EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 1996, *8* (2), 185–193

Word Frequency Affects Naming Latency in Dutch when Age of Acquisition is Controlled

Marc Brysbaert
Department of Psychology, University of Leuven, Leuven, Belgium

Morrison and Ellis (1995) claim that most evidence of frequency effects in word recognition is not genuine but an artefact of the age at which the words have been acquired. The finding that age of acquisition (AOA) has a reliable independent effect on word naming is replicated for the Dutch language. However, it is also shown that the effect of word frequency remains reliable with AOA controlled. A possible interpretation is that the English studies have been based on retrospective student ratings, whereas in the present study a more on-line measure of AOA was used.

INTRODUCTION

There is general agreement among psycholinguists that all good models of word recognition should account for the finding that words frequently seen in texts are processed faster than words that are rarely encountered (for a review, see Monsell, 1991). Frequency effects are thought to be due to (1) differences in activation levels, (2) the existence of a frequency-dependent verification stage, or (3) a combination of activation and verification.

One of the first activation models was Morton's (1969) "logogen" model. Morton assumed that each word detector (logogen) has its own recognition threshold, which is a function of the frequency of the word. So given a certain amount of sensory evidence, thresholds for high-fre-

Requests for reprints should be addressed to Marc Brysbaert, Department of Psychology, University of Leuven, B-3000 Leuven, Belgium. E-mail: MarcBrysbaert@psy.kuleuven.ac.be.

This work was supported by the Belgian "Nationaal Fonds voor Wetenschappelijk Onderzoek", of which the author is a research associate. The experiment was run by Christel Geerinck under the supervision of Ingrid Gielen and Marc Brysbaert. I am grateful to Fernanda Ferreira and Gery d'Ydewalle for their helpful comments on an earlier draft of the manuscript.

quency words are reached earlier than those for low-frequency words. Another well-known activation model is the "interactive-activation" model of McClelland and Rumelhart (1981). As with the logogen model, this model has a node for each familiar word and an activation level that depends on the frequency of the word. Later connectionist models replaced single-word units by layers of more primitive interconnected units, and had the frequency effects situated in the association strengths between units both within and across layers (e.g. Seidenberg & McClelland, 1989).

The simplest verification models (e.g. Forster, 1976) state that persons compare the orthographic input to a lexicon of word forms, and that this happens serially according to the frequency of the word forms. High-frequency word forms are compared before low-frequency word forms, which reduces the average time needed for the search. Later verification models either incorporated a more sophisticated search algorithm (e.g. Taft, 1979) or a parallel component (e.g. Forster, 1992) to explain how a search among some 50,000 alternatives can be achieved within less than 100 msec (Forster, 1992).

The final class of models (e.g. Paap, Newsome, McDonald, & Schvaneveldt, 1982) uses a combination of activation and verification to explain frequency effects. The idea is that word identification happens in two discrete stages. The first stage ("activation") is a fast parallel activation process that does not terminate in a unique identification, but establishes a small set of candidates which meet some fairly crude criterion of compatibility with the stimulus features. In the second stage ("verification"), these candidates are examined one at a time in order of their frequency.

Given the pivotal role of word frequency in current models of word recognition, it was quite surprising to come across a paper by Morrison and Ellis (1995) in which it was claimed that most experimental evidence of frequency was not genuine but an artefact of the age at which the words had been acquired. In particular, Morrison and Ellis (1995) claimed that when the variance due to age of acquisition (AOA) is partialled out, word frequency does not affect naming latencies at all, although frequency has some independent effect on lexical decision times. Morrison and Ellis explained the difference between naming and lexical decision by assuming that the latter, in addition to being based on orthographic and phonological information, also makes use of semantic information and that the access to the semantic system is the origin of the frequency effect.

The importance of Morrison and Ellis's (1995) claim is two-fold. First, it points to a possible misunderstanding of the origin of processing differences between words. Second, it suggests that a whole category of word recognition models may be inadequate, as the AOA effect is not easily explained within the framework of connectionist learning models based on back-propagation (Morrison & Ellis, 1995). This is because the back-propagation algorithm makes newly learned patterns overwrite and obliterate pre-existing patterns (e.g. McCloskey & Cohen, 1989; Ratcliff, 1990), exactly the opposite of the AOA effect.

