Journal of Experimental Psychology:
Learning, Memory, and Cognition
2004, Vol. 30, No. 2, 550–554

0278-7393/04/$12.00 DOI: 10.1037/0278-7393.30.2.550

The Effect of Age of Acquisition in Visual Word Processing: Further Evidence for the Semantic Hypothesis

Mandy Ghyselinck and Roel Custers
Ghent University

Marc Brysbaert
Ghent University and University of London

The authors investigated whether the meaning of visually presented words is activated faster for early-acquired words than for late-acquired words. They addressed the issue using the semantic Simon paradigm. In this paradigm, participants are instructed to decide whether a stimulus word is printed in uppercase or lowercase letters. However, they have to respond with a verbal label ("living" or "nonliving") that is either congruent with the meaning of the word (e.g., saying "living" to the stimulus *DOG*) or incongruent (e.g., saying "nonliving" to the stimulus *dog*). Results showed a significant congruency effect that was stronger for early-acquired words than for late-acquired words. The authors conclude that the age of acquisition is an important variable in the activation of the meaning of visually presented words.

In the past decade, researchers have revived the question regarding the extent to which the frequency effect in visual word recognition is a confound of the age at which words have been acquired (hence called age of acquisition [AoA]). Several hypotheses have been advanced to explain both the origin of the AoA effect and its relationship to the frequency effect (see Chalard, Bonin, Meot, Boyer, & Fayol, 2003; Lewis, Chadwick, & Ellis, 2002; Morrison, Hirsh, & Duggan, 2003).

One hypothesis is that part of the AoA effect originates from the semantic system. According to this explanation, the order of acquisition has a lasting effect on the time needed to activate the meanings of words. Empirical evidence for this idea comes from the word-associate generation task (Brysbaert, Van Wijnendaele, & De Deyne, 2000; van Loon-Vervoorn, 1989). In this task, participants have to say the first word that comes to mind on seeing a word. Participants are much faster to generate an associate to early-acquired words than to later-acquired words. Interestingly, there is no analogue frequency effect in the word-associate generation task when stimuli are controlled for AoA.

Theoretical support for a semantic involvement in the AoA effect comes from simulations with models based on both distributed and localist representations. The distributed account attributes the AoA effect to differences in the connection weights between the units of the orthographic and the semantic layers; the localist account attributes it to the organization of the semantic system.

Mandy Ghyselinck and Roel Custers, Department of Experimental Psychology, Ghent University, Ghent, Belgium; Marc Brysbaert, Ghent University, Ghent, Belgium, and Royal Holloway, University of London, London, United Kingdom.

Roel Custers is now at the Department of Experimental Clinical and Health Psychology, Ghent University.

This research was made possible by a Research Council of Ghent University Bijzonder Onderzoeksfonds project grant.

Correspondence concerning this article should be addressed to Mandy Ghyselinck, Department of Experimental Psychology, Henri Dunantlaan 2, B-9000 Ghent, Belgium. E-mail: mandy.ghyselinck@UGent.be

Three-layer neural network models with distributed representations show an advantage for early-trained items if the network is trained in such a way that the early stimuli continue to be presented when the later stimuli are introduced (Ellis & Lambon Ralph, 2000). This is because a neural system loses plasticity in the learning process. When the network is young, connection weights between the different layers are distributed around the mean of 0.5, and stimuli can cause large shifts in the weights. As the network gets older, the weight shifts tend to become smaller because the connection strengths are already close to one of the extremes (either 0.0 or 1.0). Therefore, the weight shifts induced by later-acquired words will never be as substantial as those induced by early-learned ones. As a consequence, the words that are learned early in training will be more influential for the final structure of the network. This advantage can survive huge differences in cumulative frequency.

Zevin and Seidenberg (2002) more or less reached the same conclusion, but they emphasized much more that the emergence of an AoA effect depends on the sort of task that has to be performed. The acquisition order is particularly important when the mapping between input and output is arbitrary (i.e., when no generalization of early-trained patterns to later-trained patterns is possible). Otherwise, the regularities learned for early-acquired patterns can be transferred to later-acquired patterns. Specifically with respect to visual word recognition, Zevin and Seidenberg (2002) argued that the mapping of orthography to phonology in English is not arbitrary enough to give rise to an AoA effect in visual word naming (because many onsets and rhymes of words are consistent between early-learned and late-learned words; e.g., the rhymes of *CAT* and *SPAT*). The fact that AoA nevertheless affects word-naming latencies in English has been explained by Zevin and Seidenberg by referring to the fact that some naming latencies are semantically mediated (in particular, those of words with inconsistent spelling-to-sound mappings). Because there are very few regularities in the mappings from spelling to meaning and from meaning to sound (words that are written similarly rarely have related meanings), AoA is expected to play a significant role in tasks that require the activation of meaning.

