Journal of Memory and Language 89 (2016) 37–54

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Linking recognition and production: Cross-modal transfer effects between picture naming and lexical decision during first and second language processing in bilinguals

Eva Van Assche [a,*], Wouter Duyck [a], Tamar H. Gollan [b]

[a] University of Ghent, Ghent, Belgium
[b] University of California, San Diego, USA

ARTICLE INFO

Article history:
Received 4 November 2014
revision received 26 February 2016
Available online 20 April 2016

Keywords:
Visual word recognition
Word production
Lexical processing

ABSTRACT

The present study examined the extent to which word production and recognition rely on shared representations in lexical access by examining cross-modality transfer effects and frequency effects in a training paradigm. Participants were trained in reading high- and low-frequency words in a lexical decision task and were subsequently tested in producing picture names and vice versa, both in their second (Experiment 1) and in their first language (Experiment 2). The same pattern of results was found for first and second language processing. Both tasks showed strong, within-modality repetition effects with faster responses and smaller frequency effects for repeated items. Training with repeated lexical decision, sped responses, and reduced the size of the frequency effects in subsequent picture naming. In contrast, training with repeated picture naming sped responses in lexical decision, but did not significantly decrease frequency effects. The results imply an amodal representation (lemma) that is shared between production and recognition and is not sensitive to word frequency. Also, they imply that a frequency sensitive phonological representation (lexeme) is activated automatically during visual word recognition.

Introduction

Oral language processing involves four basic functions: reading, speaking, listening, and writing. People use these functions to convey meaning in communication. The word *coat*, whether written or spoken aloud, refers to the same object and activates the same basic concepts. Comprehension and production are intrinsically linked to each other, but also involve different cognitive processes. In psycholinguistic research, these processes are often investigated separately (i.e., by different investigators, in separate research studies, and in separate sessions or even separate scientific meetings). To the limited extent that comprehension and production have been studied together in monolinguals, no clear consensus has emerged as to what extent shared representations and processes are involved (e.g., Dell & Gordon, 2003; Monsell, 1987; Roelofs, 2003). Similarly, in the bilingual domain, functional interactions between comprehension and production have rarely been investigated (but see Gollan et al., 2011).

The present study was designed to examine the extent to which production and recognition rely on shared representations in lexical access by examining cross-modality transfer and frequency effects in a training paradigm. Specifically, participants repeatedly read or produced high- and low-frequency words and then switched modalities in a test phase (in which they read the words they had trained with picture naming, and produced the words they had trained with reading). This training is an experimental

* Corresponding author.
E-mail address: eva.vanassche@ugent.be (E. Van Assche).

http://dx.doi.org/10.1016/j.jml.2016.02.003
0749-596X/

induction of additional language exposure. It has been argued that training effects in word production and visual word recognition can be explained by the same mechanisms as word frequency effects (e.g., Monsell, 1991; Wheeldon & Monsell, 1992). Thus, the main questions investigated were: will frequently reading a word later make it easier to produce the same word, and, similarly, will producing a word make it easier to later recognize that word in reading?

In addition to considering overall speed, we were interested in considering the size of the frequency effect, which is often considered to be a signature of lexical access (e.g., Almeida, Knobel, Finkbeiner, & Caramazza, 2007; Forster & Chambers, 1973; Levelt, Roelofs, & Meyer, 1999; Murray & Forster, 2004; Rayner, 1998). If recognition and production activate shared representations, cross-modal training effects should arise. Furthermore, if frequency sensitive lexical representations are accessed, the frequency effect should decrease in magnitude with training (e.g., Griffin & Bock, 1998; Scarborough, Cortese, & Scarborough, 1977), given that each additional exposure has a smaller effect on access speed (e.g., McCusker, 1977). Therefore low-frequency words benefit more from training than high-frequency words.

Production

In research on language production, there is general agreement that lexical access involves two major steps (e.g., Bock, 1987; Dell & O'Seaghdha, 1992; Levelt, 1989). The first step is the mapping of meaning onto an abstract representation of a word. The second step involves mapping this abstract representation onto the word's phonological characteristics. The distinction of two steps of lexical access is present in most models of speech production (e.g., Dell, 1986; Levelt et al., 1999; Rapp & Goldrick, 2000, but see Caramazza, 1997). A prominent model of lexical access in speech production is the WEAVER++ model (e.g., Levelt et al., 1999; Roelofs, 1992, 1997; Roelofs & Meyer, 1998). According to this model, speech production begins with the selection of a concept, after which lexical selection takes place with the retrieval of a syntactic representation (a lemma) from the mental lexicon. In subsequent processing steps, the word form is accessed so that morphological and phonological forms are activated. These phonological representations must be encoded to phonetic representations, which specify how the word should be articulated. In the final step, the phonetic plan is executed and the word is articulated. This multi-stage model adopts the spreading activation principle so that concepts and lemmas similar to the target also become activated and compete for selection. In cascading models (Dell, 1986; Rapp & Goldrick, 2000), such as the interactive two-step model of word production (Dell, 1986; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997), phonological encoding can begin before word selection is completed. In bilinguals, concepts activate lexical representations in the target language as well as in the non-target language (e.g., Colomé, 2001; Costa, Caramazza, & Sebastian-Galles, 2000; De Bot, 1992; Green, 1986; Hermans, Bongaerts, De Bot, & Schreuder, 1998). Representations and processes in bilingual production models are similar to those invoked in the monolingual models.

There is disagreement regarding the locus of the frequency effect in models of word production. Two-stage models such as WEAVER++ attribute frequency effects mainly to phonological encoding (e.g., Jescheniak & Levelt, 1994; Levelt et al., 1999), whereas cascade models assume that frequency effects arise in both word selection and phonological processing (e.g., Dell, 1990). There is much evidence in favor of the phonological-level locus of frequency effects (Dell, 1990; Jescheniak & Levelt, 1994). For instance, Jescheniak and Levelt (1994) studied the processing of high- and low-frequency homophones and showed that the frequency effect arises in accessing the word form (phonological retrieval) rather than the lemma. Similarly, studies of speech errors also support a phonological-level locus of the frequency effect (e.g., Dell, 1990).

However, although there is general agreement that a major locus of the frequency effect is phonological encoding, frequency effects do not necessarily need to be mutually exclusive arising *only* during phonological encoding in lexical access. There is evidence suggesting that frequency also affects lemma access (but note that some studies failed to find conclusive evidence for frequency-sensitive lemmas; e.g., Jescheniak & Levelt, 1994). For example, frequency affects grammatical gender decision to pictures suggesting frequency is represented at the level of grammatical encoding (e.g., Navarette, Basagni, Alario, & Costa, 2006). Other studies have also suggested multiple frequency-sensitive levels of lexical access (e.g., Gollan et al., 2011; Kittredge, Dell, Verkuilen, & Schwartz, 2008; Knobel, Finkbeiner, & Caramazza, 2008). Thus, it seems that frequency in word production models might be represented primarily in the second step of lexical access in which phonological encoding occurs (lexeme access), but also (though to a lesser degree) in the first step of lexical access where meaning is mapped to a lemma (e.g., Kittredge et al., 2008). Assuming that lemmas are shared between production and recognition (e.g., Levelt et al., 1999; for a different view see Caramazza, 1997), in the present study this would imply that training should both speed responses, and reduce the size of the frequency effect, in both training directions (recognition to production and vice versa). Such a result would suggest that the same representations (lemmas) are accessed in both production and recognition, and that these amodal representations are also frequency-sensitive.

Recognition

In the domain of visual word cognition, similar semantic and phonological representational levels as in production have been proposed to explain how readers derive meaning from printed words. In the dual-route theory of Coltheart, Rastle, Perry, Ziegler, and Langdon (2001), word recognition proceeds via two distinct, but interactive procedures: the lexical and non-lexical routes. In the lexical route, reading relies on the activation of whole-word orthographic and phonological representations of known words in the mental lexicon. These representations can directly activate semantic representations. Additionally, a non-lexical route involves a procedure that derives phonology by relying on a grapheme-to-phoneme conversion process. This route can process novel letter strings as well as regular words. The two routes are assumed to proceed in parallel with both excitatory and inhibitory connections between and within routes at the orthographic, phonological and semantic levels. A computational model of the dual-route theory is the Dual Route Cascaded model (Coltheart et al., 2001). The lexical orthographic part of the model is a generalization of the Interactive Activation model (Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981). Representations in the orthographic lexicon and in the phonological lexicon are frequency-sensitive (Coltheart et al., 2001).

In contrast to the dual-route architecture, connectionist models (e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989) assume a single procedure operating over distributed orthographic and phonological representations to read both regular and irregular words. In Plaut et al. (1996), word frequency effects result in weight changes that serve to reduce the error on that word and thus its activation time. In bilinguals, visual word recognition involves the activation of lexical and semantic representations in the mental lexicon that contains representations from both languages (e.g., Dijkstra & Van Heuven, 2002). Frequency in bilingual visual word recognition models is represented in the same way as in monolingual interactive activation models.

Even though phonology is not necessary for silent reading, there is extensive evidence that phonology is often activated rapidly and automatically even in tasks that require only visual word recognition and no pronunciation (see Frost, 1998, for a review). Many studies have shown that readers cannot ignore phonology when they read and that phonological representations are activated automatically and early during visual word recognition (e.g., Duyck, 2005; Grainger & Ferrand, 1996; Perfetti & Bell, 1991; Van Orden, 1987). However, in the case of visual word recognition in the lexical decision task, there is no clear consensus on whether phonological representations are automatically activated. Gerhand and Barry (1999) argue that both orthographic and phonological representations are activated in the lexical decision task, in a cascading fashion (e.g., Jared & Seidenberg, 1991; Monsell, 1991). These sources of information are all considered in the lexical decision process for which a particular level of lexical activation is set (e.g., Grainger & Ferrand, 1994). Orthographic representations will be activated before phonological ones so that orthographic information will first reach the threshold. It may be that for high-frequency words, orthographic activation is too fast for phonological information to have an effect. For low-frequency words, orthographic activation is slower, leaving more room for phonological information to exert an effect on lexical decision. Automatic phonological activation may also differ for deep and shallower orthographies (Frost, 1998). There is consistent evidence that phonology is more involved in lexical decisions in shallower orthographies (e.g., Feldman & Turvey, 1983).

The assumption of automatic activation of phonology during reading has important implications for the present study given that activation of phonology could provide the crucial link between modalities that may lead to substantial transfer of training effects from reading to speaking. Given the assumption that frequency effects arise primarily during access to phonology in speech production, frequency effects in picture naming might be expected to shrink after training with lexical decision if lexical decision activates these same phonological representations that are needed in production.