However, a weak aspect of Morrison and Ellis's (1995) paper (and all previous studies cited in the paper) is that the AOA measures were based on student ratings. Most of the time these ratings are based on the work of Gilhooly and Logie (1980), who asked university students to estimate the AOA of different words on a 7-point scale ranging from below 2 years of age to more than 13 years of age. Although the reliability of these ratings is very high (from 0.96 to 0.99 in six studies reviewed by Gilhooly & Watson, 1981), it is clear that such retrospective judgements may involuntarily involve a frequency component. Raters may tend to underestimate the AOA of frequent words and overestimate the AOA of rare words. The reported correlations between tabulated frequencies and AOA estimates are considerable and range from 0.40 (Rubin, 1980) to 0.71 (Gilhooly & Logie, 1982). In the absence of more objective measures, it is not clear to what extent the size of these correlations is due to real covariation between frequency of occurrence and AOA in English, and how much is due to response biases.

A large-scale study in the Dutch-speaking countries (the Netherlands and the northern part of Belgium) facilitated the extraction of a more direct measure of AOA (Kohnstamm, Schaerlaekens, de Vries, Akkerhuis, & Froonincksx, 1981). This study involved a representative sample of teachers who were asked to indicate which words they believed 6-year-olds should understand (i.e. passive knowledge). In particular, for Belgium, 40 teachers of the last year of kindergarten and 41 teachers of the first year of primary school were given a list of 6785 words and were asked which ones they thought children should know when they move from kindergarten to primary school. The dependent variable was the percentage of teachers who marked the words, henceforth referred to as the AOAT measure (age of acquisition based on teachers' judgements).

Previous research by van Loon-Vervoorn (1989) showed that the correlations between AOATs and student ratings on an 8-point scale range from 0.87 (for verbs, 52 words) to 0.92 (for nouns, 44 words). Two further studies indicated that the correlations between AOATs and tabulated word frequencies range from 0.13 to 0.33, based on 500 and 300 words, respectively. Note that the correlations between AOAT and frequency in Dutch tend to be lower than those between AOA and frequency in English (see above).

The study reported below investigated what effect word frequency has on naming latency with AOAT controlled. For practical reasons, the study was undertaken with third-year primary school children rather than with university students. The use of younger subjects has the advantage that the impact of AOA and word frequency can be examined while the verbal system is still developing. The disadvantage is that the effect of word frequency may be underestimated because of the more limited experience with language.

METHOD

Participants

The subjects were 22 female third-year primary school children (8–9 years old), who attended two different schools. None had reading problems, as indicated by their teachers.

Materials

The materials comprised 204 Dutch words of three ($n = 41$) or four letters ($n = 163$). Care was taken to use words that were headwords (i.e. without inflections), and that were likely to be known by 8-year-olds. Frequency estimates of the words were obtained from the CELEX database with a total of 42,380,000 counts (Burnage, 1990). They were converted into logarithmic values by taking the natural logarithm of the frequency plus one. The mean log frequency value was 7.61 and the standard deviation 1.17. The AOAT measures representing the percentage of teachers indicating that the target word should be known by 6-year-olds (see above) were based on Kohnstamm et al. (1981); they had a mean value of 78.2% and a standard deviation of 23.9%. The correlation between the frequency measure and AOAT was 0.04 ($n = 204$, NS), the correlation between frequency and word length 0.04, and the correlation between AOAT and word length 0.03. The lack of correlations between the predictors ensured that the effect of each variable could be assessed independently of the other.

Procedure

The words were divided into two lists and randomised for each subject separately (Brysbaert, 1991). They were presented on a visual display unit, which was connected to an IBM-XT. The participants were asked to look at a gap between two vertically aligned lines in the middle of the screen. The words were presented horizontally so that the second letter fell between the lines. Previous research had indicated that the second letter is the optimal viewing position for naming short Dutch words (Brysbaert, 1994). The words were presented in lower case and remained visible until the child reacted. Reaction times were measured with a voice key connected to the game port and registered to the nearest millisecond using software routines published by Bovens and Brysbaert (1990). The experimenter sat next to the observer and typed in whether the naming and the time registration were correct. It took on average 20 min for the participants to complete a list. They were allowed to take a short break between the lists. At the beginning of the experiment, 20 practice trials were given.

RESULTS

Data from trials in which a word was incorrectly named or the reaction time incorrectly registered were discarded from the analyses described below. Reaction times longer than 2000 msec were also deleted. This resulted in a total of 232 missing observations (5.2% of all data). Rather than calculating the correlations and using tabulated values, I followed the procedure recommended by Lorch and Myers (1990) for assessing the reliability of linear regression weights in repeated measurement designs. It consists of determining the regression weights for each person, and calculating group *t*-tests on the values obtained to see whether they differ reliably from zero. Preliminary analyses indicated that there was no interaction between word frequency and AOAT (cf. Forster, 1992); therefore, the analysis is limited to the main effects.

