Journal of Memory and Language **47,** 91–124 (2002)
doi:10.1006/jmla.2001.2840

Lexical Decision and Word Naming in Bilinguals: Language Effects and Task Effects

Annette M. B. de Groot, Susanne Borgwaldt, Mieke Bos, and Ellen van den Eijnden

University of Amsterdam, Amsterdam, The Netherlands

We studied effects of 18 variables on Dutch and English lexical decision, standard word naming, and delayed word naming performed by Dutch–English bilinguals. The stimuli were 440 Dutch nouns and their English translations. In lexical decision in both languages the frequency variables were the most important predictors and the semantic variables also played a role. In addition, English lexical decision was affected by length. In contrast, onset structure was the most important predictor in Dutch naming, and it was also relatively important in English naming. In Dutch and English standard naming, word length and the intensity of the onset were also important variables and the word's neighborhood also played a role. The meaning variables affected performance in English naming but not in Dutch naming. These results are discussed in terms of differences in (1) task requirements, (2) orthographic depth, and (3) our participants' language proficiency in Dutch and English.

Key Words: word recognition; lexical decision; word naming; bilingualism; orthographic depth; language proficiency.

For nearly a century research on reading, with an emphasis on visual word recognition, has figured prominently in the research agenda of cognitive psychology. All of the influential models of word recognition that have been developed during that era have primarily been informed by the outcomes of experiments with monolingual native speakers of English. Yet, English is but one of the world's numerous languages and—due to the increasing globalization of many aspects of modern life—the incidence of bilingualism and multilingualism in individual language users may already outnumber the incidence of monolingualism. The implication may be that the present models of word recognition cannot be generalized unqualifiedly beyond the English monolingual case. Indeed, studies that have looked at word recognition in different languages/orthographies (e.g., Katz & Feldman, 1983; Katz & Frost, 1992) have revealed subtle dependencies between orthography and word-recognition procedures, and because a bilingual is not two monolinguals in one person but a unique and specific speaker–hearer (Grosjean, 1989), it is possible that word recognition in bilinguals differs from that in monolinguals.

Quantitative as well as qualitative differences may exist between monolingual and bilingual word recognition. Support for quantitative differences comes from the increasing number of studies that suggest that bilingual word recognition is initially "nonselective," which means that a language input activates word representations in both of the bilingual's language subsystems, the contextually appropriate as well as the contextually inappropriate (De Groot, Delmaar, & Lupker, 2000; Dijkstra, Jaarsveld, & Ten Brinke, 1998; Grainger & Dijkstra, 1992; Tzelgov, Henik, Sneg, & Baruch, 1996). Because, obviously, the two lexical subsystems of a bilingual together store more word representations than a monolingual lexicon, initial nonselectivity implies that upon presentation of a word there will generally be more lexical competition in a bilingual than in a monolingual. Qualitative differences are assumed by a well-known model

We owe much gratitude to Marijke Engels, Sarah Slaghuis, and Jeroen Raaijmakers for statistical advice and to Walter van Heuven for extracting the orthographic neighborhoods of the stimulus words from the CELEX database. Part of the work reported in this article was presented at the 2nd International Symposium on Bilingualism, April 14–17, 1999 (Newcastle upon Tyne, UK) and at the XIth ESCOP meeting, September 1–4, 1999 (Gent, Belgium).

Address reprint requests to A. M. B. de Groot, Department of Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands. E-mail: adegroot@fmg.uva.nl.

0749-596X/02 $35.00

of lexical access in bilinguals, the Bilingual Interactive Activation (BIA) model (Dijkstra & Van Heuven, 1998). It posits a representation level (the level of "language nodes") that is unique to the bilingual memory system and that is involved in resolving the competition between lexical items that are activated during the word-recognition process. The conclusion to be drawn from such quantitative and qualitative differences between monolingual and bilingual word recognition is that a universal account of word recognition requires a detailed understanding of how word recognition comes about in both monolinguals and bilinguals/multilinguals. The present study sought to increase such understanding regarding the bilingual case. Specifically, the goal of this study was to further our knowledge of the word features that affect word recognition in one particular bilingual population's two languages.

Eighteen variables were included in this study, 7 of which were taken from an earlier study by one of us on the determinants of word translation between two languages (De Groot, Dannenburg, & Van Hell, 1994). All of the variables concern characteristics of the stimulus words. Because it was not feasible to manipulate all 18 variables in a factorial experiment, a correlational approach was adopted.