The relation between production and recognition

Although the domains of production and recognition have been mostly investigated separately, some studies have focused on the relationship between production and recognition (e.g., Barry, Hirsh, Johnston, & Williams, 2001; Liberman & Mattingly, 1985; Monsell, 1987; Roelofs, 2003). Monsell (1987) examined the relationship between speech comprehension and speech production considering evidence from cross-modal repetition priming studies. Participants first read words aloud and then switched to auditory recognition in an auditory lexical decision task of the same words that were produced earlier. Critically, these words were recognized more quickly than new words presented in the auditory recognition task. These results suggest that phonological representations are shared between recognition and production. Similarly, Barry et al. (2001) showed priming from naming printed words to naming pictures with the same names and argued that this effect arises at the stage of retrieval of the lexical phonological representations (i.e., the second step in the production models reviewed above). Studies investigating interference effects from auditory distractor words on picture naming times (e.g., Roelofs, Meyer, & Levelt, 1996) also found results pointing to overlapping word representations in production and recognition. Roelofs (1992) suggested that form information activated in speech recognition activated corresponding morpheme and shared lemma representations in the production system but specified separate output lexemes for production and input lexemes for speech recognition.

Evidence from dual-task studies (e.g., Shallice, McLeod, & Lewis, 1985) and neuropsychological studies reached the opposite conclusion, i.e., they showed evidence for separate representations for speech recognition and production. For instance, double dissociations between modalities have been observed in brain-damaged patients who have problems with auditory word comprehension but not with word production (e.g., Hillis, 2001). This implies at least partly separate systems for recognition and production, as well as modality-specific form representations. Additionally, some of the evidence cited in favor of shared representations is open to alternative interpretations. For example, the production task in Monsell (1987) was reading words aloud, which also contains a recognition component, so that the claimed cross-modal training effect may actually have been a within-modality training effect (from visual to auditory recognition). Barry et al. (2001) did use picture naming but primed it with another production task (reading words aloud), instead of a recognition task. In order to maximally dissociate production and recognition, the present study used visual lexical decision and picture naming, which does not entail word recognition at any processing stage.

Frequency in production and recognition

Both production and recognition in L1 and L2 exhibit strong frequency effects such that words that are used frequently are processed more quickly than words that are not used as often (e.g., Forster & Chambers, 1973; Rayner & Duffy, 1986). This is often explained by changes to the activation levels of representations and the connections between them. High-frequency words have higher levels of resting activation than low-frequency words. This relation between frequency and lexical access speed follows a logarithmic function (e.g., McCusker, 1977; Murray & Forster, 2004): a small increase in frequency of use has a large effect on lexical access time for low-frequency words but only a small effect for high-frequency words.

Frequency effects in recognition and production are typically larger in the non-dominant than in the dominant language (e.g., Duyck, Vanderelst, Desmet, & Hartsuiker, 2008; Gollan, Montoya, Cera, & Sandoval, 2008). This reflects the fact that words in L2 are used less often than words in the first language (L1) (i.e., the Frequency Lag hypothesis, Gollan et al., 2011; also known as Weaker Links, Gollan et al., 2008). Given the logarithmic function of frequency and access time, words in the monolingual lexicon are closer to ceiling levels of lexical accessibility than words in L1 (the native and first language), which in turn are closer to ceiling than words in L2 (the later learned second language) in the bilingual lexicon. For this reason, we first investigated the effects of recognition and production training in L2. The larger frequency effects in L2 and the expected larger effects of repetition (e.g., Francis, Augustini, & Sáenz, 2003) would increase the chances of finding cross-modal transfer effects.

Gollan et al. (2011) used the frequency effect to compare lexical access mechanisms in comprehension and production. Participants varying in the level of English proficiency (monolingual, early Spanish-English bilingual and late Dutch-English bilingual) were presented with high- and low-frequency words presented in isolation (i.e., single words), or in a low-constraint or in a high-constraint sentence. Lexical access in isolation was investigated using a picture naming and lexical decision task. The results showed robust frequency effects in both production and comprehension, but frequency effects were significantly larger in production than in comprehension, even when taking into account slower response times in speaking relative to reading. Gollan et al. considered a number of reasons why the frequency effect might be larger in production, but of particular relevance in the present context was the possibility that production entails more frequency sensitive processing stages than recognition (e.g., Kittredge et al., 2008; Knobel et al., 2008).

The present study

Participants in the current study were trained repeatedly (15 times in total) on three consecutive days in lexical decision and picture naming, with two different sets of high-and low-frequency target words. After cycling through 15 repetitions (5 presentations per day) of training, participants switched tasks and completed the same items they had practiced in one modality in the other modality (for both sets of items trained; i.e., in both tasks).

Within both tasks, we expected to observe significant training effects; i.e., that responses would become faster with repetition of the items (cf. Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Griffin & Bock, 1998). This may originate from accessing shared semantic concepts in both modalities leading to faster activation of semantics after training in the other modality. We also considered this to be an experimental manipulation of word frequency. Within the WEAVER ++ speech production model (e.g., Levelt et al., 1999; Roelofs & Meyer, 1998) and the Dual Route Cascaded recognition model (Coltheart et al., 2001) discussed in the introduction, differences in word frequency are assumed to reflect changes in baseline activation of lexical representations. In production, evidence has accumulated that the main locus of frequency effects is the phonological level, but the lemma representational level is also frequency-sensitive (e.g., Gollan et al., 2011; Kittredge et al., 2008). In recognition, frequency effects arise from orthographic and phonological levels (Coltheart et al., 2001). We expected that within-modality training effects would be stronger for low- than for high-frequency targets and thus that the frequency effect would decrease in magnitude with training in production and recognition (e.g., Forster & Davis, 1984; Gollan et al., 2011; Griffin & Bock, 1998; Scarborough et al., 1977; but see Levelt et al., 1999).

When considering the cross-modal transfer effects, the main question is which representations are shared between modalities allowing cross-modal transfer (a) in terms of overall speeding of reaction times and (b) affecting the size of the frequency effects. It is important to first spell out which representations are activated during the picture naming and lexical decision tasks. Fig. 1 presents the different processing components for speech production and recognition.

In production, speaking involves (1) the retrieval of a concept, (2) the activation of a lemma, (3) the activation of morpho-phonological representations, and (4) the retrieval of a phonetic representation (e.g., Indefrey & Levelt, 2004; Levelt et al., 1999). Note that orthographic representations may also become active during word production (Rastle, McCormick, Bayliss, & Davis, 2011). However, at this moment, there is no conclusive evidence for the involvement of orthography in speech production (e.g., Damian & Bowers, 2009).

In recognition, it is clear that upon presentation of a word, orthographic representations are activated and,

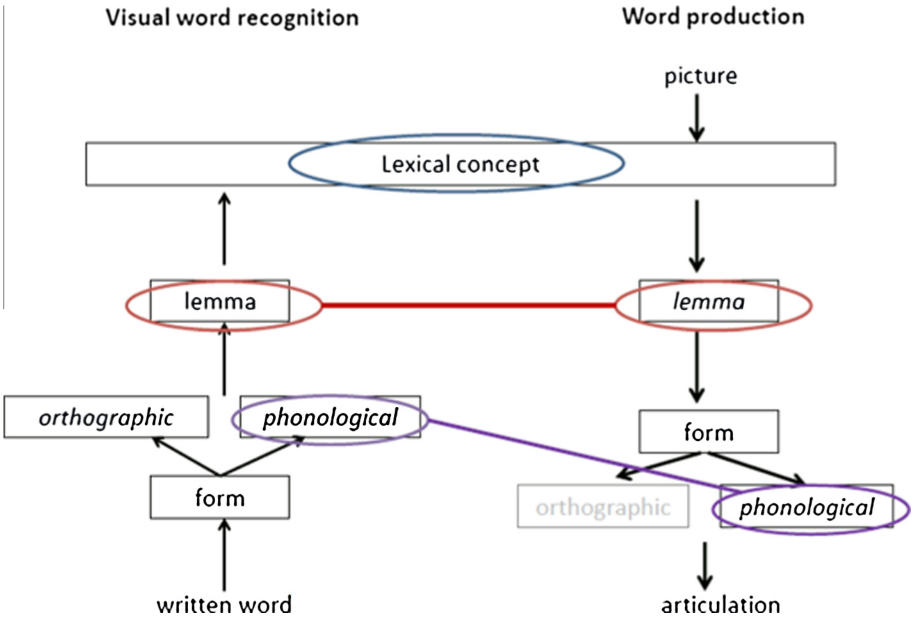

Fig. 1. Network of processing levels involved in visual word recognition and word production. Different connections across modalities indicate possible shared representations across modalities. Representational levels printed in italics can be frequency-sensitive.

depending on the task requirements, lead to conceptual activation (e.g., Coltheart et al., 2001). There is also much evidence for the involvement of phonology in visual word recognition (e.g., Frost, 1998; Grainger & Holcomb, 2009). However, though lemmas have played a central role in theories of speech production, and Levelt et al. (1999) assumed that lemmas are shared between production and recognition, lemmas (i.e., amodal lexical representations) are rarely used in models of lexical access in recognition. Nonetheless, lemma activation may be relevant in reading (e.g., Indefrey & Levelt, 2004).

Across tasks, cross-modal transfer effects in terms of the overall processing speed can arise at several representational levels. Given the assumption of shared lemmas between production and recognition (e.g., Levelt et al., 1999; for a different view see Caramazza, 1997), this would predict that training in the other modality should speed responses. Such a result would suggest that the same representations (lemmas) are accessed in both production and recognition. In addition, Fig. 1 shows that transfer effects may also arise from access to shared, or at least closely linked, conceptual and phonological representations for both production and recognition. However, note that conceptual and phonological activation is not necessarily involved in lexical decision. Lexical decisions may be based more quickly solely on the orthographic code. This would predict that activation of shared, or at least closely linked, conceptual, lemma and phonological representations in production after recognition training will lead to overall speeding of reaction times. However, speeding should be smaller in recognition after production training because lexical decision does not necessarily require access to conceptual, lemma and phonological representations.

Predictions for the transfer effect in terms of the size of the frequency effect depend on whether the representations that are shared between recognition and production are frequency-sensitive representations. If recognition involves access to frequency-sensitive lemma and phonological representations activated during production training, then a decrease in frequency effect after training with production would be expected to transfer to recognition. However, given that the main locus of frequency effects is the phonological level in production (e.g., Jescheniak & Levelt, 1994), and given that lexical decision does not necessarily involve access to phonological representations, transfer in terms of the size of the frequency effect from production to recognition may be weak.

We predicted that cross-modal transfer in terms of the size of the frequency effect should be strongest going from lexical decision to production. Written word recognition likely involves automatic activation of phonology (e.g., Frost, 1998) and phonological representations provide the main locus of frequency effects in production. Thus, frequency effects in picture naming should shrink significantly after training with lexical decision.

