THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2006, 59 (10), 1681–1690

Psychology Press
Taylor & Francis Group

Short article

Does word frequency affect lexical selection in speech production?

Eduardo Navarrete
GRNC, Parc Científic Universitat de Barcelona, & Hospital Sant Joan de Déu, Spain

Benedetta Basagni
GRNC, Parc Científic Universitat de Barcelona, & Hospital Sant Joan de Déu, Spain; Centro Cardinal Ferrari di Fontanellato, Parma, Italy

F.-Xavier Alario
GRNC, Parc Científic Universitat de Barcelona, & Hospital Sant Joan de Déu, Spain; CNRS & Université de Provence, Marseille, France

Albert Costa
GRNC, Parc Científic Universitat de Barcelona & Hospital Sant Joan de Déu, Spain

We evaluated whether lexical selection in speech production is affected by word frequency by means of two experiments. In Experiment 1 participants named pictures using utterances with the structure "pronoun + verb + adjective". In Experiment 2 participants had to perform a gender decision task on the same pictures. Access to the noun's grammatical gender is needed in both tasks, and therefore lexical selection (lemma retrieval) is required. However, retrieval of the phonological properties (lexeme retrieval) of the referent noun is not needed to perform the tasks. In both experiments we observed faster latencies for high-frequency pictures than for low-frequency pictures. This frequency effect was stable over four repetitions of the stimuli. Our results suggest that lexical selection (lemma retrieval) is sensitive to word frequency. This interpretation runs against the hypothesis that a word's frequency exerts its effects only at the level at which the phonological properties of words are retrieved.

One of the most robust effects in speech production is that of word frequency. Word frequency affects the speed and accuracy with which words are produced in both normal and aphasic speakers

Correspondence should be addressed to Albert Costa, Dept. Psicologia Bàsica, Universitat de Barcelona, P. Vall d'Hebron, 171, 08035 Barcelona, Spain. E-mail: acosta@ub.edu

This research was supported by a grant from the Spanish Government (SEJ 2005-00409/PSIC), a grant from the European Science Foundation (BFF2002–10379-E), the McDonnell grant "Bridging Mind Brain and Behavior", and a grant from the HFSP (RGP68/2002). F.-Xavier Alario was supported by a grant from the "Programa de Movilidad" (ref. SB2003–0322) from the Spanish government.

http://www.psypress.com/qjep
DOI:10.1080/17470210600750558

(Caramazza & Hillis, 1990; Dell, 1990; Oldfield & Wingfield, 1965). Although there is agreement on assuming that part of this effect originates when retrieving words from the lexicon (e.g., Alario, Costa, & Caramazza, 2002; Dell, 1990; Griffin & Bock, 1998; Wingfield, 1968), there is still some debate regarding its precise locus. The aim of the present study is to assess whether lexical selection (lemma retrieval) is sensitive to word frequency.

In Levelt's model (Levelt, 1989; Levelt, Roelofs, & Meyer, 1999), lexical access entails the retrieval of two distinct lexical representations: the lemma node (the syntactic and semantic properties of the lexical item) and the lexeme node (the morpho-phonological properties; see also Dell, 1986). In this model, word frequency may affect the retrieval of the lemma and/or the retrieval of the lexeme. Two important results reported by Jescheniak and Levelt (1994) helped to constrain these hypotheses.

First, Jescheniak and Levelt (1994) observed similar naming latencies for low-frequency homophone words (e.g., "nun") and for words matched for the cumulative frequency of the two meanings of the homophone (e.g., "nun" + "none"). That is, the actual frequency of "nun" is the result of its own specific-word frequency plus the specific-word frequency of its homophone twin ("none"). It is as if "nun" inherits the frequency of "none". This observation was used to claim that homophones share a lexeme and that word frequency mostly affects lexeme retrieval.[1]

The second critical observation comes from an experiment in which participants were asked to provide the gender value of pictures' names. Jescheniak and Levelt (1994) argued that participants perform the gender decision task by exploring the properties of the lemma without retrieving the corresponding lexeme. Based on this assumption, they further argued that, if word frequency mostly affects the retrieval of the lexeme, gender decision times should be independent of word frequency. The results were consistent with this prediction. Although they observed a word frequency effect the first time the pictures were presented, the effect vanished in subsequent presentations. This transient frequency effect contrasts with the robustness over repetitions of the frequency effect in picture naming (e.g., Levelt, Praamstra, Meyer, Helenius, & Salmelin, 1998), and it was attributed to a different origin from that of the genuine frequency effect observed in naming. These findings led the authors to conclude that lemma retrieval seems insensitive to word frequency: "If [. . .] the locus of the robust word frequency effect is the lemma threshold activation, we should find a robust frequency effect in gender decision. That did not happen" (p. 840).