The participants were all native speakers of Dutch with Dutch (L1) as their strongest language and English (L2) as their strongest foreign language. All of them started their formal English school training at about age 12 or 13 and had received some preparatory training in English at age 11 or 12. In a number of other respects our participants shared the same English language background, and the fluency and functional use of English did not vary substantially among participants. Our use of the word "bilingual" does not imply that the people concerned have no knowledge of any other language than Dutch and English. In fact, most of them had been trained in a number of other modern languages as well, in most cases French and German. The term "bilingual" simply means to highlight that a combination of two of their languages, the two strongest languages, is the focus of the present investigation.

THE VARIABLES

Semantic Variables

We included three semantic variables: imageability, context availability, and definition accuracy. Imageability distinguishes words by how easily they arouse a mental image. This variable is typically confounded with concreteness, which distinguishes words by the extent to which they can be experienced by the senses. Words that arouse a mental image easily are therefore called "concrete," whereas words for which it is hard to arouse an image are called "abstract." Imageability is also usually confounded with the second variable included here, context availability (Schwanenflugel, Harnishfeger, & Stowe, 1988; Van Hell & De Groot, 1998). The context availability of words is a measure of how easy it is to think up a context for them. For concrete words this is relatively easy, whereas for abstract words it is hard. Definition accuracy discriminates words by how easily language users think they could come up with a definition for them. It too is confounded with imageability, concreteness, and context availability (e.g., De Groot et al., 1994). The critical factor underlying all of these semantic variables may be the information density in the memory representations (Van Hell, 1998).

Frequency Variables

Our fourth and fifth variables, familiarity and frequency, are measures of a word's frequency of use. Familiarity is a subjective measure based on participants' ratings of the number of times they have experienced the word. This variable is usually highly correlated with objective frequency, which reflects the number of times the word occurs in a language corpus.

Length Variables

Four further variables all concern word length. One of these is simply the number of letters in the word and a second expresses word length in terms of the number of phonemes. The third length variable is the number of spoken syllables of the word, whereas a fourth expresses length in terms of number of morphemes. The more letters, phonemes, syllables, or morphemes a word contains, the more structurally complex it may be. In this respect, the length variables may be confounded with structural complexity.

Neighborhood Variables

Two lexical-neighborhood variables were included. One concerned the size of the word's within-language orthographic neighborhood (e.g., the number of a Dutch word's Dutch neighbors); the second was the size of the word's cross-language orthographic neighborhood (e.g., the number of a Dutch word's English neighbors). A word's orthographic neighborhood (henceforth referred to as the word's neighborhood) was established by determining how many other words exist (here, in Dutch or in English) that only differ from the target word in one letter position.

Onset Variables

Two variables concerned the structure of the word's onset. One of these two was the number of consonants in the onset of the first (or only) syllable of the spoken word. The scores on this variable ranged from 0 (for words starting with a vowel) to 3 (for words starting with three consonants). The second onset variable was a measure of the degree to which the onset of the word's first syllable conforms to language-universal preference structures. This measure of preferred onset structure was composed of two components—onset complexity and the number of violations of language-universal sonority constraints in the onset. The onset-complexity component takes account of the fact that in all natural languages syllables of the structure CV(C) exist and that words/syllables starting with a vowel are also very common across the world's languages. Languages that allow syllables with more than one consonant in the onset are much less common (e.g., Jakobson & Halle, 1956; Fudge & Shockey, 2000). The onset-complexity component of the preferred-onset-structure variable distinguishes among words in terms of the number of consonants in the word's onset and is, therefore, related to the onset variable described above (although in this case words beginning with a vowel or with a single consonant received the same score). The importance of onset complexity in word naming—which is one of the tasks that we employed in this study—was demonstrated by Frederiksen and Kroll (1976) and by Kawamoto and Kello (1999), although, surprisingly, the former authors showed the expected result of slower processing of words with two-consonant onsets than of words with one-consonant onsets, whereas the latter authors, contrary to prediction, obtained the opposite (and counterintuitive) pattern of results, with shorter responses for the words with complex onsets. Kawamoto and Kello (1999) explained this result in terms of the constraint the second letter/phoneme of a complex onset imposes on the first letter/phoneme of the onset. This constraint, in turn, was attributed to the sonority principle (Selkirk, 1985).