Other authors have also pointed to the semantic system as a possible origin of the AoA effect in visual word recognition, but these authors were thinking more in terms of the organization of the semantic system rather than the weights of the connections to and from the system. This is because these authors worked within the framework of localist models, which postulate a single node for each meaningful unit (words, concepts, semantic features) rather than the distributed representations on which neural networks are based (see, e.g., Bowers, 2002; for a discussion of localist versus distributed representations in visual word processing). For example, Steyvers and Tenenbaum (2003) presented a mathematical model that simulated the organization of a growing semantic network. The network consists of interconnected nodes that represent concepts, and it develops according to a principle Steyvers and Tenenbaum previously observed in semantic structures. Basically, the principle implies that new concepts are added to the network by connecting them to existing nodes (concepts) as a function of the number of connections each node already has. This preferential-attachment principle makes the prediction that early-acquired nodes have a more central position in the network because, on average, they have more connections than later-acquired nodes (which are attached to them).

Despite the fact that the semantic system has been suggested as a major contributor to the AoA effect in visual word recognition, the empirical evidence remains rather weak. To the best of our knowledge,[1] the evidence is limited to the word-associate generation task, which can be questioned because the choice of associations may be based more on co-occurrences of word forms than on the meaning of the words (i.e., the word *cat* is given as the first associate of *dog* not because both nouns refer to animals but because both words often co-occur in discourse).

One reason why so few behavioral data exist may be that it is difficult to find a suitable visual word-processing task. First, the response latencies must not be too long. Otherwise, it can be argued that the AoA effect was not due to the semantic system but, for instance, to the fact that the phonology of the words was activated as part of good task performance (see Morrison & Ellis, 1995, and Gerhand & Barry, 1998, for such an explanation of the AoA effect in the lexical decision task). Second, some methodological issues reduce the chances of finding a reliable AoA effect in semantic categorization tasks. Brysbaert et al. (2000), for instance, argued that results must not be averaged over the two response categories because binary manual decisions are usually translated into a yes–no decision, with different response criteria for the no trials than for the yes trials. Another methodological caveat that has to be taken into account is that the frequency effect (and presumably the AoA effect) is reduced under primed conditions relative to unprimed conditions because the priming effect is stronger for difficult words than for easy words (Becker, 1979). This, for instance, reduces the chances of finding a strong AoA effect in a category verification task, in which a category is presented first (e.g., birds) followed by a target word (e.g., robin, heron, sword, or canoe) of which the participant has to decide whether it is an exemplar of the previously shown category or not.

We believe we have found a way to circumvent these methodological problems. It is based on a semantic variant of the classic Simon paradigm, first reported by De Houwer (1998). In the Simon paradigm, participants are asked to make a spatial response to a nonspatial stimulus characteristic (e.g., press the left key when a red light is shown), while ignoring the location of the stimulus (e.g., to the left or the right of the fixation location). This typically results in faster responses when the stimulus location is congruent with the response code (i.e., a red light presented to the left) than when it is incongruent (red light presented to the right), even though the location of the stimulus is irrelevant for correct task performance. De Houwer (1998) showed that a similar effect is obtained when the irrelevant stimulus property concerns the meaning of the stimulus words. He presented stimuli in uppercase or lowercase letters and asked the participants to say “animal” when the stimulus was presented in uppercase and “human” when the stimulus was presented in lowercase. Responses were faster relative to a neutral condition when the participant’s verbal response was congruent with the meaning of the stimulus (e.g., saying “animal” to the stimulus *CAT*), and they were slower when the participant’s verbal response was incongruent with the meaning of the stimulus (e.g., saying “human” to the stimulus *cat*), even though the meaning of the stimulus was irrelevant for correct task performance.

The congruency effect found with the semantic Simon paradigm can only be due to the automatic activation of the semantic information conveyed by the stimulus word, which interferes with the response label. This opens a nice way to examine the extent to which the activation of semantic information is influenced by AoA. If early-acquired words activate their meaning faster than late-acquired words, either because their orthographic–semantic connections are better (neural network account) or because early-acquired concepts in the semantic system have more connections (localist account), then the congruency effect should be stronger for earlier-acquired words than for later-acquired words. We present the data of an experiment that tested this prediction. In this experiment, participants were asked to say “living” or “nonliving” to words presented in uppercase or lowercase that could refer to either living creatures (e.g., robin, heron) or nonliving entities (e.g., sword, canoe).