In the present article we focus on this last result. We further test whether the frequency effect is absent in tasks in which the phonological content of the target is not produced. The rationale of our experiments is exactly the same as that used by Jescheniak and Levelt (1994). We assume that the tasks presented below can be performed by exploring the properties of the lemma, without lexeme retrieval. Although such an assumption may be controversial (Starreveld & La Heij, 2004), it has been embraced not only by Jescheniak and Levelt (1994) but also by other researchers in the field, leading to important conclusions about the nature of lexical access. For example, by using a paradigm involving the gender decision task, Van Turennout, Hagoort, and Brown (1998) estimated that access to the syntactic properties of words precedes access to their phonological properties by about 40 ms.

[1] The homophone frequency effect has been shown to be a rather elusive effect. In fact, several investigations have failed to replicate the effect in several languages (English and Mandarin Chinese: Caramazza, Costa, Miozzo, & Bi, 2001; French: Bonin & Fayol, 2002). Furthermore, note that the presence of homophone frequency effects does not necessarily mean that word frequency is located at the lexeme level. For example, Dell (1990) found a similar homophone frequency effect in slips of the tongue and located the frequency effect in the links between lemma and lexeme levels. Whether or not one accounts for such an effect by locating the word frequency effect at the lemma level depends on other assumptions about the processing dynamics in lexical access (e.g., interactivity; further discussion of this issue can be found in Alario et al., 2002; Caramazza et al., 2001; Levelt, 2002).

In the following, then, we interpret our results in the theoretical framework used by Jescheniak and Levelt, and we defer any discussion of the validity of its assumptions (and the corresponding theoretical implications) to the General Discussion.

A recent investigation by Finocchiaro and Caramazza (2006) made use of this same framework but reached a different conclusion from that drawn by Jescheniak and Levelt (1994). In one of the experiments, participants were asked to produce Italian utterances with the structure "verb + pronominal gender-marked clitic form" (e.g., *portalo*, literally "bring it_{masc}" when referring to a *treno*, $train_{masc}$; or *portala*, literally "bring it_{fem}" when referring to a *sedia*, $chair_{fem}$). The retrieval of the appropriate clitic (*lo* or *la*) requires access to the noun's grammatical gender (masculine or feminine). In this context, the retrieval of the referent noun's lexeme is irrelevant for performing the task (the noun is not produced, and its phonological properties are irrelevant for the selection of the clitic form). Thus, in the pertinent dimensions, the gender decision task and the pronominal clitic naming task are very similar since both (a) entail the retrieval of the noun's lemma node, and (b) do not entail the retrieval of the noun's lexeme node. According to the rationale used by Jescheniak and Levelt, naming latencies should be independent of the frequency value of the picture's name. Contrary to this prediction, naming latencies were faster when the frequency of the clitic's referent was high than when it was low. This observation contrasts with the results reported by Jescheniak and Levelt and suggests that lexical selection is affected by word frequency.

Given these contrasting results, it is premature to reach a definite conclusion about the emergence of the frequency effect in tasks where the phonological properties of target words are not retrieved/produced. Of course, one possible reason for these contrasting results could lie in different task requirements for gender decision and pronoun naming. That is, one could claim that, contrary to the rationale that we have been following, the pronoun-naming task actually entails the retrieval of the phonological form of the referent noun (through a mechanism that would need to be specified). If this were to be the case, Jescheniak and Levelt's (1994) rationale would predict that the frequency effect would be absent in the gender decision, but present in the pronoun production task.

Before speculating between the different processes and mechanisms that might be behind this contrasting pattern of results it is important to assess the reliability of the two phenomena (lack of word frequency effect in gender decision and presence in pronoun naming) using the same language, materials, and laboratory. This is precisely the objective of our investigation, in which native speakers of Spanish were asked to perform a pronoun-naming task and a gender decision task while the frequency of the target nouns was manipulated.