For the second component of the preferred-onset-structure variable, we assigned words a score based on the number of violations of the sonority principle. The sonority principle assumes a hierarchy of speech sounds. From lowest to highest sonority the ranking is as follows: plosives, fricatives, nasals, liquids, glides, and vowels. This sonority hierarchy can be used to describe the preferred structure of a syllable (Selkirk, 1985): The nucleus of a syllable has the highest sonority, while the sonority of the surrounding consonants typically increases gradually from the onset of the syllable and typically decreases gradually toward the end of the syllable's coda. For example, in the word *breast,* /r/ is more sonorous than /b/, and /s/ is more sonorous than /t/. But violations of this preferred syllable structure occur (such as in the English word *strength,* where /t/ is less sonorous than is /s/). The second component of the preferred-onset-structure variable counts the number of such violations (see *Materials* for details).

To summarize, both components of the preferred-onset-structure variable take into account the incidence of particular onset structures across the world's natural languages: Simple onsets are more common than complex onsets (that consist of more than one consonant), and onsets that conform to the sonority principle are more common than onsets that violate it. Because the first component of this onset variable is similar to the other onset variable included in this study (see above), the two onset variables may be expected to correlate highly with one another.

Consonant-Cluster Variables

A plausible reason complex onsets do not occur in all languages is that consonant clusters may be particularly hard to pronounce (Smith, 1973; see also Crystal, 1987, p. 240). Therefore, responses in tasks that require the pronunciation of the stimulus materials may be slowed by the occurrence of consonant clusters. The present word-naming task is such a task. We have therefore included two further variables that focus on consonant clusters (but now not only in the onset). One is the number of two-consonant clusters in the stimulus; the second is the number of three-consonant clusters in the stimulus. Because any word with a consonant cluster in initial position (a variable that was taken into account in the above onset variables) has at least one consonant cluster, the onset and consonant-cluster variables may be expected to correlate.

Sound-Intensity Variables

Two variables were included to reflect the fact that speech sounds, and thus also the first phones of spoken words, differ in intensity. The first used the speech-sound-intensity classification scheme presented by Fry (1979, in Crystal, 1987, p. 134), which classifies English speech sounds in terms of their average intensity in decibels (see *Materials* for details). Scores were assigned to the stimulus words depending upon the intensity level of their initial phoneme. The second intensity variable departs from the assumption that stressed syllables have a higher intensity than unstressed syllables; it distinguishes between words with stress on the first (or, in the case of monosyllabic words, the only) syllable and words with stress on any other syllable. These two variables were included because the intensity of a spoken word's onset is known to affect the moment the voice key (used in the naming tasks) will detect the response (see Kawamoto & Kello, 1999; Kessler, Treiman, & Mullennix, in press; Sakuma, Fushimi, & Tatsumi, 1997). Kessler et al. (in press) provided a fine-grained analysis of voice key effects in English word naming and concluded that they are large, pervasive, and systematic and therefore cannot be ignored in naming studies.

Cognate Status

The final variable was cognate status. It is a measure of the perceptual similarity of the words in a translation pair (e.g., the Dutch–English translation pair *peer–pear* consists of cognates, whereas the pair *bloem–flower* consists of noncognates). This variable, as well as the cross-language-neighborhood variable, enjoys a special status in this study, because effects or null effects of these variables provide information on the nature of the bilingual lexical-access process. Specifically, such effects indicate whether this process is nonselective (as suggested above) or selective (meaning that only the contextually appropriate language system responds to the input).

Many of the present set of variables are not newcomers in the word-recognition research field and, indeed, some of them (familiarity, frequency, imageability, and neighborhood) have received much attention. What makes this study, we believe, unique and important is that the effect of each one of this rather large set of variables is studied in both of the tasks that are most commonly used to study word recognition (lexical decision and word naming) and in both of the present bilingual population's two languages, while using the same stimulus set in the various task by language conditions (except, of course, that the language of the materials differs between the two language conditions). Also noteworthy is that, instead of using a small stimulus set, as is often done, our stimulus set was large and varied. It may therefore be expected to reflect word diversity in natural language material reasonably faithfully, except that only nouns were used. These features of this study provide an excellent opportunity for determining the relative contribution of each of the (clusters of) variables within and across the present tasks and languages vis à vis one another.