Experiment 1: Second language processing

In Experiment 1, we investigated transfer effects across modalities in second language processing. We reasoned that L2 processing might provide optimal conditions for cross-modal transfer effects to occur because L2 would leave more room for improvement than L1. Word processing in L2 is generally slower (e.g., Duyck, 2005; Duyck et al., 2008; Gollan et al., 2008) and further from ceiling levels of performance compared to the more proficient L1 (e.g., Gollan et al., 2005; Hanulová, Davidson, & Indefrey, 2011; Runnqvist, Strijkers, Sadat, & Costa, 2011).

Method

Participants

Twenty-three students from Ghent University participated in the English training protocol. The data from 19 subjects were analyzed. Three subjects had to be removed from the data because of insufficient L2 knowledge (they knew less than half of the low frequency items) and one participant had to be removed because of corrupted E-prime results files. They were all late Dutch-English bilinguals who were exposed to English at an early age through popular media (e.g., music, movies, etc.) and also learned English at secondary school around age 14 for about 3–4 h a week. Additionally, their university studies required them to read university textbooks in their L2 (English). Participants were paid or received course credit for their participation. The criteria for recruitment stipulated that the participants should have good knowledge of English. They completed two language proficiency tests: the LexTALE (Lemhöfer & Broersma, 2011) in Dutch and English and a self-report language questionnaire in which they were asked to rate their L1 and L2 proficiency with respect to several skills (reading, speaking and general proficiency) on seven-point Likert scales ranging from *very bad* to *very good*. Detailed scores on all proficiency measures are reported in Table 1.

Materials

The stimulus materials consisted of 80 black-and-white line-drawn pictures from Gollan et al. (2011), Snodgrass and Vanderwart (1980), and Severens, Van Lommel, Ratinckx, and Hartsuiker (2005) for the picture naming task and the English word forms of these pictures for the lexical decision task. None of the items were cognates or interlingual homographs. Half of the items had high-frequency names and half low-frequency names. The items were divided into two lists with each list containing 20 high-frequency and 20 low-frequency items. High- and low-frequency items were matched for name agreement using the mean H-statistic from Severens et al. which is an index of name agreement in Dutch. Means across the two sets of high- and low-frequency words in the two lists did not differ ($ps > .19$; *M* high-frequency words list 1 = .39; *M* low-frequency words list 1 = .69; *M* high-frequency words list 2 = .52; *M* low-frequency words list 2 = .52). The high- and low-frequency items of lists 1 and 2 are presented in Appendix A.

To enable a within-subjects test of cross-modal transfer in both directions, one list was presented during training in the lexical decision task and subsequent testing in the picture naming task, while the other list was used for training in picture naming and for subsequent testing in lexical decision. This way, the results from the first presentation of the items in one modality served as the baseline to compare with the transfer results after training. We matched the high-and low-frequency items of one list and the high- and low-frequency items of the other list on number of letters, logfrequency in the CELEX database (Baayen, Piepenbrock, & Van Rijn, 1993), Subtlex zipf frequency based on film subtitle frequencies (Brysbaert & New, 2009), and number of word neighbors ($ps > .25$). Logfrequency per million and Subtlex zipf frequency differed significantly between frequency conditions within languages ($ps < .001$). Furthermore, high- and low-frequency items within each list were matched on three other variables influencing lexical decision and picture naming times: number of letters/phonemes, neighborhood size and summated bigram frequency (calculated using the WordGen stimulus generation program, Duyck, Desmet, Verbeke, & Brysbaert, 2004) ($ps > .11$). Table 2 is a summary of the item characteristics.

Additionally, for each lexical decision block in which 20 high- and 20 low-frequency words were presented, we selected 40 orthographically regular and pronounceable nonwords. Given that the lexical decision block was repeated 15 times over three days and that there was a lexical decision task in the cross-modal test, this resulted in 640 nonwords that were matched to the high- and low-frequency target words on length, neighborhood size, and bigram frequency ($ps > .76$). Each of the 15 lexical decision and cross-modal test lexical decision blocks contained different nonwords. By selecting different nonwords in each lexical decision block, the nonwords did not become familiar to the participants, and this variable could not

Table 1
Participant characteristics in Experiments 1 and 2.

Language	Characteristic	Experiment 1	Experiment 2	*p*
L1 (Dutch)	Age in years	21.3 (2.5)	21.4 (2.9)	.52
	Lextale score[a]	92.6 (4.3)	90.8 (8.2)	.40
	Self-rated speaking[b]	6.7 (0.5)	6.4 (0.6)	.19
	Self-rated reading[b]	6.4 (0.8)	6.7 (0.6)	.18
	Self-rated general Proficiency[b]	6.3 (0.6)	6.5 (0.6)	.85
L2 (English)	Lextale score[a]	83.7 (9.7)	79.5 (7.6)	.18
	Self-rated speaking[b]	5.5 (0.8)	5.5 (0.8)	.75
	Self-rated reading[b]	5.8 (0.7)	6.2 (0.8)	.67
	Self-rated general Proficiency[b]	5.5 (0.7)	5.6 (0.8)	.67

Note. Standard deviations are displayed in parentheses.
[a] Proficiency score in percentage.
[b] Proficiency level based on self-ratings on a scale ranging from 1 (very bad) to 7 (very good). *P*-values were taken from independent samples *t*-tests testing mean differences across experiments.

Table 2
Mean lexical characteristics of the high- and low-frequency items in L1 and L2.

Language		List 1		List 2		*ps*
		HF (*n* = 20)	LF (*n* = 20)	HF (*n* = 20)	LF (*n* = 20)	
L2 English	Letters	4.85 (1.14)	5.60 (1.76)	4.90 (1.45)	5.50 (1.79)	>.12
	CELEX logfreq	1.82 (0.43)	0.80 (0.44)	1.82 (0.37)	0.85 (0.28)	<.001
	Subtlex Zipf	4.62 (0.38)	3.67 (0.59)	4.67 (0.44)	3.87 (0.52)	<.001
	N	5.50 (6.03)	3.20 (4.10)	5.60 (6.33)	3.85 (4.70)	>.16
L1 Dutch	Letters	4.70 (1.38)	5.45 (1.85)	4.90 (2.22)	5.50 (1.79)	>.15
	CELEX logfreq	1.76 (0.45)	0.76 (0.36)	1.75 (0.41)	0.80 (0.26)	<.001
	Subtlex Zipf	4.55 (0.37)	3.63 (0.30)	4.60 (0.54)	3.67 (0.28)	<.001
	N	7.15 (6.10)	4.65 (4.89)	8.20 (7.15)	5.35 (6.18)	>.16

Note. Standard deviations are indicated in parentheses. *P*-values were taken from paired samples *t*-test testing mean differences between high-frequency and low-frequency words within lists.

change the nature of word/nonword decisions with repeated training in the lexical decision task.

Procedure

Participants came to the lab on three consecutive days. They were tested individually for 50 min on the first day, 30 min on the second day, and 50 min on the third day. On day 1, half of the participants performed a lexical decision task in L2 on the set of high-and low-frequency words in list 1. They repeated this task five times. The same participants also completed a picture naming task in L2 on the set of high- and low-frequency pictures in list 2, and they repeated this task five times as well. The other half of the participants named the list 1 pictures five times and performed the lexical decision on list 2 stimuli five times. The order of task presentation was counterbalanced across participants and presentation of the pictures was fully randomized within each block.

After the first lexical decision task of day 1, participants were asked to translate the English high- and low-frequency target words to Dutch. The results showed that participants knew almost all the target words (accuracy $M = 0.97$). The large majority (92%) of the unknown words were low-frequency words (e.g., *skunk*, *tweezers*, *snail*, *peacock*). When the translation test was completed, the Dutch translation of unknown words was taught to the participants. This first lexical decision task and the translation test constituted the familiarization phase.

Similarly, after the picture naming task of day 1, participants looked at each picture with the experimenter and named each picture. This was done to check whether participants knew all the names in L2. All participants knew at least 90% of the picture names. The remaining 10% of the pictures were mostly low-frequency items (e.g., *kite*, *skunk*). Names participants did not know were provided by the experimenter.

After this familiarization phase, the training phase started in which participants were presented with the pictures four more times so that RTs were measured five times on day 1. We presented a second block of 40 trials, after which a double block of 80 trials followed. Finally, there was another block of 40 trials in which the pictures were presented for the fifth time. Participants could take a short break between blocks. The same presentation scheme was applied in the lexical decision training.

On the second and third day of the training phase, all items from lists 1 and 2 were again presented five times using the same lexical decision and picture naming tasks as on the first day. On the third day, after training, participants completed the transfer block, which was the cross-modal test. The results from the first presentation of the items of one list in the lexical decision task served as the baseline for comparison with the lexical decision results of the other list after being trained with picture naming (i.e., the cross-modal test). Similarly, the first presentation results of the items of one list in the picture naming task served as the baseline for comparison with the picture naming results of the other list after being trained with lexical decision. For instance, on day 1, participant A performed a lexical decision task on the words of list 1 and a picture naming task on the words of list 2. These results served as the baseline for frequency effects because they show the untrained reaction times for these items in reading and production.

The materials for the picture naming task were presented using E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA). Participants were instructed to name the pictures as quickly and accurately as possible. They were encouraged to speak clearly and to give their answer right away without saying "um". If they did not know the name of the picture, they could say "pass". Participants' responses were collected using a voice key (Duyck et al., 2008). An experimenter was present during the experiment for verification of the accuracy of the responses and the accuracy of voice key triggers. We audiotaped the responses for later verification. Trials began with a central fixation cross presented for 500 ms. After a 300 ms interstimulus interval, the picture was presented in the center of the computer screen. Each picture remained on the screen until the participant responded. The next trial began 1500 ms later. On each test day, 10 practice trials of five high- and five low-frequency pictures different from the experimental pictures preceded the experiment in which the set of pictures was presented five times.

For the lexical decision task, materials were presented using the same E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA) presentation software. Participants were instructed to decide on each trial whether or not the presented letter string was a real English word by pressing one of two response buttons. They were instructed to press the right button for a word response and the left button for a nonword. It was emphasized that it was important to make this decision as quickly and accurately as possible. Each trial started with the presentation of a fixation cross in the center of the screen for 800 ms. After a 300 ms interstimulus interval, the letter string was presented centrally until the participant responded. The intertrial interval was 700 ms. The first lexical decision task of each day was preceded by 10 practice trials of five high- and five low-frequency words not presented in the actual experiment.