Method

Participants

Thirty-six participants volunteered for the experiments (mean age = 22.3 years; range = 18–27). All participants had normal or corrected-to-normal eye vision, and all spoke Dutch as their first language.

Materials

We created four word lists of 22 words each. The words referred to early-acquired living things, late-acquired living things, early-acquired nonliving things, and late-acquired nonliving things. The words were matched on frequency, familiarity,[2] word length, and numbers of syllables.

[1] Note, however, that studies in other domains such as face recognition (Moore & Valentine, 1998) and picture categorization (Johnston & Barry, 2002) have reported AoA effects, which can also be taken as support for the semantic hypothesis.

[2] Zevin and Seidenberg (2002) argued that stimuli must be controlled for subjective familiarity in addition to objective frequency if one wants to interpret an AoA effect as more than a cumulative frequency effect. In a pilot study in which stimuli were not controlled for familiarity, we indeed obtained a stronger AoA effect than the one reported here.

The AoA ratings were taken from Ghyselinck, Custers, and Brysbaert (2003). Frequency measures were based on the Celex database (Baayen, Piepenbrock, & van Rijn, 1993). It is important to note that we chose only words for which each participant in the Ghyselinck et al. study (2003) indicated they knew the meaning. The familiarity ratings were collected by asking 35 undergraduates (mean age = 21.8 years; range = 19–29) to indicate on a 5-point scale for 260 words how often they had heard, seen, or used each word (1 = *never [you have never seen, heard, or used this word before]*, 5 = *very often [you see, hear, or use this word nearly every day]*). The words were presented one by one on a computer screen in a randomized order, and participants typed in their answer on the keyboard. The reliability of the ratings was assessed with the intraclass correlation of Shrout and Fleiss (1979) and amounted to .93. Details of the word lists are shown in Table 1, and the full list of experimental stimuli is given in the Appendix. Half of the stimulus set was presented in lowercase letters and half in uppercase letters, counterbalanced across participants.

Procedure

Participants were tested individually in a quiet room. They were given written instructions on the computer screen in which accuracy and speed were stressed. The task of the participant was to categorize the stimulus word as quickly as possible depending on the letter case and to ignore its actual semantic category. Half of the participants had to say “levend” (“living”) in response to lowercase letters and “levenloos” (“nonliving”) in response to uppercase letters. The other half received the opposite instructions (i.e., lowercase → nonliving, uppercase → living). On each trial several events occurred. First, a central fixation point (“+”) was presented for 500 ms followed by a 500-ms blank interval. Then the stimulus appeared in the white standard MS-DOS letter font in the middle of the screen on a black background. The stimulus stayed on the screen until the voice key registered a response. Successful voice key registration was indicated by a cross that appeared at the bottom of the screen. The experimenter coded the correctness of the response online by means of the keyboard. Stimulus presentation was randomized for each participant. Before the test items, participants received a series of 40 different practice trials (20 of each category). The intertrial interval was 1500 ms.

Results

Only correct reaction times (RTs) were included in the analyses (percentage of errors was less than 1.2%). Harmonic means of the latencies were calculated per condition and per participant (or stimulus word). We used harmonic means rather than arithmetic means following Ratcliff’s (1993) suggestions for appropriate data transformation in analyses of variances (ANOVAs). A 2 (congruency) × 2 (AoA) × 2 (semantic category) ANOVA revealed a significant main effect of congruency, $F_1(1, 35) = 21.17$, $p < .01$, $F_2(1, 168) = 44.53$, $p < .01$, a significant main effect of AoA in the analyses over items, $F_1(1, 35) = 3.56$, $p = .07$, $F_2(1, 168) = 4.22$, $p < .05$, and a nearly significant interaction between congruency and AoA, $F_1(1, 35) = 2.97$, $p = .09$, $F_2(1, 168) = 2.65$, $p = .10$.

The *t* tests showed that the 25-ms slower RTs to the early-acquired words than to the late-acquired words in the incongruent condition were significant, $t_1(35) = 2.03$, $p < .05$, $t_2(42) = 2.87$, $p < .01$. The 50-ms congruency effect for the early-acquired words was also significant, $t_1(35) = –3.59$, $p < .01$, $t_2(1, 42) = -5.82$, $p < .01$. The same was true for the 25-ms congruency effect for the late-acquired words, $t_1(35) = –3.69$, $p < .01$, $t_2(42) = –4.03$, $p < .01$. These results are given in Figure 1.