THE TASKS: LEXICAL DECISION AND WORD NAMING

The two most popular tasks in studies of word recognition are lexical decision and word naming. However, it is by now well established that neither of these tasks provides a pure measure of word recognition, and only one of them, word naming, is a task that is performed quite naturally by language users. Lexical decision is not only a word-identification task but also (and maybe primarily) a discrimination task, in which words and nonwords have to be discriminated from one another. The discrimination component of the task is influenced by a number of variables, such as the frequency/familiarity and the meaningfulness of the stimulus (e.g., Balota & Chumbley, 1984). Because such stimulus characteristics can be exploited in discriminating words from nonwords, a "word" decision may not even require word identification in all cases. Other processes, such as a feeling of familiarity, may lead to a "yes" response as well (Grainger & Jacobs, 1996). Consequently, whenever an effect of a particular variable is obtained in lexical decision, it is not immediately obvious that it affected actual word recognition.

The word-naming task as a tool to study word recognition has a similar shortcoming. Performing this task requires not only recognizing the word but also pronouncing it. Either the recognition stage of processing or the production stage, or both of them, may be the locus of a particular effect. (The production stage refers to all processes after identification but before the onset of the vocalization; Balota & Chumbley, 1990, p. 232.) In a study that sought to determine the locus of the word-frequency effect in word naming, Balota and Chumbley (1985) tackled this problem by comparing performance in a standard word-naming task (where the participants read aloud the word as quickly as possible from the moment it appears on the screen) with performance in a delayed-naming task. In the latter task the word is not to be read aloud immediately upon its presentation, but only when a particular cue (in their study, a pair of brackets surrounding the word) is presented. Interestingly, a frequency effect still materialized when the participants should have had sufficient time to recognize the word when the cue appeared (see McRae, Jared, & Seidenberg, 1990, for a replication). This finding led the authors to suggest that the output stage of word naming is sensitive to frequency. Other researchers (Forster & Chambers, 1973; Monsell, Doyle, & Haggard, 1989; Savage, Bradley, & Forster, 1990) have rejected this view, but Balota and Chumbley (1990) argued that many of the arguments raised against the delayed-naming technique were flawed. The more general implication of the finding that frequency affects the output stage in word naming is that whenever a particular variable turns out to affect processing in the standard word-naming task, the locus of the effect is undetermined.

An additional complication of the naming task is that certain naming responses may come about nonlexically, that is, by bypassing the lexicon altogether. This holds in particular for words with regular grapheme-to-phoneme correspondences. These can, in theory, be named through a process in which the word's sound is assembled by recoding the word's individual letters or letter clusters into the corresponding sounds and blending the outcome of these recoding operations into the sound structure of the whole word. In these cases no actual word recognition needs to take place. Baluch and Besner (1991) and Tabossi and Laghi (1992) have shown that the inclusion of pseudowords in the stimulus set promotes the use of nonlexical naming, whereas a stimulus set consisting of exclusively words fosters the use of a lexical route to the naming response.

In this study we adopted a converging-evidence approach (advocated by, among others, Andrews, 1997) by including both the lexical-decision task and the naming task. Similar effects of a particular variable across the two tasks will suggest that the effect concerned has its locus in the assumed common word-recognition component of the tasks (the component that Grainger & Jacobs, 1996, refer to as the "functional overlap" between the tasks). Different effects across the two tasks will suggest that this variable taps into (one of) the tasks' unique components, those not shared between the tasks.

To foster the use of the lexical route to the naming response, the naming task included words only (Baluch & Besner, 1991; Tabossi & Laghi, 1992). Furthermore, to get as detailed a picture as possible of how the present variables affect the word-recognition component of the standard-naming task, we administered a delayed-naming task as well.

THE LANGUAGES: DUTCH AND ENGLISH

In our participants L1 (Dutch) acquisition may be expected to be largely completed by the time school training in L2 (English) starts at about age 12 or 13. The reading processes learned for Dutch, including visual word recognition, may largely transfer to reading in the new language, English. For this reason and because both Dutch and English have an alphabetic orthography, we may expect similar results for the present two language conditions.

But differences between the two languages may also be expected, because Dutch and English differ in orthographic depth. In orthographically shallow languages the correspondences between graphemes and phonemes are relatively consistent, whereas orthographically deep languages are characterized by many inconsistencies in the grapheme–phoneme relations. It has been argued that these differences in orthographic depth lead to processing differences between languages, particularly in naming (e.g., Katz & Frost, 1992). The more shallow a language, the larger the role of nonlexical, assembled phonology. As a consequence, the effects of all variables that have a lexical locus should be smaller in a shallow language than in a deeper language. In support of this "orthographic-depth hypothesis" Katz and Feldman (1983) obtained a semantic-priming effect when native speakers of English (a relatively deep language) named English words but not when native speakers of Serbo-Croatian (a shallow language) named Serbo-Croatian words. In contrast, in a lexical-decision task equally large semantic-priming effects were obtained in English and Serbo-Croatian, suggesting lexical processing in both language conditions, and to the same extent.