Results

Mean reaction times (RTs) were calculated for each participant and frequency condition (high- vs. low-frequency) at each repetition in the familiarization and training phases and in the transfer block. Errors (i.e., pass-answers or wrong naming of the picture), voice key malfunctions (i.e., verbal disfluencies such as coughs or hesitations), and outliers were not included in the RT analyses. In the picture naming and lexical decision familiarization and training phases, voice key malfunctions (11.0%) and errors in picture naming (2.0%) and incorrect responses in lexical decision (2.9%) were not included in the RT analyses. In the picture naming task after training with lexical decision, 3.2% errors and 9.1% voice key malfunctions were excluded from the RT analyses. In the lexical decision task after training with picture naming, errors were excluded from the RT analyses (4.6% of the data). RTs more than 2.5 standard deviations above each participant's mean RT for high- and low-frequency items were excluded from analyses. This procedure led to an overall trimming of 2.4% of picture naming RTs and 2.5% of lexical decision RTs.

In Fig. 2, two graphs (one for each task) present the RTs in the familiarization and training phases and in the transfer block after training. For reasons of clarity, instead of presenting all 15 repetitions, we included six repetitions in the training phase (the second and fifth block on day 1, and the first and fifth blocks of days 2 and 3, i.e., repetition numbers 2, 5, 6, 10, 11, and 15). Note that the

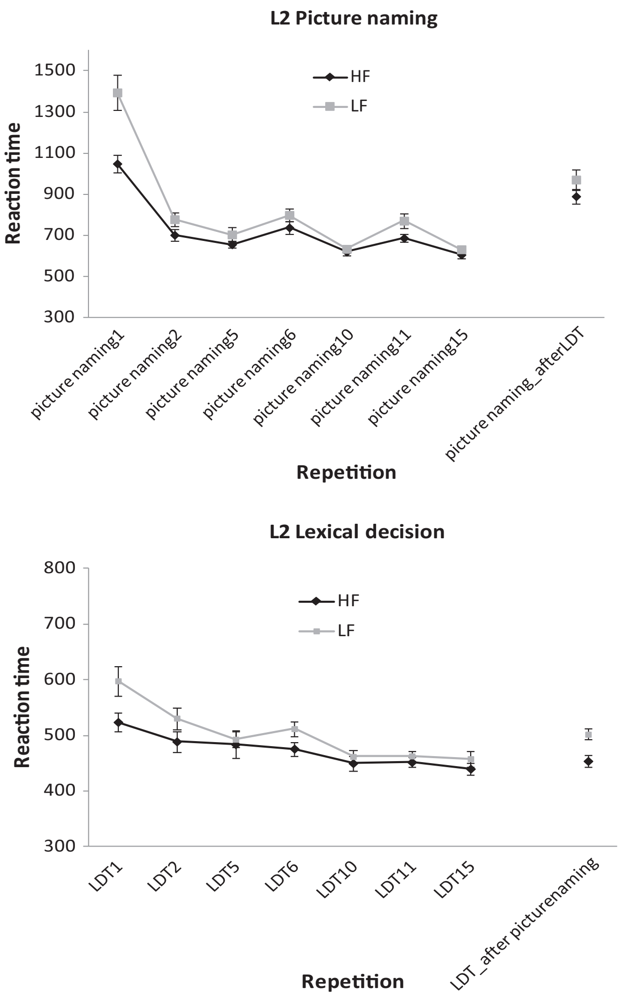

Fig. 2. Picture-naming (top panel) and lexical-decision times (lower panel) in milliseconds in L2 for high- and low-frequency words on the first, second, and fifth presentations (on day 1), sixth, tenth (on day 2), eleventh and fifteenth presentations (on day 3) for the cross-modal transfer list. Error bars are standard errors for each condition. LDT = lexical decision task, HF = high-frequency, LF = low-frequency.

familiarization, training, and transfer blocks in each graph involved reaction times to the same task.

Familiarization and training phase

As shown in Fig. 2 and Table 3, the familiarization phase in which participants were first presented with the items and subsequently performed a word knowledge test led participants to respond more quickly and with fewer errors. In both lexical decision and picture naming tasks, participants were faster in the second block than in the first block and more so for low-frequency than for high-frequency items. After the initial facilitation of response times in the familiarization phase, the RTs and frequency effects in the training phase showed further small decreases with repeated presentations.

Cross-modal transfer

The picture naming and lexical decision data were submitted to a 2 × 2 × 2 ANOVA with Task (picture naming, lexical decision), Repetition (first presentation, cross-modal test) and Frequency (high, low) as repeated measures factors in the participant analyses (F_1). In the items analyses (F_2), Repetition and Task were repeated measures factors and Frequency was a between-items factor. Participants were faster in the lexical decision than in the picture naming task [$F_1(1,18) = 172.58$, $MSE = 71{,}161$, $p < .001$; $F_2(1,77) = 841.02$, $MSE = 30{,}364$, $p < .001$; min $F'(1,26) = 143.20$, $p < .001$] and were faster on high-frequency than on low-frequency items [$F_1(1,18) = 43.82$, $MSE = 19{,}888$, $p < .001$; $F_2(1,77) = 42.69$, $MSE = 35{,}522$, $p < .001$; min $F'(1,58) = 21.88$, $p < .001$]. RTs after training in the other modality were faster than RTs on the first presentation of the items as shown by the main effect of Repetition [$F_1(1,18) = 70.04$, $MSE = 19{,}530$, $p < .001$; $F_2(1,77) = 157.96$, $MSE = 18{,}740$, $p < .001$; min $F'(1,36) = 48.52$, $p < .001$]. This transfer effect was stronger in picture naming than in lexical decision, as shown by a significant Task × Repetition interaction [$F_1(1,18) = 33.54$, $MSE = 13{,}014$, $p < .001$; $F_2(1,77) = 52.34$, $MSE = 17{,}647$, $p < .001$; min $F'(1,44) = 20.44$, $p < .001$]. Frequency effects were stronger in picture naming than in lexical decision, as indexed by a significant Task × Frequency interaction [$F_1(1,18) = 18.18$, $MSE = 17{,}406$, $p < .001$; $F_2(1, 77) = 15.69$, $MSE = 30{,}364$, $p < .001$; min $F'(1,64) = 8.42$, $p < .01$]. Transfer effects were stronger for low-frequency than for high-frequency items, indicated by a significant Repetition × Frequency interaction [$F_1(1,18) = 18.69$, $MSE = 12{,}662$, $p < .001$; $F_2(1,77) = 25.79$, $MSE = 18{,}740$, $p < .001$; min $F'(1,49) = 10.84$, $p < .01$]. These task, frequency, and transfer effects interacted in a significant Task × Repetition × Frequency interaction [$F_1(1,18) = 15.09$, $MSE = 11{,}155$, $p < .001$; $F_2(1,77) = 18.27$, $MSE = 17{,}647$, $p < .001$; min $F'(1,52) = 8.26$, $p < .01$] showing that cross-modal transfer effects and frequency effects were stronger in picture naming than in lexical decision. Given the significant interactions with Task, we submitted the picture naming and lexical decision data separately to 2 × 2 ANOVAs with the factors Repetition and Frequency. The results of these analyses are presented below, first for picture naming and then for lexical decision. Average RTs per task are shown in Fig. 2.

Table 3
Mean error percentages for picture naming and lexical decision in L2 for high- and low-frequency words on the first (on day 1), second to fifteenth presentations (on days 1–3), and for the cross-modal transfer list. LDT = lexical decision task, HF = high-frequency, LF = low-frequency.

Task		HF	LF
Picture naming	Picture naming1	3.50 (0.05)	17.5 (0.11)
	Picture naming2–15	0.20 (0.02)	2.08 (0.05)
	Picture naming after LDT	2.80 (0.05)	4.20 (0.08)
LDT	LDT1	0.26 (0.01)	7.11 (0.07)
	LDT2–15	2.37 (0.05)	3.08 (0.04)
	LDT after picture naming	2.40 (0.04)	6.80 (0.04)

Note. Standard deviations are indicated in parentheses.

Picture naming

Speakers named pictures with high-frequency names faster than pictures with low-frequency names [$F_1(1,18) = 31.30$, $MSE = 35{,}749$, $p < .001$; $F_2(1,77) = 29.68$, $MSE = 63{,}150$, $p < .001$; min $F'(1,60) = 15.23$, $p < .001$]. Picture naming times were significantly and substantially faster after training with lexical decision, a main effect of Repetition [$F_1(1,18) = 58.93$, $MSE = 28{,}419$, $p < .001$; $F_2(1,77) = 104.07$, $MSE = 34{,}547$, $p < .001$; min $F'(1,41) = 37.63$, $p < .001$]. Also, the frequency effect was significantly smaller in picture naming after training with lexical decision than it was in picture naming without prior cross-modal training as shown by a significant Frequency × Repetition interaction [$F_1(1,18) = 17.94$, $MSE = 22{,}414$, $p < .01$; $F_2(1,77) = 23.08$, $MSE = 34{,}547$, $p < .001$; min $F'(1,50) = 10.09$, $p < .01$]. However, planned comparisons revealed that both high-frequency [$F_1(1,18) = 24{,}94$, $MSE = 8{,}733$, $p < .001$; $F_2(1,77) = 14.75$, $MSE = 34{,}547$, $p < .001$; min $F'(1,78) = 9.27$, $p < .01$] and low-frequency pictures [$F_1(1,18) = 44.16$, $MSE = 42{,}099$, $p < .001$; $F_2(1,77) = 111.18$, $MSE = 34{,}547$, $p < .001$; min $F'(1,34) = 31.61$, $p < .001$] were produced more quickly after training with lexical decision than without training. Mean RTs and size of the cross-modal and frequency effects are presented in Table 4.

Analyses of errors (incorrect answers and pass-answers) for picture naming showed fewer errors for high- than for low-frequency targets [$F_1(1,18) = 49.49$, $MSE = 0.01$, $p < .001$; $F_2(1,78) = 10.37$, $MSE = 0.02$, $p < .001$; min $F'(1,96) = 8.57$, $p < .01$]. As in the RT analyses, the frequency effect was smaller in picture naming after training with lexical decision [$F_1(1,18) = 29.66$, $MSE = 0.004$, $p < .01$; $F_2(1,78) = 12.29$, $MSE = 0.01$, $p < .001$; min $F'(1,89) = 8.69$, $p < .01$], and participants produced fewer errors after cross-modal training [$F_1(1,18) = 8.77$, $MSE = 0.01$, $p < .01$; $F_2(1,78) = 16.80$, $MSE = 0.01$, $p < .001$; min $F'(1,39) = 5.76$, $p < .05$]. Overall, error rates were very low, especially after the first presentation of the items when participants were told the names of unknown items. The mean error percentages in picture naming are presented in Table 3.

Lexical decision

Lexical decision times were faster for high-frequency than for low-frequency words [$F_1(1,18) = 44.54$, $MSE =$

Table 4
Size of the cross-modal effects for reaction time (in ms), frequency (in ms) and percentage adjusted frequency effect in picture naming and lexical decision in Experiments 1 and 2.