EXPERIMENT 1: WORD FREQUENCY EFFECTS IN PRONOMINAL SENTENCE PRODUCTION

In this experiment participants were asked to name pictures by means of utterances of the structure "pronoun + verb + adjective", where both the pronoun and the adjective agree with the grammatical gender of the referent noun: *Esta*$_{fem}$ *es nueva*$_{fem}$ [literally, $this_{fem}$ (one) is new_{fem}], *Este*$_{masc}$ *es nuevo*$_{masc}$ [literally, $this_{masc}$ (one) is new_{masc}]. The phonological properties of the referent noun are irrelevant for the selection of the other words. However, access to the lemma node of the referent noun is needed to retrieve the gender inflections. Thus, following the exact rationale used by Jescheniak and Levelt (1994) production latencies should be independent of the frequency value of the pictures' names.

Method

Participants

A total of 16 native speakers of Spanish, students at the University of Barcelona, took part in this experiment.

Materials

A total of 48 pictures of common objects (items) were selected: 24 with high-frequency names (HF: 734 occurrences per million on average) and 24 with low-frequency names (LF: 30 occurrences per million on average). Both sets were matched for number of syllables (HF = 2.2, LF = 2.2) and number of phonemes (HF = 4.8, LF = 5.1). In the HF set, 12 picture names were feminine, and 12 were masculine; in the LF set, 10 picture names were feminine and 14 masculine (See Appendix).

To assess whether picture identification was comparable between the two sets, we followed Jescheniak and Levelt's (1994) strategy and ran a word–picture recognition pretest. In this task, 16 other participants were presented with a printed word followed by a picture. They had to decide whether both items corresponded to the same concept. Identification times were similar for the two sets (HF = 475 ms, LF = 481 ms; Fs < 1).

To avoid massive repetition of a single verbal response—for example, "This (one) is new"—two characteristics of depiction of the objects were manipulated: distance (close–far) and appearance (new–old). For each item, four pictures were created, yielding four different responses.

The distance dimension was manipulated by locating the objects in two different positions on the picture: objects located in the inferior left angle were interpreted as being close to the participant's point of view while objects located in the superior right angle of the plane were interpreted as being distant. Participants were instructed to use the pronouns *Este–Esta* ($\text{This}_{\text{masc}}$–$\text{This}_{\text{fem}}$, respectively) for close objects and the pronouns *Aquel–Aquella* ($\text{That}_{\text{masc}}$–$\text{That}_{\text{fem}}$, respectively) for distant objects.

The appearance dimension was manipulated by keeping the original thin outline pictures or by blurring the contours of pictures (see Alario & Caramazza, 2002). Participants were instructed to use the adjectives *Nuevo–Nueva* (New_{masc}–New_{fem}, respectively) for pictures with thin outlines and the adjectives *Viejo–Vieja* (Old_{masc}–Old_{fem}, respectively) for pictures with blurred outlines.

The instruction was to name the resulting 192 pictures by making reference to both dimensions—for example, "*Este es nuevo*" ($\text{This}_{\text{masc}}$ is new_{masc}; see Figure 1). Four additional items

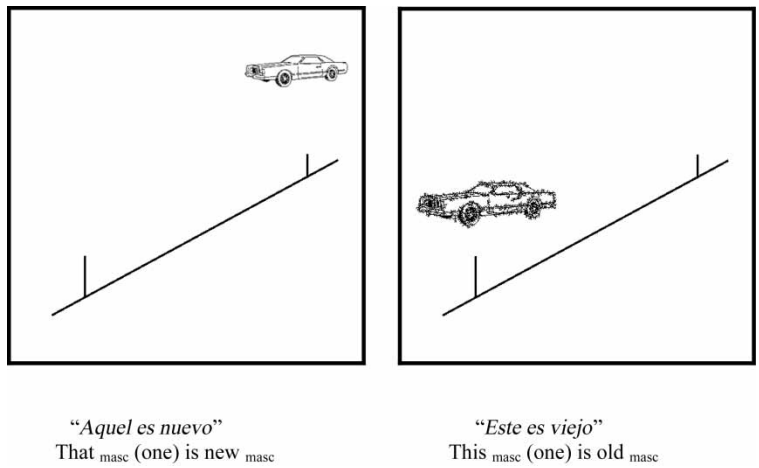

Figure 1. *Two examples of the experimental stimuli are presented in Figure 1. In Experiment 1 participants were instructed to use the adjective* Nuevo/a *(new) to refer to objects with thin outlines (left picture) and to use the adjective* Viejo/la *to refer to objects with blurred outlines (right picture). Participants were instructed to indicate the distance dimension using the pronoun* Aquel/lla *(that) to refer to objects far from the participant's point of view (left picture) and the pronoun* Este/a *(this) to refer to objects close to the participant's point of view (right picture). For each object, four pictures were created.*

were used as warm-up trials at the beginning of each experimental block.