The orthographic-depth hypothesis was tested in cross-language work on monolingual native speakers of the languages concerned. It would be interesting to see whether bilinguals adapt their word-processing procedures to the depth of the language they are currently processing. Dutch is a more shallow language than is English. Therefore, all effects with a lexical locus should be smaller in Dutch naming than in English naming. In the present study, the semantic variables imageability, context availability, and definition accuracy, which clearly have a lexical source, may therefore be expected to show differential effects in the two language conditions. Effects of frequency/familiarity are also usually localized in the lexicon, at least partly (e.g., Baluch & Besner, 1991), and the same holds for neighborhood effects in at least some tasks (see Andrews, 1997, for a review). Whether neighborhood effects in word naming have a lexical source or, instead are due to the neighborhood characteristics affecting nonlexical naming processes is a topic of current debate (Andrews, 1997, p. 449; Peereman & Content, 1995). If they indeed originate in the lexicon, effects of frequency/familiarity and of neighborhood may therefore also be expected to be smaller in Dutch naming than in English naming.

Katz and Feldman's (1983) data suggested that, irrespective of language depth, lexical decision always involves lexical access, including semantic activation. For this task we might therefore expect to find equally large effects in Dutch and English of all variables that tap lexical processes. However, studies by De Groot (1989; Experiment 4), James (1975), and Kroll and Merves (1986; Experiment 2) suggested that in lexical decision word meaning is only reliably accessed and exploited when infrequent words are presented. The activated meaning appears to be used as an additional source of evidence for the decision that this particular uncommon visual form indeed represents a word. Presumably, semantic information is not exploited when a frequent word is presented, either because its common form itself arouses a strong enough feeling of familiarity to judge the stimulus to be a word or because this common form on its own activates the word's lexical representation beyond its recognition threshold (cf. Grainger & Jacobs, 1996).

If this analysis is correct, we may expect to find larger effects of the present semantic variables in English lexical decision than in Dutch lexical decision. The reason is that the materials we used in English were translations of the materials in Dutch. Because our participants' bilingualism was not completely balanced but, instead, their native language, Dutch, was stronger than their nonnative English, the subjective frequency of the materials may be expected to be generally higher in Dutch than in English. As a consequence, semantic information might be exploited less often in Dutch lexical decision than in English lexical decision.

To summarize, both in word naming and in lexical decision larger effects of semantic variables may be expected to occur in English than in Dutch. In word naming the different orthographic depths of Dutch and English would be held responsible for this language effect, whereas in lexical decision the present participants' unequal proficiency in Dutch and English would be regarded the source of the effect.

As to length in terms of the number of letters, this variable has reliably been shown to affect word naming (Henderson, 1982). This result has led researchers to suggest that word naming indeed involves a phonological assembly process, especially in shallow languages. A (sequentially operating) grapheme-to-phoneme recoding process may be expected to take longer the more letters the word contains. The effect of length in lexical decision appears less reliable (cf. Forster & Chambers, 1973; Frederiksen & Kroll, 1976; Whaley, 1978), a finding that suggests that this assembly process is exploited less often in lexical decision. If the earlier analysis of the relation between orthographic depth and route to the naming response is correct, we might expect larger length effects in Dutch naming than in English naming.

Finally, regarding the onset variables, number-of-consonant-clusters variables, and intensity variables, the prediction is that an effect of these variables will occur in naming but not in lexical decision. We assume differences in pronunciation difficulty to be the source of the effect of the onset variables and the number-of-consonant-clusters variables, and the source of the effect of the intensity variables to be the loudness of the pronunciation. Because only word naming involves a pronunciation response, only this task should respond to these variables.

To summarize, the primary goal of this study was to further our understanding of the word-recognition process by looking at the effects of a large set of word variables on word processing in the two most popular word-recognition tasks and in each of the two languages in Dutch–English bilinguals. Our knowledge about each of the two tasks' suitability as a tool to study the word-recognition process may also be expected to increase. The focus of the data analyses will be on both task effects and language effects.