		L2		L1	
		Picture naming	Lexical decision	Picture naming	Lexical decision
Mean RT	Presentation 1	1217 (238)	560 (89)	795 (111)	494 (62)
	Transfer task	936 (176)	478 (39)	745 (109)	446 (51)
	Transfer effect	*281*	*82*	*50*	*48*
Frequency effect	Presentation 1	388	73	104	39
	Transfer task	97	48	27	28
	Transfer effect	*291*	*25*	*77*	*11*
Percentage adjusted FE	Presentation 1	30%	12%	14%	8%
	Transfer task	10%	10%	4%	7%
	Transfer effect	*20%*	*2%*	*10%*	*1%*

Note. Standard deviations are indicated in parentheses.

1,545, $p < .001$; $F_2(1,78) = 58.80$, $MSE = 2{,}728$, $p < .001$; min $F'(1,49) = 25.34$, $p < .001$], and faster after training with picture naming than on the first presentation on day 1 [$F_1(1,18) = 31.39$, $MSE = 4{,}125$, $p < .001$; $F_2(1,78) = 160.70$, $MSE = 1{,}816$, $p < .001$; min $F'(1,25) = 26.26$, $p < .001$]. Cross-modal training did not decrease the size of the frequency effect in lexical decision [$F_1(1,18) = 2.10$, $MSE = 1{,}403$, $p = .17$; $F_2(1,78) = 4.50$, $MSE = 1{,}816$, $p < .05$; min $F'(1,37) = 1.43$, $p = .24$]. Planned comparisons showed that this cross-modal transfer effect was present for both high-frequency [$F_1(1,18) = 43.27$, $MSE = 1{,}082$, $p < .01$; $F_2(1,78) = 55.70$, $MSE = 1{,}816$, $p < .001$; min $F'(1,50) = 24.35$, $p < .001$] and low-frequency words [$F_1(1,18) = 19.24$, $MSE = 4{,}445$, $p < .01$; $F_2(1,78) = 109.50$, $MSE = 1{,}816$, $p < .001$; min $F'(1,25) = 16.36$, $p < .01$]. Mean RTs and size of the cross-modal and frequency effects are presented in Table 4.

There were fewer errors for high- than for low-frequency words [$F_1(1,18) = 38.34$, $MSE = 0.002$, $p < .001$; $F_2(1,78) = 16.62$, $MSE = 0.01$, $p < .001$; min $F'(1,88) = 11.59$, $p < .01$]. Cross-modal training did not influence the number of errors in lexical decision [$F_1(1,18) = 1.27$, $MSE = 0.001$, $p = .27$; $F_2(1,78) = 0.76$, $MSE = 0.004$, $p = .38$; min $F' < 1$] nor the size of the frequency effect in lexical decision [$F_1(1,18) = 1.31$, $MSE = 0.002$, $p = .27$; $F_2(1,78) = 1.17$, $MSE = 0.004$, $p = .28$; min $F' < 1$]. Overall, error rates were very low, especially after the first presentation of items. The mean error percentages in lexical decision are presented in Table 3.

Proportionally adjusted frequency effects

Above, we reported cross-modal transfer effects in overall response speed for both production and recognition. However, the frequency effect became smaller only in production but not in recognition. However, it is difficult to compare responses across the two tasks because picture naming times are much slower than lexical decision times. Therefore, to test whether the apparent differences in the magnitude of cross-modal transfer effects might reflect between-task differences in baseline response times we calculated proportionally adjusted frequency effects for each participant in the first presentation during training and in the cross-modal test as follows: (low frequency – high frequency)/[(low frequency + high frequency)/2]. We then analyzed these proportional frequency values in a 2 × 2 ANOVA with Modality (production, recognition) and Repetition (first presentation, cross-modal test) as repeated factors.

Fig. 3 shows the results of this analysis and proportionally adjusted frequency effects are presented in Table 4. Proportionally adjusted frequency effects were smaller in recognition than in production (a main effect of Modality) [$F(1,18) = 5.71$, $MSE = 0.02$, $p < .05$], and smaller in the cross-modal test compared to first presentation (a main effect of Repetition) [$F(1,18) = 17.11$, $MSE = 0.01$, $p < .001$].

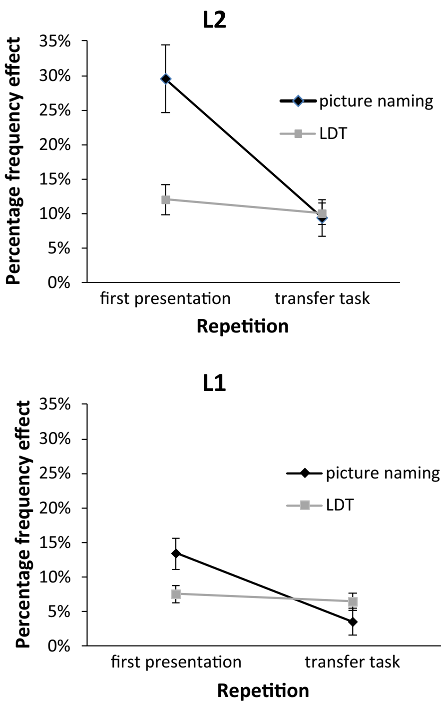

Fig. 3. The size of the frequency effect in the first presentation and the cross-modal test in picture naming and lexical decision in L2 (upper panel) and L1 (lower panel) as a proportion of baseline response times.

Of greatest interest, cross-modal training facilitated picture naming much more than lexical decision, as indexed by a significant interaction between Modality and Repetition [$F(1,18) = 10.70$, $MSE = 0.01$, $p < .01$]; there was a decrease in the size of the proportionally adjusted frequency effect in production [$F(1,18) = 17.69$, $MSE = 0.02$, $p < .001$], but not in recognition [$F(1,18) = 0.62$, $MSE = 0.01$, $p = .43$].

Discussion

Picture naming and lexical decision times exhibited strong training effects (i.e., responses became faster with repetition). Frequency effects decreased with repeated presentation of the items (e.g., Gollan et al., 2005; Griffin & Bock, 1998) and this decrease was stronger in picture naming than in lexical decision. Looking at cross-modal transfer, responses in picture naming after training with lexical-decision were faster and frequency effects were smaller compared to responses without any prior training. Similarly, training in production led to faster responses but did not significantly change frequency effects in lexical decision (although this effect was significant in the items analysis). An analysis on proportionally adjusted frequency effect confirmed the apparent difference between tasks in the nature of cross-modal transfer effects. Both high- and low-frequency targets benefited from cross-modal training, but only in production, low-frequency targets benefitted more from cross-modal training than high-frequency targets.

Experiment 2: First language processing

In Experiment 1, we obtained clear cross-modal transfer effects that interacted with task characteristics. In Experiment 1, we assessed whether the same pattern of results would be found in the dominant language, which is likely to show faster reaction times. Our analysis of proportional frequency effects in Experiment 1 led us to conclude that the difference between tasks in cross-modal transfer was not an artifact of differences in baseline response times but a difference in the degree of overlapping representations between recognition and production. Thus, we expected that Experiment 2 would replicate the results of Experiment 1, but also that training and transfer effects should be smaller because words in L1 are closer to ceiling levels of lexical accessibility.

Method

Participants

Twenty-four students from Ghent University who did not participate in Experiment 1 were recruited for participation in the Dutch experiment. They were all late Dutch-English bilinguals and completed the same two language proficiency tests as the bilinguals in Experiment 1. There was no difference in mean general L1 and L2 proficiency for the participants of Experiments 1 and 2 (independent samples *t*-test yielded *ps* > .18). Participant characteristics are reported in Table 1.

Materials

The materials were the same 80 pictures as used in Experiment 1. The Dutch names of these pictures were used in the L1 lexical decision task. The division and matching of the high- and low-frequency items within and across the two lists was the same as in Experiment 1 (*ps* > .15). As in Experiment 1, logfrequency and Subtlex zipf frequency differed significantly between frequency conditions in Dutch (*ps* < .001). A summary of the stimulus characteristics is given in Table 2. For the Dutch lexical decision tasks, 640 nonwords were selected which were matched to the high-and low-frequency target words on length, neighborhood size, and bigram frequency (*ps* > .38). The high- and low-frequency items of lists 1 and 2 are presented in Appendix A.

Procedure

The procedure was the same as in Experiment 1, but both tasks were completed in L1 (instead of in L2). The word knowledge tests after the first lexical decision and picture naming task of day 1 were similar to the word knowledge tests in Experiment 1. Given that this experiment was in L1, we anticipated that participants would probably know all the words but to keep the procedure comparable to Experiment 1, we tested knowledge of the targets in the same way. The test involved translating the words from Dutch to English after the first lexical decision task and naming the pictures in Dutch after the first picture naming task. The results from the word knowledge test for picture naming showed that participants knew all the targets in their first language. They also knew all the Dutch targets in the lexical decision task and could translate most of the words to English (*M* = .93 correct translations).

Results and discussion

Mean RTs were calculated for each participant and frequency condition (high- vs. low-frequency) at each repetition. Incorrect responses in the lexical decision task (4.2% overall), errors in the picture naming training task (i.e., pass-answers and wrong naming of the picture) (0.8%) and voice key malfunctions in the familiarization and training phases (12.4%) were not included in the RT analyses. In the lexical decision task after training with picture naming, errors were excluded from the RT analyses (5.6% of the data). In the picture naming task after training with lexical decision, errors (2.6%) and voice key malfunctions (12.8%) were excluded from the RT analyses. Data trimming procedures were the same as in Experiment 1 and resulted in the exclusion of 4.4% of the picture naming RTs and 2.8% of the lexical decision RTs. As for the analyses of Experiment 1, we ran a $2 \times 2 \times 2$ ANOVA with the factors Task, Repetition and Frequency. Next, we ran separate 2×2 ANOVAs on the lexical decision and picture naming data to investigate the cross-modal transfer and frequency effects separately for each task. Fig. 4 shows two graphs (one for each task) with the average RTs in the familiarization and training phases and in the transfer block after training.