All three variables made a reliable contribution to naming times. The regression weight for word length was 16.3 msec per letter [$t(21) = 2.89$, $P < 0.01$]; for word frequency it was –10.5 [$t(21) = -3.24$, $P < 0.01$] and for AOAT –0.58 [$t(24) = -3.37$, $P < 0.01$]. The intercept was 783 msec. Taking four times the standard deviation as a rough measure of the extent of the effect, it was found that the magnitude of the frequency effect ($4 \times 1.17 \times 10.5 = 49$ msec) was of the same order as the magnitude of the AOAT effect ($4 \times 23.9 \times 0.58 = 55$ msec).

To examine whether the reliable independent effect of word frequency may have been due to the fact that the AOAT measure does not permit one to make fine distinctions between the very first words learned during childhood (i.e. the words that all 6-year-olds should know according to their teachers), I repeated the analysis for the 128 words with an AOAT measure lower than 95%. This analysis essentially repeated the findings of the analysis based on all 204 words: word length = 22.0 [$t(21) = 3.37$, $P < 0.01$], log frequency = –12.1 [$t(21) = -3.25$, $P < 0.01$] and AOAT = –0.50 [$t(21) = -2.95$, $P < 0.01$].

DISCUSSION

This study produced two main findings. First, it replicated the results of Morrison and Ellis (1995) by showing that age of acquisition has a reliable effect on word naming latencies when the variance due to word frequency is controlled for. As far as I know, this is the first demonstration of the AOA effect on naming in a language other than English. As such, it adds further credit to Morrison and Ellis's (1995) claim that future studies and theories of word recognition should take this variable into account.

The present study, however, also shows that in Dutch and with the AOA measure used, word frequency still has a reliable independent effect on naming latencies. This deviates clearly from Morrison and Ellis's finding. Several interpretations may be proposed. First, it could be that the AOAT measure is not sensitive enough to capture all variance due to age of acquisition. It could be hypothesised, for instance, that the AOAT measure, being heavily focused on 6-year-olds, excludes fine distinctions that can be made between words learned around the age of 1–4 years. This explanation is unlikely, however, given that exactly the same picture with similar regression weights and reliability gradients is obtained when the analysis is limited to the 128 words with AOAT less than 95%. Furthermore, the schools made use of instruction lists to know which words the pupils should be able to read in each grade (and therefore the orthography of which should be taught during that year). When this variable was introduced into the regression analyses rather than the AOAT measure, the same result was obtained—word frequency had a reliable independent effect.

Another interpretation of the difference between the present results and those of Morrison and Ellis (1995) is that their study may have lacked the power to reveal the effect of frequency. This lack of power may not only have been due to the number of stimuli involved (i.e. 24 instead of 204), but also to the fact that Morrison and Ellis tested university students rather than 8- to 9-year-old children. Although in principle working with adults should enhance the importance of word frequency (because the frequency effect arises from repeated encounters with the same words), I repeatedly found that children are influenced more by visual and linguistic factors in word processing than adults are (e.g. Brysbaert & Meyers, 1993). This finding can be illustrated with the variable of word length. For adults, the regression weight usually fluctuates around 3–5 msec per letter (Brysbaert, 1992; Hudson & Bergman, 1985); in the present study, it was 16 msec per letter. The larger impact of this variable for children than for adults may imply that the effect of a factor (even of word frequency) is more easy to demonstrate in children.

A final reason for the divergence between Morrison and Ellis (1995) and the experiment reported here may be that the Gilhooly and Logie (1980) ratings are not pure measures of AOA. As indicated in the Introduction, the retrospective AOA estimates of university students may be biased by the frequency of occurrence of the words, so that the AOA measure unintendedly includes a frequency component (or a familiarity component; see Gernsbacher, 1984). This bias would tend to increase the impact of the "combined" AOA–frequency measure at the expense of the pure frequency measure. It would certainly be interesting to repeat Morrison and Ellis's (1995) study with AOAT measures, to see whether AOA in English really can explain all variance in naming times previously associated with word frequency.

