GERHAND, S., & BARRY, C. (1998). Word frequency effects in oral reading are not merely age-of-acquisition effects in disguise. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 267-283.

GERHAND, S., & BARRY, C. (1999). Age of acquisition, word frequency, and the role of phonology in the lexical decision task. *Memory & Cognition*, **27**, 592-602.

GILHOOLY, K. J., & GILHOOLY, M. L. M. (1979). Age-of-acquisition effects in lexical and episodic memory tasks. *Memory & Cognition*, **7**, 214-223.

KUČERA, H., & FRANCIS, W. H. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

LACHMAN, R., SHAFFER, J. P., & HENNRIKUS, 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.

LEWIS, M. B. (1999). Age of acquisition in face categorisation: Is there an instance-based account? *Cognition*, **71**, B23-B39.

MCCLELLAND, J. L., MCNAUGHTON, B. L., & O'REILLY, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, **102**, 419-457.

MESCHYAN, G., & HERNANDEZ, A. (2002). Age of acquisition and word frequency: Determinants of object-naming speed and accuracy. *Memory & Cognition*, **30**, 262-269.

MONAGHAN, J., & ELLIS, A. W. (2002a). Age of acquisition and the completeness of phonological representations. *Reading & Writing*, **15**, 759-788.

MONAGHAN, J., & ELLIS, A. W. (2002b). What exactly interacts with spelling–sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 183-206.

MOORE, V., & VALENTINE, T. (1999). The effects of age of acquisition in processing famous faces and names: Exploring the locus and proposing a mechanism. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the Twenty-First Annual Meeting of the Cognitive Science Society* (pp. 416-420). Mahwah, NJ: Erlbaum.

MORRISON, C. M., & ELLIS, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 116-133.

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.

SAFFRAN, E. M., & SCHWARTZ, M. F. (1994). Of cabbages and things: Semantic memory from a neuropsychological perspective. A tutorial review. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 507-536). Cambridge, MA: MIT Press, Bradford Books.

SNODGRASS, J. G., & 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.

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

VAN LOON-VERVOORN, W. A. (1989). *Eigenschappen van basiswoorden*. Lisse: Swets & Zeitlinger.

VITKOVITCH, M., & TYRRELL, L. (1995). Sources of disagreement in object naming. *Quarterly Journal of Experimental Psychology*, **48A**, 822-848.

YAMAZAKI, M., ELLIS, A. W., MORRISON, C. M., & LAMBON RALPH, M. A. (1997). Two age of acquisition effects in the reading of Japanese Kanji. *British Journal of Psychology*, **88**, 407-421.

ZEVIN, J. D., & SEIDENBERG, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory & Language*, **47**, 1-29.

APPENDIX A
Names of the Objects Presented in Experiments 1A and 1B

Early Acquired		Late Acquired	
Inside	Outside	Inside	Outside
banana	bus	broom	barrel
button	chicken	cigarette	barn
coat	donkey	desk	caterpillar
comb	duck	envelope	lobster
dress	elephant	jug	motorbike
drum	giraffe	needle	mountain
paintbrush	leaf	pepper	mushroom
stool	snake	pineapple	peacock
strawberry	snowman	screwdriver	penguin
television	tiger	suitcase	squirrel
tomato	tree	toaster	tortoise
toothbrush	wheel	watermelon	yacht

APPENDIX B
Names of the Objects Presented in Experiments 2A and 2B

Early Acquired		Late Acquired	
Smaller	Bigger	Smaller	Bigger
ant	bicycle	belt	anchor
bird	coat	caterpillar	barrel
butterfly	airplane	cherry	broom
frog	drum	cigar	deer
comb	giraffe	thimble	desk
leaf	stool	doorknob	guitar
lemon	snowman	glasses	harp
nail	swan	mushroom	camel
necklace	sweater	lettuce	peacock
paintbrush	television	pepper	penguin
pear	tiger	pineapple	piano
pipe	train	pliers	suitcase
strawberry	wheel	screw	tortoise
whistle	zebra	tie	mountain