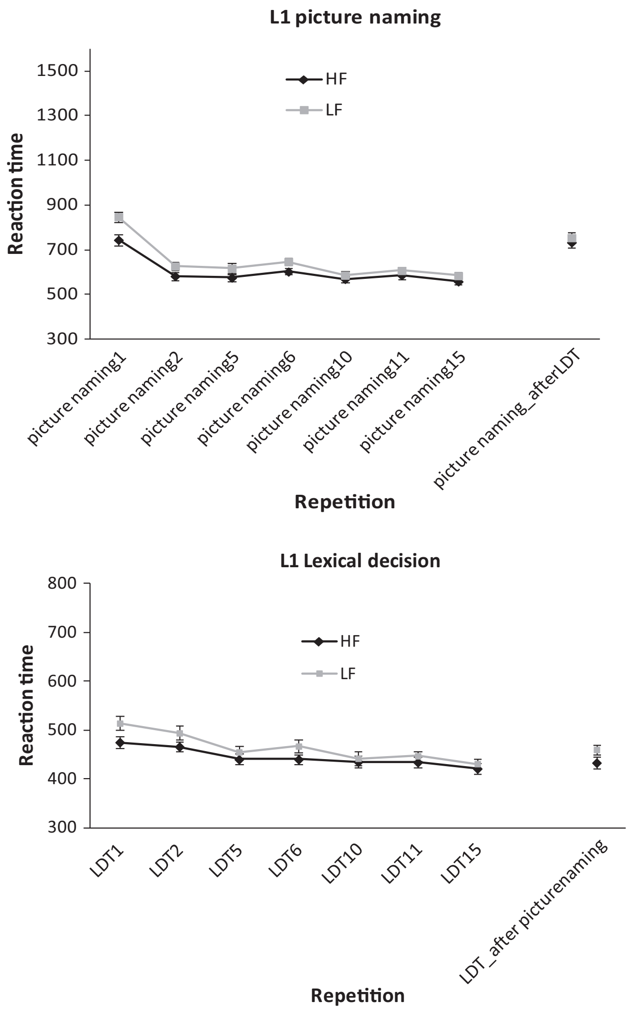

Fig. 4. Picture-naming (top panel) and lexical-decision times (lower panel) in milliseconds in L1 for high- and low-frequency words on the first, second, and fifth presentations (on day 1), sixth, tenth (on day 2), eleventh and fifteenth presentations (on day 3) for the cross-modal transfer list. Error bars are standard errors for each condition. LDT = lexical decision task, HF = high-frequency, LF = low-frequency.

Familiarization and training phases

As indicated in Fig. 4 and Table 5, and similar to Experiment 1, familiarization led participants to respond more quickly and reduced error rates. In both lexical decision and picture naming tasks, participants were faster in the second block than in the first block and more so for low-frequency than for high-frequency items. In the training phase, there was a further small decrease in RTs and frequency effects when items were repeatedly presented.

Table 5
Mean error percentages for picture naming and lexical decision in L1 for high- and low-frequency words on the first (on day 1), second to fifteenth presentations (on days 1–3), and for the cross-modal transfer list. LDT = lexical decision task, HF = high-frequency, LF = low-frequency.

Task		HF	LF
Picture naming	Picture naming1	2.03 (0.03)	6.90 (0.06)
	Picture naming2–15	0.34 (0.01)	0.43 (0.02)
	Picture naming after LDT	2.28 (0.04)	3.80 (0.05)
LDT	LDT1	1.05 (0.03)	2.92 (0.04)
	LDT2–15	2.53 (0.05)	3.82 (0.06)
	LDT after picture naming	3.96 (0.06)	7.29 (0.07)

Note. Standard deviations are presented in parentheses.

Cross-modal transfer

Reaction times were slower in picture naming than in lexical decision [$F_1(1,23) = 241.97$, $MSE = 17{,}763$, $p < .001$; $F_2(1,78) = 390.67$, $MSE = 19{,}105$, $p < .001$; min $F'(1,54) = 149.42$, $p < .001$] and were faster for high- than for low-frequency items [$F_1(1,23) = 71.20$, $MSE = 1{,}643$, $p < .001$; $F_2(1,78) = 7.67$, $MSE = 25{,}743$, $p < .01$; min $F'(1,92) = 6.92$, $p < .001$]. Responses were faster after training in the other modality compared to the first presentation [$F_1(1,23) = 30.51$, $MSE = 3{,}763$, $p < .001$; $F_2(1,78) = 48.39$, $MSE = 4{,}233$, $p < .01$; min $F'(1,55) = 18.71$, $p < .001$]. These transfer effects were stronger for low- than for high-frequency items as shown by a significant Repetition × Frequency interaction [$F_1(1,23) = 13.28$, $MSE = 1{,}761$, $p < .01$;

$F_2(1,78) = 17.05$, $MSE = 4{,}233$, $p < .001$; min $F'(1,62) = 7.47$, $p < .01$]. In contrast to Experiment 1 where frequency effects were stronger in picture naming than in lexical decision, the effect of frequency did not interact with Task [$F_1(1,23) = 8.96$, $MSE = 1{,}399$, $p < .01$; $F_2(1,78) = 1.40$, $MSE = 19{,}105$, $p = .24$; min $F'(1,96) = 1.21$, $p = .27$] for L1 processing. Also in contrast to Experiment 1 where the transfer effects were stronger in picture naming than in lexical decision, there was no such interaction of Task × Repetition in Experiment 2 [$F_1 < 1$; $F_2(1,78) = 1.83$, $MSE = 3{,}574$, $p = .18$; min $F' < 1$]. The marginally significant interaction of Task × Repetition × Frequency indicated that the cross-modal transfer effects and the frequency effects were stronger in picture naming than in lexical decision [$F_1(1,23) = 6.08$, $MSE = 2{,}223$, $p < .05$; $F_2(1,78) = 8.07$, $MSE = 3{,}574$, $p < .01$; min $F'(1,61) = 3.47$, $p = .07$].

Picture naming

Participants produced pictures with high-frequency names more quickly than pictures with low-frequency names [$F_1(1,23) = 38.71$, $MSE = 2{,}662$, $p < .001$; $F_2(1,78) = 4.30$, $MSE = 42{,}982$, $p < .05$; min $F'(1,92) = 3.87$, $p = .05$]. Cross-modal training sped picture naming times [$F_1(1,23) = 12.30$, $MSE = 4{,}753$, $p < .01$; $F_2(1,78) = 19.61$, $MSE = 7{,}258$, $p < .001$; min $F'(1,55) = 7.56$, $p < .01$]. The frequency effect was significantly and substantially smaller in picture naming after training with lexical decision than it was in picture naming without prior cross-modal training, a significant Frequency × Repetition interaction [$F_1(1,23) = 10.61$, $MSE = 3{,}413$, $p < .01$; $F_2(1,78) = 13.25$, $MSE = 7{,}258$, $p < .001$; min $F'(1,63) = 5.89$, $p = .02$]. The training effect was significant only for low-frequency [$F_1(1,23) = 23.78$, $MSE = 3{,}925$, $p < .01$; $F_2(1,78) = 32.54$, $MSE = 7{,}258$, $p < .001$; min $F'(1,60) = 13.74$, $p < .001$] and not for high-frequency targets [$Fs < 1$]. Table 4 presents the means and size of the cross-modal reaction time and frequency effects.

The analyses of errors (incorrect answers and pass-answers) showed that error rates did not differ between high- and low-frequency words [$F_1(1,23) = 8.52$, $MSE = 0.003$, $p < .01$; $F_2(1,78) = 4.32$, $MSE = 0.01$, $p < .05$; min $F'(1,95) = 2.87$, $p = .09$]. Training in lexical decision did not decrease the frequency effect in error scores [$F_1(1,23) = 3.05$, $MSE = 0.002$, $p = .09$; $F_2(1,78) = 3.38$, $MSE = 0.004$, $p = .07$; min $F'(1,67) = 1.60$, $p = .21$], nor the amount of errors [$F_1(1,23) = 2.17$, $MSE = 0.002$, $p = .15$; $F_2(1,78) = 3.33$, $MSE = 0.004$, $p = .07$; min $F'(1,56) = 1.31$, $p = .26$]. Mean error percentages in picture naming are presented in Table 5.

Lexical decision

RTs were faster for high- than for low-frequency words [$F_1(1,23) = 69.64$, $MSE = 380$, $p < .001$; $F_2(1,78) = 20.90$, $MSE = 1{,}871$, $p < .001$; min $F'(1,101) = 16.08$, $p < .001$]. Cross-modal training sped RTs in the lexical decision task [$F_1(1,23) = 26.29$, $MSE = 2{,}144$, $p < .001$; $F_2(1,78) = 127.38$, $MSE = 534$, $p < .001$; min $F'(1,33) = 21.79$, $p < .001$]. Differing from cross-modal training effects on picture naming, but resembling cross-modal effects found in L2, training in picture naming did not reduce the size of the frequency effect in lexical decision, a non-significant Frequency × Repetition interaction [$F_1(1,23) = 1.18$, $MSE = 571$, $p = .29$; $F_2(1,78) = 11.23$, $MSE = 534$, $p < .01$; min $F'(1,28) = 1.07$, $p = .31$]. The decrease in RTs in the lexical decision task after training with picture naming was present for both high-frequency [$F_1(1,23) = 25.55$, $MSE = 875$, $p < .001$; $F_2(1,78) = 30.72$, $MSE = 534$, $p < .001$; min $F'(1,64) = 13.95$, $p < .001$] and low-frequency targets [$F_1(1,23) = 18.85$, $MSE = 1{,}839$, $p < .001$; $F_2(1,78) = 109.86$, $MSE = 534$, $p < .001$; min $F'(1,31) = 16.09$, $p < .001$]. Means and size of the cross-modal reaction time and frequency effects are presented in Table 4.

There were fewer errors with high-frequency than with low-frequency targets [$F_1(1,23) = 9.83$, $MSE = 0.002$, $p < .01$; $F_2(1,78) = 5.85$, $MSE = 0.002$, $p = .02$; min $F'(1,90) = 3.67$, $p = .06$]. Contrary to the picture naming results, cross-modal training slightly increased error rates in lexical decisions (increase of 3.5%) [$F_1(1,23) = 17.27$, $MSE = 0.002$, $p < .001$; $F_2(1,78) = 6.20$, $MSE = .002$, $p = .01$; min $F'(1,100) = 4.56$, $p = .04$], but given the low error rates and the fact that we did not find this unexpected increase in Experiment 1, we do not interpret this finding. Cross-modal training did not reduce the frequency effect in errors [$Fs < 1$]. Mean error percentages are presented in Table 5.

Proportionally adjusted frequency effects

Fig. 3 shows the results of an analysis where we calculated proportionally adjusted frequency effects for each participant in the first presentation during training and cross-modal test. Table 4 presents the means across conditions. Proportionally adjusted frequency effects were smaller in the cross-modal test compared to the first presentation (a main effect of Repetition) [$F(1,23) = 9.40$, $MSE = 0.01$, $p < .01$], but did not differ across tasks [$F(1,23) = 1.08$, $MSE = 0.01$, $p = .31$]. However, as found in Experiment 1, cross-modal training benefitted picture naming much more than lexical decision, a significant interaction of Task × Repetition [$F(1,23) = 6.36$, $MSE = 0.01$, $p = .02$]. Cross-modal training decreased the size of the proportionally adjusted frequency effect in production [$F(1,23) = 10.14$, $MSE = 0.01$, $p < .01$], but not in recognition [$F(1,23) = 0.44$, $MSE = 0.003$, $p = .51$].