Design and procedure

The experiment included four blocks of 50 trials (48 experimental trials + 2 warm-up trials). For a given block: (a) Each item appeared only once; (b) each format was represented the same number of times, resulting in 12 trials of each format (this/new; this/old; that/new; that/old); (c) the order of the trials was randomized with restrictions (no more than three consecutive trials with the same pronoun, the same adjective, the same gender, or nouns from the same frequency group; two successive items were never semantically nor phonologically related). Participants were randomly assigned to one of eight different block orders.

In the familiarization phase, participants were presented with two exemplars of each object (new and old) placed side by side, and they were asked to name the item aloud. They were given feedback about the intended name when appropriate. A training block (48 experimental targets + 2 warm-up targets) was then administered. Finally, the four experimental blocks were presented with short pauses between them.

An experimental trial involved the following events: (a) a fixation point (1,250 ms); (b) a blank screen (500 ms); (c) the target picture (presented until a response was given, or a 2,000-ms deadline was reached, whichever came first); (e) a question mark (presented until the spacebar was pressed). Participants were asked to name the pictures as fast and as accurately as possible using sentences such as "*Este es nuevo*" ($\text{This}_{\text{masc}}$ is new_{masc}), "*Aquella es vieja*" (That_{fem} is old_{fem}), and so on. Response latencies were measured from the onset of the target presentation. The entire experimental session lasted 35 minutes and was controlled by DMDX software (Forster & Forster, 2003).

Results and discussion

Recording failures were excluded from the analyses (2.2%). Three types of response were considered as errors: (a) production of utterances different from those expected by the experimenter; (b) verbal dysfluencies (stuttering, utterance repairs, etc.); (c) naming latencies below 300 ms or above 3 standard deviations from the participant's mean. Overall, 14.7% of the data points were excluded (see Table 1). The statistical analyses included two factors: "word frequency" (HF vs. LF) and "repetition" (1, 2, 3, and 4).

In the error analysis, the main effect of word frequency was significant, $F_1(1, 15) = 19.40$; $MSE = 51.25$; $p < .01$; $F_2(1, 46) = 9.21$; $MSE = 34.17$; $p < .01$. The effect of repetition was also significant, $F_1(3, 45) = 8.20$; $MSE = 14.70$; $p < .02$; $F_2(3, 138) = 4.40$; $MSE = 9.80$; $p < .05$. Error rates for the HF set were lower than those for the LF set and decreased over repetitions. The interaction between these two factors was not significant (both Fs < 1).

In the naming latencies analysis, the main effects of word frequency, $F_1(1, 15) = 26.62$; $MSE =$

Table 1. *Mean naming latencies[a], standard deviations, and error rates[b] for each of the conditions in Experiment 1*

	Repetition											
	1			*2*			*3*			*4*		
Noun frequency	*M*	*SD*	*E%*	*M*	*SD*	*E%*	*M*	*SD*	*E%*	*M*	*SD*	*E%*
High	819	128	13.3	784	115	8.3	773	117	10.2	748	107	8.0
Low	863	145	16.9	826	143	16.7	803	101	13.0	801	112	14.0
Effect (low − high)	44		3.6	42		8.4	30		2.8	53		6.0

[a]In ms. [b]E%, in percentages.

57,650; $p < .01$; $F_2(1, 46) = 10.09$; $MSE = 77{,}516$; $p < .01$, and repetition, $F_1(3, 45) = 7.38$; $MSE = 75{,}449$; $p < .02$; $F_2(3, 138) = 61.45$; $MSE = 119{,}338$; $p < .01$, were significant. Naming latencies for the HF set were lower than those for the LF set and decreased over repetitions. The interaction between the two factors was not significant (Fs < 1).