As indicated by Morrison and Ellis (1995), the reliable effect of AOA poses serious problems for many existing models of word recognition. Ironically, this is especially true for the (connectionist) models that obtain the frequency effect as an intrinsic outcome of the learning algorithm (Monsell, 1991). Other models that account for the effect of frequency in terms of differences in activation level or search order can change the terminology more easily and make the activation levels and/or the search order dependent on the age of acquisition (in addition to frequency). Indeed, such a model already exists (Forster, 1992). According to this model, word forms are stored in bins of 20–50 words, which are checked in parallel when an orthographic input needs to be recognised (serial verification within the bin). The idea is that during early lexical acquisition, there may be only a single bin, and that words are added to it in order of their acquisition. When the bin gets too large, a new bin is added, and further words are added to the new bin. So initial order is based purely on the order of acquisition, but this initial order is subsequently partially revised by the relative frequencies of the words. One option is that a word, having been encountered, moves halfway between the top of the bin and its current position. This re-ordering would lead to word-frequency effects in addition to AOA effects, exactly the pattern of results obtained in the present experiment (and in the study described by Forster, albeit with a lexical decision task). Other word recognition models (e.g. without a serial verification stage) can certainly be developed or adapted to explain the present findings. What needs to be taken into account is that the order of acquisition has a more profound effect than assumed until now, but not so profound as to rule out every later order re-arrangement or activation appreciation due to the encountering of words in real life.

Manuscript received and accepted July 1995

REFERENCES

Bovens, N., & Brysbaert, M. (1990). IBM PC/XT/AT and PS/2 Turbo Pascal timing with extended resolution. *Behavior Research Methods, Instruments and Computers*, *22*, 332–334.

Brysbaert, M. (1991). Algorithms for randomness in the behavioral sciences: A tutorial. *Behavior Research Methods, Instruments and Computers*, *23*, 45–60.

Brysbaert, M. (1992). *Interhemispheric transfer in reading*. Unpublished doctoral thesis, University of Leuven.

Brysbaert, M. (1994). Interhemispheric transfer and the processing of foveally presented stimuli. *Behavioural Brain Research*, *64*, 151–161.

Brysbaert, M., & Meyers, C. (1993). The optimal viewing position for children with normal and with poor reading abilities. In S.F. Wright & R. Groner (Eds), *Facets of dyslexia and its remediation*, pp. 107–123. Amsterdam: Elsevier.

Burnage, G. (1990). *Celex: A guide for users*. Nijmegen: Centre for Lexical Information.

Forster, K.I. (1976). Accessing the mental lexicon. In R.J. Wales & E. Walker (Eds), *New approaches to language mechanisms*. Amsterdam: North-Holland.

Forster, K.I. (1992). Memory-addressing mechanisms and lexical access. In R. Frost & L. Katz (Eds), *Orthography, phonology, morphology, and meaning*, pp. 413–434. Amsterdam: Elsevier.

Gernsbacher, M.A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*, 256–281.

Gilhooly, K.J., & Logie, R.H. (1980). Methods and designs: Age of acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods and Instruments*, *12*, 395–427.

Gilhooly, K.J., & Logie, R.H. (1982). Word age-of-acquistiton and lexical decision making. *Acta Psychologica*, *50*, 21–34.

Gilhooly, K.J., & Watson, F.L. (1981). Word age-of-acquisition effects: A review. *Current Psychological Research*, *1*, 269–286.

Hudson, P.T.W., & Bergman, M.W. (1985). Lexical knowledge in word recognition: Word length and word frequency in naming and lexical decision tasks. *Journal of Memory and Language*, *24*, 46–58.

Kohnstamm, G.A., Schaerlaekens, A.M., de Vries, A.K., Akkerhuis, G.W., & Froonincksx, M. (1981). *Nieuwe streeflijst woordenschat*. Lisse: Swets and Zeitlinger.

Lorch, R.F., Jr, & Myers, J.L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *16*, 149–157.

McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*, 375–407.

McCloskey, M., & Cohen, N.J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In G.H. Bower (Ed.), *The psychology of learning and motivation*, Vol. 24, pp. 109–164. New York: Academic Press.

Monsell, S. (1991). The nature and locus of word frequency effects in reading. In D. Besner & G.W. Humphreys (Eds), *Basic processes in reading: Visual word recognition*, pp. 148–197. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

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 and Cognition*, *21*, 116–133.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, *76*, 165–178.

Paap, K.R., Newsome, S.L., McDonald, J.E., & Schvaneveldt, R.W. (1982). An activation–verification model for letter and word recognition. *Psychological Review*, *89*, 573–594.

Ratcliff, R. (1990). Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychological Review*, *97*, 285–308.

Rubin, D.C. (1980). 51 properties of 125 words: A unit analysis of verbal behavior. *Journal of Verbal Learning and Verbal Behavior*, *19*, 736–755.

Seidenberg, M.S., & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523–568.

Taft, M. (1979). Lexical access via an orthographic code: The Basic Orthographic Syllable Structure. *Journal of Verbal Learning and Verbal Behavior*, *18*, 21–39.

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