In conclusion, as in Experiment 1, training yielded quicker responses and a reduction of the size of the frequency effect. Cross-modal training sped picture naming times (but only for low-frequency names, unlike Experiment 1) and also reduced the size of the frequency effect. Cross-modal training also sped responses for both high- and low-frequency words in lexical decision, but did not reduce the size of the frequency effect.

General discussion

Using the cross-modal training paradigm, the present study explored the extent to which visual word recognition and speech production share common representations. We trained bilinguals in reading high-and low-frequency words in a lexical decision task and subsequently tested them in a picture naming task, and vice versa. The same experiment was done first in the second language and then in the first language of bilinguals. Both picture naming and lexical decision tasks exhibited training effects: responses were faster when the items were presented repeatedly. Both tasks also exhibited strong frequency effects. The size of the frequency effect diminished with repeated presentation in both picture naming and lexical decision, in L1 and in L2.

To investigate cross-modal transfer, we compared responses for the first presentation of the stimuli in each task with the responses after training in the other modality. For example, the first presentation of the list 1 words in lexical decision on day 1 was compared to the lexical decision of the matched list 2 words on day 3 after training with picture naming. Cross-modal training sped responses in recognition and production. This facilitation was stronger for picture naming than for lexical decision in L2, but not in L1. Training with lexical decision reduced the size of the frequency effects in subsequent picture naming, in both L1 and in L2. However, cross-modal training appeared to reduce the size of the frequency effect much more in picture naming than in lexical decision: this reduction was highly robust in both L1 and L2 for picture naming, but in lexical decision it was not significant. This difference between tasks was confirmed by a subsequent analysis that adjusted for baseline differences between tasks in the length of overall response times using a calculation of the frequency effect as a proportion of overall RTs. This analysis revealed a significant interaction between task and repetition (i.e., cross-modal training effect) in the size of the proportionally adjusted frequency effect; training with repeated lexical decision led to significantly smaller frequency effects in picture naming in L2 and L1, whereas training with repeated picture naming did not change the size of the frequency effects in lexical decision in either L2 or L1.

Within-modality repetition

In both picture naming and lexical decision tasks, the familiarization phase had the expected effect of facilitating the reaction times and decreasing the frequency effect. The training paradigm with repetition of items over three days constituted an experimental induction of frequency, changing the activation levels of representations and leading to faster access and processing. This can be explained by the same mechanisms as word frequency effects (e.g., Monsell, 1991; Wheeldon & Monsell, 1992). Frequency effects are long-lasting, possibly reflecting changes in baseline resting levels of activation. In the present study, training effects were especially evident in production in which there was a sharp decrease in response times when comparing first to second presentations in both languages. This is in agreement with previous studies that also showed a strong decrease in naming times from the first to the second presentation and a smaller decrease for each additional repetition (e.g., Gollan et al., 2005; Griffin & Bock, 1998; Oldfield & Wingfield, 1965).

We cannot exclude that the training effects we observed partly reflect temporary boosts in activation due to recent exposure to the same item, versus changes in baseline activation levels. However, such temporary residual activation is thought to decay rapidly as the interval increases between the two repeated items (e.g., Coltheart et al., 2001; Levelt et al., 1999). Given that we had on average 40 intervening pictures in picture naming and 80 intervening words/nonwords in lexical decision and given that participants were trained on three different days, the effect of residual activation on training and frequency effects should be rather small, and observed training effects are likely to be long-lasting structural changes.

Linking recognition and production: cross-modal transfer

The finding of cross-modal facilitation in both training directions (for both L1 and L2) is consistent with the notion of shared lexical representations of which baseline activation levels are boosted by training. As shown in Fig. 1, cross-modal facilitation may originate from shared, or connected, conceptual, lemma and phonological representations. These representations may be shared between modalities but need not be activated to the same extent across tasks. For example, activation of an orthographic representation may be sufficient for lexical decision (e.g., Coltheart, Davelaar, Jonasson, & Besner, 1977), and therefore, conceptual, lemma, and phonological representations might not be activated as much in lexical decision as in picture naming. However, picture naming must necessarily involve conceptual, lemma, and phonological representations (e.g., Levelt et al., 1999). This might explain why transfer effects were stronger in picture naming than in lexical decision in Experiment 1. In Experiment 2, there was no difference in transfer effects between training directions. This is likely to be a consequence of the fact that representations of the more proficient language are closer to ceiling performance, with less additional room for facilitation effects.

Cross-modal training reduced the size of the frequency effect in production (in both L1 and L2) but it did not significantly change frequency effects in visual word recognition. Thus, although one might expect that training with a more difficult task should always result in greater transfer effects and stronger learning (e.g., Thompson, 2007), in this case the arguably more difficult task (i.e., picture naming) produced weaker (rather than stronger) effects. Importantly, our proportional analyses of the frequency effect demonstrated that differences in transfer effects across modalities were unlikely to be an artifact of baseline differences in overall response times (which were much longer in the modality that showed greater transfer effects, i.e., picture naming).

The modulation of the frequency effect in production after training in recognition indicates that production accessed the same frequency-sensitive representations as recognition training. Frequency-sensitive representations accessed in recognition and in production can be the lemma (e.g., Levelt et al., 1999), or the phonological representations (e.g., Coltheart et al., 2001; see Fig. 1), or both. Above, we argued that access to shared conceptual, lemma and phonological representations may account for cross-modal effects in terms of overall processing speed. This facilitation in processing speed was present in both training directions, but the decrease in frequency effects was only present in production after cross-modal training. This indicates that representations that are not sensitive to frequency should account for the cross-modal transfer effects in both training directions. This can be the conceptual level but possibly also the lemma level given that the main locus of frequency effects in production is the phonological level (e.g., Jescheniak & Levelt, 1994). Thus, it appears that the phonological representation that is activated or assembled during reading (e.g., Frost, 1998) is the same representation – or at least activates the same representation – that is needed in production.

There was no modulation of the frequency effects in recognition after production training. This can be explained by assuming that frequency-sensitive phonological representations in production were not needed in lexical decision which only requires activation of orthographic representations (e.g., Coltheart et al., 1977). Baseline activation levels of orthographic representations were not increased during production training, so that frequency effects in recognition after production training did not decrease.

The cross-modal transfer effects we report indicate that much can be gained by direct comparisons across modalities. Although historically, there has been a clear separation between researchers who study recognition and those who study production, there are theoretical accounts (e.g., Liberman & Mattingly, 1985; Pickering & Garrod, 2013) and previous studies (e.g., Monsell, 1987) pointing to interactions between recognition and production, which are consistent with the current findings. Most of these studies have focused on production and auditory recognition and used tasks such as cross-modal repetition priming and interference studies using auditory distractor words and picture naming (e.g., Monsell, 1987; Roelofs et al., 1996). However, no clear consensus has been reached as to what extent shared representations and processes are involved (e.g., Roelofs, 2003). The motor theory of speech perception for instance (Liberman & Mattingly, 1985) assumes that we perceive phonetic elements as the intended phonetic gestures of the speaker, represented in the brain as motor commands. Similarly, Pickering and Garrod (2013) propose a tight coordination between the production and comprehension systems of speaker and listener in dialog. This account assumes that listeners generate predictions at specific representational levels (e.g., sound, meaning) about speaker's utterances by relying on their own production system. Also, there is neuropsychological evidence in favor of shared systems for comprehension and production (e.g., Menenti, Gierhan, Segaert, & Hagoort, 2011; Neuhaus & Penke, 2008; Silbert, Honey, Simony, Poeppel, & Hasson, 2014).

Implications of language effects

We investigated cross-modal transfer effects first in L2 and then in L1 processing because we reasoned that L2 processing might provide better conditions for cross-modal transfer effects to occur. Word processing in L2 is generally slower (e.g., Duyck, 2005; Duyck et al., 2008; Gollan et al., 2008) and further from ceiling levels of performance compared to the L1 (e.g. Gollan et al., 2005; Hanulová et al., 2011; Runnqvist et al., 2011). In general, cross-modal transfer effects were stronger in L2 than in L1 processing. For instance, in L1 picture naming, only the low-frequency items benefitted from cross-modal training, whereas in L2, both high-and low-frequency words benefitted. These results are consistent with the idea that lexical representations in L1 are closer to ceiling levels of activation than representations in L2 (Duyck et al., 2008; Gollan et al., 2005; Gollan et al., 2008; Gollan et al., 2011).

Limits to cross-modal transfer effects

The present results reveal that recognition and production rely on shared representations in lexical access. However, cross-modal training effects were not as strong as training effects within the same modality as can be seen in Figs. 2 and 4. There is always a slight increase in reaction times when switching modalities. There are several possible explanations for this. First, within a single task training may originate from faster peripheral processes outside the lexical system. For instance, production training may improve speech motor programs (e.g., Levelt et al., 1999). Second, within-modality training involves no task switch, whereas in cross-modality effects, testing occurs after a change in task. Waszak, Hommel, and Allport (2003) suggested that part of task-shift costs do not arise from control mechanisms, but arise from stimulus-response bindings created in another task in which the same stimulus was presented. In their study, participants switched between picture naming and word reading. The targets were either repeated or not in both tasks. Shift-costs were larger for targets presented in both tasks, than for targets presented in only one task. This shift cost lasted even when up to 100 trials intervened between the first and second task for that item. Waszak et al. suggested that each item in the picture naming task had been linked to its specific picture-naming stimulus-response binding. If participants then encountered the same item in the word-reading task, this triggered retrieval of the previously created picture-naming stimulus-response binding, and thus delayed response responses because of conflict with the stimulus-response binding in the word-reading process. For the present design, this shift cost for repeated items may have counteracted cross-modal facilitation effects.

Conclusions

Taken together, the cross-modal training results we observed in both picture naming and lexical decision imply the existence of amodal lexical representations that are not sensitive the word frequency accessed during production and recognition. The stronger decrease in frequency effect in picture naming than in lexical decision after cross-modal training also has modality specific implications. Reading activates frequency sensitive representations (phonology) that must be accessed in production, but that do not facilitate lexical access in recognition. Thus, although phonology may be rapidly and automatically activated during visual word recognition, it does not play an important role in the access process itself – which is instead a process more specific to the orthographic input lexicon (e.g., Murray & Forster, 2004; Rapp, 2001; Shalom & Poeppel, 2008). In this respect, the present study revealed how investigation of the link between recognition and production can shed light both on the answer to the question itself, as well as to questions more specific to theories within each subfield.