The results of this experiment are clear: The production of utterances of the structure "pronoun + verb + adjective" is affected by the frequency of the referent noun. Furthermore, the effect is stable across the four repetitions of the items (Figure 2). This observation replicates the frequency effect observed by Finocchiaro and Caramazza (2006), and it is at odds with the predictions derived from Jescheniak and Levelt's (1994) argumentation. Before deriving conclusions about the sensitivity of lemma access to the frequency manipulation on the sole basis of a pronoun production task, we need to assess the robustness of Jescheniak and Levelt's influential findings, which have not been replicated yet.

EXPERIMENT 2: WORD FREQUENCY EFFECTS IN GENDER DECISION

Participants were asked to make a button press decision on the grammatical gender of the names of pictures. If the gender decision task is not sensitive to word frequency, then the frequency effect observed in Experiment 1 should be absent in this experiment. On the contrary, if the effect that we found in Experiment 1 is not tied to the specific properties of the task that we used, then we should also observe it in the present experiment.

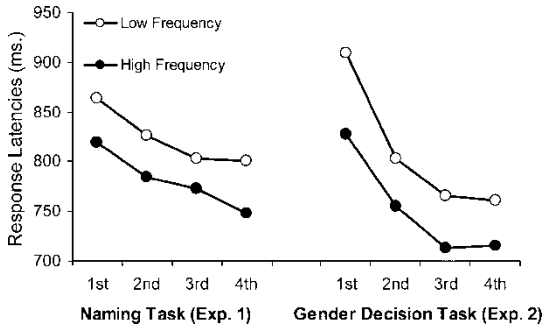

Figure 2. *Average response latencies for high-frequency and low-frequency sets broken by repetition in Experiments 1 and 2. Word frequency effects were present across the four repetitions of the experimental materials.*

Method

A total of 16 participants from the same population as that in Experiment 1 participated in the experiment. The same materials were used. Most experimental details were the same. To follow closely the procedure used by Jescheniak and Levelt (1994) the following modifications in the procedure were made. The training block did not involve experimental pictures but was conducted over 20 filler pictures. Participants were asked to decide whether the picture's name was masculine or feminine, ignoring its position on the display (close vs. distant) and its appearance (normal vs. blurred). Responses were given by pressing the two buttons of a joystick with the two index fingers. The assignment of response side to gender value was counterbalanced across participants.

Results and discussion

Erroneous responses were identified and processed as in Experiment 1 (6.8% of trials, Table 2).

In the error analysis, the main effects of word frequency, $F_1(1, 15) = 22.72$; $MSE = 50.00$; $p < .01$; $F_2(1, 46) = 10.12$; $MSE = 33.33$; $p < .01$, and repetition, $F_1(3, 45) = 43.16$; $MSE = 40.00$; $p < .01$; $F_2(3, 138) = 22.33$; $MSE = 26.66$; $p < .01$, were significant. The interaction between these two factors was also significant, $F_1(3, 45) = 26.49$; $MSE = 11.55$; $p < .01$; $F_2(3, 138) = 6.45$; $MSE = 7.70$; $p < .02$, revealing that the difference for error rates between HF and LF sets was larger for the first repetition than for the three others. The frequency effect reached significance in the three first repetitions (all ps $< .05$) and was marginal for the last repetition (all ps $< .1$).

Table 2. *Mean latencies[a], standard deviations, and error rates[b] for each of the conditions in Experiment 2*

	Repetition											
	1			*2*			*3*			*4*		
Noun frequency	*M*	*SD*	*E%*	*M*	*SD*	*E%*	*M*	*SD*	*E%*	*M*	*SD*	*E%*
High	827	107	5.9	755	117	3.9	713	127	4.4	715	124	2.6
Low	909	130	14.8	803	114	9.9	766	116	8.3	761	120	4.6
Effect (low − high):	82		8.9	48		6.0	53		3.9	46		2.0

[a]In ms. [b]E%, in percentages.