The Simon effect was the same for the words belonging to the category of living things (congruent = 560 ms, incongruent = 598 ms) as for the words belonging to the category of nonliving things (congruent = 564 ms, incongruent = 601 ms).

Discussion

According to the neural network account (with distributed representations), the activation of word meanings has become a prime candidate for the origin of the AoA effect in visual word processing because the mappings between spellings and meanings and between sounds and meanings are arbitrary, so that no learning from early-acquired items can transfer to the learning of late-acquired items. Combined with the loss of plasticity in learning systems, this results in stronger connections to and from the meanings of early-acquired words compared with later-acquired words. According to localist accounts, meanings of early-acquired concepts can be activated more easily than those of later-acquired concepts because early-acquired words take a more central position in the network and have more connections with other nodes within the network. At the same time, however, we noted that there was not much compelling empirical evidence for the semantic hypothesis.

We presented results from an experiment that corroborated the semantic hypothesis. In this experiment, participants had to give a verbal response to the visual appearance of a stimulus words (printed in uppercase vs. lowercase). In half of the trials, the response was congruent with the meaning of the stimulus word (e.g., saying “living” to *DEER* or “nonliving” to *cave*); in the other half it was incongruent (e.g., saying “living” to *HARP* or “nonliving” to *finch*). Although the meaning of the stimulus word had to be ignored for good task performance, we found a congruency effect: Responses were faster in the congruent trials than the

Table 1
Characteristics of the Stimulus Lists: Age of Acquisition, Logarithm of Frequency, Word Familiarity, Word Length, and Number of Syllables

Variable	AoA	Log(freq)	WF	WL	NS
Early acquired, living	6.4	2.1	3.0	6.6	1.9
Early acquired, nonliving	6.0	2.1	3.0	7.6	2.1
Late acquired, living	9.2	1.9	2.9	6.5	2.1
Late acquired, nonliving	9.8	2.1	2.9	6.4	2.1

Note. AoA = age of acquisition (in years); Log(freq) = logarithm of frequency; WF = word familiarity; WL = word length; NS = number of syllables.

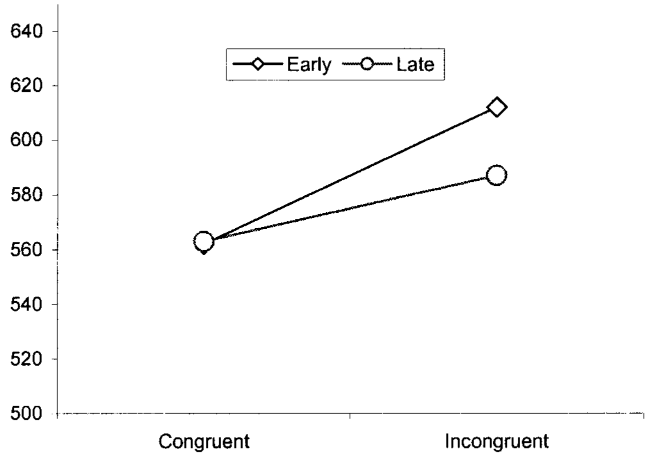

Figure 1. Reaction times for congruent and incongruent trials as a function of age of acquisition.

incongruent trials presumably because the meaning of the target word was activated automatically and interfered with the meanings of the verbal responses that were to be produced. In addition, the congruency effect was twice as large for early-acquired words as for late-acquired words (see Figure 1), in line with our hypothesis that the meaning is activated faster for first-learned words than for later-learned words.

To prevent confusion about our theoretical position, we stress that we do not interpret our findings as evidence for the claim that the AoA effect in visual word recognition (or indeed any other task) is solely due to the meaning of the stimuli. The neural network account (Ellis & Lambon Ralph, 2000; see also Lewis, 1999, for the cumulative frequency hypothesis, which makes an analogue prediction) has made it clear that the effect of AoA is an emerging property of learning systems and is unlikely to be limited to a single stage. However, what our data do show is that the AoA effect in word-processing tasks is not totally due to the activation of word forms but also to the activation of word meanings.

References

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database (CD-ROM).* Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Becker, C. A. (1979). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 252–259.

Bowers, J. S. (2002). Challenging the widespread assumption that connectionism and distributed representations go hand-in-hand. *Cognitive Psychology, 45,* 413–445.

Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica, 104,* 215–226.

Chalard, M., Bonin, P., Meot, 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.

De Houwer, J. (1998). The semantic Simon effect. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 51A,* 683–688.

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.

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, and Cognition, 24,* 267–283.

Ghyselinck, M., Custers, R., & Brysbaert, M. (2003). Age-of-acquisition ratings for 2,332 Dutch words from 49 different semantic categories. *Psychologica Belgica, 43,* 181–241.

Johnston, R. A., & Barry, C. (2002, November). *Age-of-acquisition effects on two picture categorization tasks.* Poster presented at the 43rd Annual Meeting of the Psychonomic Society, Kansas City, MO.

Lewis, M. B. (1999). Age of acquisition in face categorisation: Is there an instance-based account? *Cognition, 71,* 23–39.

Lewis, M. B., Chadwick, A. J., & Ellis, H. D. (2002). Exploring a neural-network account of age-of-acquisition effects using repetition priming of faces. *Memory and Cognition, 30,* 1228–1237.

Moore, V., & Valentine, T. (1998). Naming faces: The effect of AoA on speed and accuracy of naming famous faces. *Quarterly Journal of Psychology, 51A,* 458–513.

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, 21,* 116–133.

Morrison, C. M., Hirsh, K. W., & Duggan, G. B. (2003). Age of acquisition, ageing, and verb production: Normative and experimental data. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 56,* 705–730.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114,* 510–532.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428.

Steyvers, M., & Tenenbaum, J. (2003). *The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth.* Manuscript submitted for publication.

van Loon-Vervoorn, W. A. (1989). *Eigenschappen van basiswoorden.* Lisse, The Netherlands: Swets and Zeitlinger.

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.

(*Appendix follows*)

Appendix

Stimuli

Early-acquired		Late-acquired	
Living	Nonliving	Living	Nonliving
Dolfijn (dolphin)	Badmuts (swimming cap)	Adder (viper)	Abdij (abbey)
Eekhoorn (squirrel)	Blokfluit (recorder)	Adelaar (eagle)	Atoombom (atom bomb)
Egel (hedgehog)	Draaimolen (merry-go-round)	Bloedzuiger (leech)	Aula (auditorium)
Gorilla (gorilla)	Fluit (flute)	Fazant (pheasant)	Beitel (chisel)
Goudvis (goldfish)	Grot (cave)	Havik (goshawk)	Cello (cello)
Hert (deer)	Hobbelpaard (rocking horse)	Kakkerlak (cockroach)	Doedelzak (bagpipes)
Ijsbeer (polar bear)	Jojo (yo-yo)	Kameleon (chameleon)	Fundament (foundation)
Inktvis (cephalopod)	Kinderwagen (baby buggy)	Karper (carp)	Harp (harp)
Kalf (calf)	Klompen (wooden shoes)	Koolmees (coletit)	Kano (canoe)
Krokodil (crocodile)	Knikkers (marbles)	Lama (llama)	Koepel (dome)
Neushoom (rhinoceros)	Koord (cord)	Poema (puma)	Limousine (limousine)
Nijlpaard (hippopotamus)	Kruiwagen (wheelbarrow)	Ratelslang (rattlesnake)	Panty (tights)
Papegaai (parrot)	Luchtballon (hot air balloon)	Reiger (heron)	Pilaar (pillar)
Parkiet (parakeet)	Poppenhuis (doll's house)	Rog (ray)	Piramide (pyramid)
Roodborstje (robin)	Poppenkast (puppet theatre)	Salamander (salamander)	Rasp (grater)
Sprinkhaan (grasshopper)	Schepje (small spoon)	Schorpioen (scorpion)	Scharnier (hinge)
Tijger (tiger)	Slinger (swing)	Snoek (pike)	Scooter (scooter)
Veulen (foal)	Stal (stable)	Spreeuw (starling)	Tandem (tandem)
Vlo (flea)	Tractor (tractor)	Tonijn (tunnyfish)	Tol (top)
Wesp (wasp)	Trompet (trumpet)	Valk (falcon)	Vrachtschip (cargo ship)
Wolf (wolf)	Tuinstoel (garden chair)	Vink (finch)	Vuurwapen (firearm)
Zebra (zebra)	Zwaard (sword)	Zalm (salmon)	Zuil (pillar)

Note. Stimuli were presented in Dutch. English equivalents are shown in parentheses.

Received October 30, 2002
Revision received July 17, 2003
Accepted July 27, 2003 ■