Acknowledgment

This research was made possible by the Research Foundation – Flanders (FWO Belgium), of which Eva Van Assche is a postdoctoral research Fellow and by grants from the National Institute on Deafness and Other Communication Disorders (011492), the National Institute of Child Health and Human Development (050287; 077140; 079426; 051030), and the National Science Foundation (BCS1457519).

Appendix A

English words/picture names in Experiment 1 and Dutch words/picture names in Experiment 2.

List number	Experiment 1 English		Experiment 2 Dutch	
	LF	HF	LF	HF
1	Arrow	Bird	Pijl	Vogel
1	Ashtray	Bottle	Asbak	Fles
1	Ax	Car	Bijl	Auto
1	Butterfly	Chair	Vlinder	Stoel
1	Cherry	Cheese	Kers	Kaas
1	Corn	Church	Maïs	Kerk
1	Donkey	Duck	Ezel	Eend
1	Fox	Eye	Vos	Oog
1	Frog	Glasses	Kikker	Bril
1	Kite	Knife	Vlieger	Mes
1	Mailbox	Lion	Brievenbus	Leeuw
1	Moose	Pig	Eland	Varken
1	Pumpkin	Queen	Pompoen	Koningin
1	Scissors	Rabbit	Schaar	Konijn
1	Skunk	Rope	Stinkdier	Touw
1	Snail	Smoke	Slak	Rook
1	Spider	Snake	Spin	Slang
1	Tweezers	Towel	Pincet	Handdoek
1	Waiter	Window	Ober	Raam
1	Whale	Witch	Walvis	Heks
2	Bra	Bag	Beha	Zak
2	Candle	Boy	Kaars	Jongen
2	Comb	Cage	Kam	Kooi
2	Couch	Chain	Zetel	Ketting
2	Deer	Chicken	Hert	Kip
2	Eagle	Coat	Arend	Jas
2	Knight	Dog	Ridder	Hond
2	Lobster	Egg	Kreeft	Ei
2	Owl	Flower	Uil	Bloem
2	Parrot	Horse	Papegaai	Paard
2	Peacock	Key	Pauw	Sleutel
2	Pineapple	Mirror	Ananas	Spiegel
2	Saw	Money	Zaag	Geld
2	Scarf	Monkey	Sjaal	Aap
2	Squirrel	Mountain	Eekhoorn	Berg
2	Turtle	Nurse	Schildpad	Verpleegster
2	Umbrella	Shower	Paraplu	Douche
2	Wig	Spoon	Pruik	Lepel
2	Wizard	Stairs	Tovenaar	Trap
2	Zipper	Thumb	Rits	Duim

References

Almeida, J., Knobel, M., Finkbeiner, M., & Caramazza, A. (2007). The locus of the frequency effect in picture naming: When recognizing is not enough. *Psychonomic Bulletin & Review, 14*, 1177–1182.
Baayen, H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database (CD-ROM)*.Philadelphia: University of Pennsylvania, Linguistic Data Consortium.
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 and Language, 44*, 350–375.
Bock, J. K. (1987). Coordinating words and syntax in speech plans. In A. Ellis (Ed.), *Progress in the psychology of language* (pp. 337–390). London: Erlbaum.
Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*, 977–990.
Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14*, 177–208.
Colomé, A. (2001). Lexical activation in bilingual's speech production: Language-specific or language independent? *Journal of Memory and Language, 45*, 721–736.
Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). Hillsdale, NJ: Erlbaum.
Coltheart, M., Rastle, K., Perry, C., Ziegler, J., & Langdon, R. (2001). DRC: A dual route cascade model of visual word recognition and reading aloud. *Psychological Review, 108*, 204–256.
Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1283–1296.
Damian, M. F., & Bowers, J. S. (2009). Assessing the role of orthography in speech perception and production: Evidence from picture-word interference tasks. *European Journal of Cognitive Psychology, 21*, 581–598.
De Bot, K. (1992). A bilingual production model: Levelt's speaking model adapted. *Applied Linguistics, 13*, 1–24.
Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*, 283–321.
Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes, 5*, 313–349.
Dell, G. S., & Gordon, J. K. (2003). Neighbors in the lexicon friends or foes? In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 9–38). Berlin Germany: Mouten de Gruyter.
Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition, 42*, 287–314.
Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and non aphasic speakers. *Psychological Review, 104*, 801–838.
Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*, 175–197.
Duyck, W. (2005). Translation and associative priming with cross-lingual pseudohomophones: Evidence for nonselective phonological

activation in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1340–1359.
Duyck, W., Anseel, F., Szmalec, A., Mestdagh, P., Tavernier, A., & Hartsuiker, R. J. (2008). Improving accuracy in detecting acoustic onsets. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 1317–1326.
Duyck, W., Desmet, T., Verbeke, L., & Brysbaert, M. (2004). Wordgen: a tool for word selection and nonword generation in Dutch, German, English and French. *Behavior Research Methods, Instruments & Computers, 36*, 488–499.
Duyck, W., Vanderelst, D., Desmet, T., & Hartsuiker, R. J. (2008). The frequency-effect in second-language visual word recognition. *Psychonomic Bulletin & Review, 15*, 850–855.
Feldman, L. B., & Turvey, M. T. (1983). Word recognition in Serbo-Croatian is phonologically analytic. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 288–298.
Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12*, 627–635.
Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 680–698.
Francis, W. S., Augustini, B. K., & Sáenz, S. P. (2003). Repetition priming in picture naming and translation depends on shared processes and their difficulty: Evidence from Spanish–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1283–1297.
Frost, R. (1998). Towards a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123*, 71–99.
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.
Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58*, 787–814.
Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition, 33*, 1220–1234.
Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General, 140*, 186–209.
Grainger, J., & Ferrand, L. (1994). Phonology and orthography in visual word recognition: Effects of masked homophone primes. *Journal of Memory and Language, 33*, 218–233.
Grainger, J., & Ferrand, L. (1996). Masked orthographic and phonological priming in visual word recognition and naming: Crosstasks comparisons. *Journal of Memory and Language, 35*, 623–647.
Grainger, J., & Holcomb, P. J. (2009). Watching the word go by: On the time-course of component processes in visual word recognition. *Lang Linguist Compass, 3*, 128–156.
Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103*, 518–565.
Green, D. W. (1986). Control, activation and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language, 27*, 210–223.
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.
Hanulová, J., Davidson, D. J., & Indefrey, P. (2011). Where does the delay in L2 picture naming come from? Neurocognitive evidence on second language word production. *Language and Cognitive Processes, 26*, 902–934.
Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition, 1*, 213–229.
Hillis, A. (2001). The organization of the lexical system. In B. Rapp (Ed.), *The handbook of cognitive neuropsychology: What deficits reveal about the human mind* (pp. 185–210). Ann Arbor: MI, Psychology Press.
Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*, 101–144.
Jared, D., & Seidenberg, M. S. (1991). Does word identification proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General, 120*, 358–394.
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.
Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology, 25*, 463–492.
Knobel, M., Finkbeiner, M., & Caramazza, A. (2008). The many places of frequency: Evidence for a novel locus of the lexical frequency effect in word production. *Cognitive Neuropsychology, 25*, 256–286.
Lemhöfer, K., & Broersma, M. (2011). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods, 44*, 325–343.
Levelt, W. J. M. (1989). *Speaking: From intention to articulation*.Cambridge, MA: MIT Press.
Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavior and Brain Sciences, 22*, 1–38.
Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech production revised. *Cognition, 21*, 1–36.
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.
McCusker, L.M. (1977, April). Some determinants of word recognition: Frequency. In *Paper presented at the 24th annual convention of the southwestern psychological association*. Fort Worth, TX.
Menenti, L., Gierhan, S. M. E., Segaert, K., & Hagoort, P. (2011). Shared language: Overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. *Psychological Science, 22*, 1173–1182.
Monsell, S. (1987). On the relation between lexical input and output pathways for speech. In A. Allport, D. MacKay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production: Relationships between listening, speaking, reading, and writing* (pp. 273–311). London: 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.
Murray, W. S., & Forster, K. I. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review, 111*, 721–756.
Navarette, E., Basagni, B., Alario, F. X., & Costa, A. (2006). Does word frequency affect lexical selection in speech production? *The Quarterly Journal of Experimental Psychology, 59*, 1681–1690.
Neuhaus, E., & Penke, M. (2008). Production and comprehension of wh-questions in German Broca's aphasia. *Journal of Neurolinguistics, 21*, 150–176.
Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology, 17*, 273–281.
Perfetti, C. A., & Bell, L. C. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language, 30*, 473–485.
Pickering, M. J., & Garrod, S. (2013). An integrated theory or language production and comprehension. *Behavioral and Brain Sciences, 36*, 329–392.
Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*, 56–115.
Rapp, B. (2001). *The handbook of cognitive neuropsychology: What deficits reveal about the human mind*.Ann Arbor, MI: Psychology Press.
Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107*, 460–499.
Rastle, K., McCormick, S. F., Bayliss, L., & Davis, C. J. (2011). Orthography influences the perception and production of speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 1588–1594.
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.
Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*, 191–201.
Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42*, 107–142.
Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition, 64*, 249–284.
Roelofs, A., & Meyer, A. (1998). Metrical structure in planning the production of spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 922–939.
Roelofs, A., Meyer, A. S., & Levelt, W. J. M. (1996). Interaction between semantic and orthographic factors in conceptually driven naming:

Comment on Starreveld and La Heij (1995). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 246–251.
Roelofs, A. (2003). Modeling the relation between the production and recognition of spoken word forms. In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 115–158). Berlin: Mouton de Gruyter.
Runnqvist, E., Strijkers, K., Sadat, J., & Costa, A. (2011). On the temporal and functional origin of L2 disadvantages in speech production: A critical review. *Frontiers in Psychology, 2*.
Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 1–17.
Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523–568.
Severens, E., Van Lommel, S., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica, 119*, 159–187.
Shallice, T., McLeod, P., & Lewis, K. (1985). Isolating cognitive modules with the dual-task paradigm: Are speech perception and production separate processes? *The Quarterly Journal of Experimental Psychology, 37*, 507–532.
Shalom, D. B., & Poeppel, D. (2008). Functional anatomic models of language: Assembling the pieces. *Neuroscientist, 14*, 119–127.
Silbert, L. J., Honey, C. J., Simony, E., Poeppel, D., & Hasson, U. (2014). Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proceedings of the National Academy of Sciences, 111*, E4687–E4696.
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 & Verbal Behavior, 23*, 67–83.
Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and meaning. *Memory and Cognition, 15*, 181–198.
Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs. *Cognitive Psychology, 46*, 361–413.
Wheeldon, L. R., & Monsell, S. (1992). The locus of repetition priming of spoken word production. *Quarterly Journal of Experimental Psychology, 44*, 723–761.