In the latencies analysis (Figure 2), both main effects were significant: word frequency, $F_1(1, 15) = 49.19$; $MSE = 105{,}336$; $p < .01$; $F_2(1, 46) = 6.46$; $MSE = 191{,}329$; $p < .02$; repetition, $F_1(3, 45) = 63.13$; $MSE = 295{,}792$; $p < .01$; $F_2(3, 138) = 180.42$; $MSE = 485{,}070$; $p < .01$. The interaction between the two variables was not reliably significant, $F_1(3, 45) = 2.32$; $MSE = 4{,}187$; $p > .14$; $F_2(3, 138) = 4.59$; $MSE = 12{,}350$; $p < .04$. The frequency effect was significant for all repetitions (all ps $< .05$).

The results of this experiment reveal that the word frequency effect remains stable over four repetitions, therefore contrasting with the transient frequency effect reported by Jescheniak and Levelt (1994).

GENERAL DISCUSSION

We reported two experiments assessing whether lexical selection (lemma retrieval) is sensitive to word frequency. In Experiment 1, participants performed a gender-marked pronominal naming task (pronoun + verb + adjective), and in Experiment 2 they performed a gender decision task. In both cases, responses were faster for pictures with high-frequency names than for pictures with low-frequency names. Importantly, this word frequency effect was present across the four repetitions of the same stimuli, suggesting that the effect is as robust as it is in regular picture naming (see, e.g., Experiment 1 in Jescheniak & Levelt, 1994, and also Levelt et al., 1998).

These results are consistent with those reported by Finocchiaro and Caramazza (2006) and inconsistent with the null effect observed by Jescheniak and Levelt (1994). Furthermore, they extend the observation of word frequency effects to the gender decision task. This is important as it reveals that the word frequency effect in the pronoun-naming task is not tied to some task-specific property or demand. Rather, the effect reveals a more general property of the speech production system.

Our observations contradict the hypothesis predicting the absence of a word frequency effect in tasks where lexical selection is needed but lexeme retrieval is not. Thus, it appears that for those tasks where only lemma retrieval is logically needed, word frequency effects are present. This empirical generalization indicates that lexical selection is sensitive to word frequency.

The presence of a word frequency effect in these experiments has implications for the functional architecture of the speech production system. Levelt et al. (1999) argued that the absence of a genuine word frequency effect in the gender decision task, as well as the presence of a homophone frequency effect in the naming task, requires postulating two different levels of lexical representation: the lemma and the lexeme levels. The argument based on the absence of a genuine word frequency effect in the gender decision task goes as follows. Given that the word frequency

effect observed in word production is not located at the level at which the grammatical representations of a word are selected (the lemma level), one should postulate another lexical level of representation at which frequency exerts its effects (the lexeme level). As stated by Levelt et al. (1999): "A further argument for an independent lemma representation derives from experiments by Jescheniak and Levelt (1994) [. . .] What matters [. . .] is that gender and form properties of the word bear markedly different relations to word frequency" (p. 14). The argument based on the homophone frequency effect goes as follows. Given that homophone words inherit the frequency of their twins, then they should share the same lexeme node. In this way, each time the homophonic form is used, its frequency count is increased for both of its meanings. However, homophone words need to be represented at some linguistic level independently, since they have different semantic and grammatical information. This distinction is postulated at the lemma level.

The results reported here invalidate the first argument, since the observation upon which it is postulated appears to be wrong. We obtained word frequency effects in tasks that require access only to the grammatical properties of the target word. The second argument also appears to be problematic, since it is postulated on the basis of the homophone frequency effect, an effect that has been shown to be rather difficult to replicate (see Footnote 1). Thus, caution needs to be exercised when using these different frequency effects to argue in favour of the existence of a two-layer architecture in language production.[2]

Before concluding, two points need to be highlighted. First, our interpretation of the present results follows the rationale developed by Jescheniak and Levelt (1994). That is, we assume that the two tasks used here are not affected by the retrieval of phonological properties. However, this assumption may be wrong, and one may claim that these tasks are actually affected by the speed with which the phonological properties of the target words are retrieved. If that were to be the case, we should reconsider the theoretical implications of our study and also those of Jescheniack and Levelt's study. Furthermore, not only would these results be irrelevant to inform us about the locus of the frequency effect. We would also need to reconsider the conclusions reached by the influential study conducted by Van Turennout et al. (1998; and other similar investigations) in which the same rationale was used. Regardless of the results of future research bearing on the validity of this assumption, our observations reveal that the results obtained by Jescheniak and Levelt can no longer be used to argue about the locus of the frequency effect.

Secondly, it is important to raise the possibility that our observations and those of Finocchiaro and Caramazza (2006) contrast with Jescheniak and Levelt's (1994) results because of the different languages used in these studies. One may argue that while Dutch speakers can explore the grammatical properties of nouns, Spanish speakers cannot, hence the latter would produce their responses by exploring the nouns' phonological properties. However, it is unclear what is the theoretical motivation sustaining this cross-linguistic hypothesis. That is, although, taking this step would capture the contrasting effects in the gender decision, it is unclear what properties of Spanish prevent speakers from inspecting the grammatical gender in such language in the same way that Dutch speakers do.

To conclude, our results suggest that lexical selection is sensitive to the frequency with which a word is retrieved from the lexicon (Caramazza, 1997; Dell, 1990). This does not imply that word frequency affects only one level of representation. Actually, it is likely that the frequency with which any given representation is consulted affects the speed and reliability of its retrieval.

[2] This is not to say that there may not be other reasons to postulate a difference between lemma and lexeme levels of representation. A discussion of this issue is beyond the scope of the present article.

Original manuscript received 14 February 2006
Accepted revision received 31 March 2006

REFERENCES

Alario, F.-X., & Caramazza, A. (2002). The production of determiners: Evidence from French. *Cognition, 82*, 179–223.

Alario, F.-X., Costa, A., & Caramazza, A. (2002). Frequency effects in noun phrase production: Implications for models of lexical access. *Language and Cognitive Processes,17*, 299–319.

Bonin, P., & Fayol, M. (2002). Frequency effects in the written and spoken production of homophonic picture names. *European Journal of Cognitive Psychology, 14*, 289–314.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14*, 177–208.

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1430–1450.

Caramazza, A., & Hillis, A. E. (1990). Where do semantic errors come from? *Cortex, 26*, 95–122.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*, 283–321.

Dell, G. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes, 5*, 313–349.

Finocchiaro, C., & Caramazza, A. (2006). The production of pronominal clitics: implications for theories of lexical access. *Language and Cognitive Processes, 21*, 141–180.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments & Computers, 35*, 116–124.

Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language, 38*, 313–338.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 824–843.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

Levelt, W. J. M. (2002). Picture naming and word frequency: Comments on Alario, Costa and Caramazza. *Language and Cognitive Processes, 17*, 663–671.

Levelt, W. J. M., Praamstra, P., Meyer, A. S., Helenius, P., & Salmelin, R. (1998). A MEG study of picture naming. *Journal of Cognitive Neuroscience, 10*, 553–567.

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

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology, 17*, 273–281.

Starreveld, P. A., & La Heij, W. (2004). Phonological facilitation of grammatical gender retrieval. *Language and Cognitive Processes, 19*, 677–711.

Van Turennout, M., Hagoort, P., & Brown, C. M. (1998). Brain activity during speaking: From syntax to phonology in 40 milliseconds. *Science, 280*, 572–574.

Wingfield, A. (1968). Effects of frequency on identification and naming of objects. *American Journal of Psychology, 81*, 226–234.

APPENDIX

Pictures used in the experiments

Masculine		*Feminine*	
High-frequency names	*Low-frequency names*	*High-frequency names*	*Low-frequency names*
Coche (car)	Globo (hot air balloon)	Casa (house)	Canoa (canoe)
Barco (ship)	Trineo (sledge)	Iglesia (church)	Jarra (pitcher)
Tren (train)	Iglú (igloo)	Ventana (window)	Rana (frog)
Vaso (glass)	Castillo (castle)	Puerta (door)	Foca (seal)
Caballo (horse)	Candado (lock)	Boca (mouth)	Gorra (cap)
Perro (dog)	Pomo (doorknob)	Mano (hand)	Piña (pineapple)
Gato (cat)	Pez (fish)	Silla (chair)	Fresa (strawberry)
Pie (foot)	Collar (necklace)	Cama (bet)	Trompeta (trumpet)
Corazón (heart)	Jersey (pullover)	Mesa (table)	Tuerca (nut)
Reloj (watch)	Chaleco (vest)	Luna (moon)	Flecha (arrow)
Libro (book)	Limón (lemon)	Nube (cloud)	
Teléfono (telephone)	Violín (violin)	Caja (box)	
	Bate (baseball bat)		
	Pincel (paintbrush)		