METHOD

Participants

In all, 120 first-year psychology students from the University of Amsterdam took part, 40 of whom performed the lexical-decision task, 40 the standard-naming task, and 40 the delayed-naming task. Of the 40 participants in each task condition, 20 received the Dutch materials and 20 received the English materials. Within both the lexical-decision and the naming parts of the study the participants were randomly assigned to the different conditions (the two language conditions in lexical decision and the four language by type-of-task conditions in naming).

All participants had Dutch as their native and strongest language and English as their strongest foreign language. They had all received English training in secondary school for at least 6 years, 3 to 4 h a week, starting at age 12 or 13. The university training of the participants required them to read mainly English textbooks. All participants had been exposed to English outside school from childhood, e.g., while watching television and movies and while listening to music. This extensive experience with English has resulted in rather high levels of English fluency, especially in English comprehension. Van Hell (1998, p. 18), who performed a large-scale study of the present population's

English comprehension skills, concluded that it was similar to that of Johnson and Newport's (1989) adult native speakers of Korean or Chinese who had lived in the United States since they were 8–10 years old.

We performed a test to assess the participants' English proficiency: The participants were asked to rate on a scale from 1 to 7 their comprehension and production abilities in English compared to the corresponding abilities in Dutch (1, very low; 7, the same as in Dutch). The ratings of the participants in the four naming conditions were entered in a 2 (task, delayed naming versus standard naming) by 2 (language, Dutch versus English) by 2 (type of skill, comprehension versus production) analysis of variance. This analysis showed a significant effect of type of skill ($p < .0001$), with comprehension ability receiving a higher score (mean, 6.07) than production ability (mean, 5.39). The main effects of task and language were not significant, nor were any of the interactions ($p > .10$ in all cases). Because the raw comprehension and production ratings for the Dutch and English lexical-decision conditions were no longer available, an ANOVA including the ratings of the participants in the lexical-decision conditions could not be performed. However, the mean ratings and standard deviations for these conditions were comparable to those for the naming conditions. Overall, the ratings suggest that the participants in the six language by task conditions were equally proficient in English.

Materials

In the Dutch lexical-decision condition 840 stimuli were presented, of which 440 were words (all nouns) and 400 were pseudowords that complied with Dutch orthography and phonology, but that carried no meaning. Similarly, in the English lexical-decision condition 840 stimuli were used, 440 words and 400 pseudowords. The 440 words in English were the dominant translations of the words in Dutch. The pseudowords in this condition were orthographically and phonologically regular in English. The 440 words and 400 pseudowords of each language condition were ordered alphabetically, words and pseudowords separately. Subsequently, they were divided into two lists per language, by assigning the alphabetically ordered stimuli alternately to List 1 and List 2. Every participant performed the lexical-decision task to all 840 stimuli of one of the language conditions, Dutch or English, in two separate sessions, held on 2 separate days. In each session the materials of one of the two lists were presented. The order in which the lists were presented alternated among the participants. The reason for splitting up the materials into two lists and presenting the lists on 2 separate days was to prevent fatigue. There were two practice lists, one for each of the two language conditions. Each practice list contained 11 words and 9 pseudowords.

The stimulus materials of the standard- and delayed-naming tasks were the 440 Dutch words and their English translations presented in the lexical-decision task. An additional 48 words, 24 Dutch and 24 English, served as practice materials. Every participant in the naming tasks received either all Dutch words or all English words in a single session.

The word stimuli of the present lexical-decision and naming tasks were the same 440 Dutch words and their translations that we used earlier in a study on word translation (De Groot et al., 1994). From this earlier study, for each word a rating on a 7-point scale was available on imageability (IMA), context availability (CA), definition accuracy (DEF), and familiarity (FAM). The 7 end of the scale indicated that the word was easy to imagine, to contextualize, to define, and highly familiar, respectively. All ratings had been provided by samples of participants taken from the same bilingual population as tested here. From these earlier studies a rating was also available of the cognate status of each Dutch–English translation pair (COG). These ratings indicated the degree to which the words in the translation pair were perceptually similar (1, completely dissimilar; 7, very similar).

Length information for all Dutch and English words (number of letters, NUML; number of phonemes, NUMP; number of syllables, NUMS; and number of morphemes NUMM) was extracted from the Nijmegen (The Netherlands) CELEX corpus (Baayen, Piepenbrock, & Van Rijn, 1993). The CELEX database does not contain morphemic-length information for all words. Where this information was missing we coded the morphemic length ourselves. Dutch word lengths ranged from 3 to 12 letters, from 2 to 11 phonemes, from 1 to 4 syllables, and from 1 to 3 morphemes. English word lengths ranged from 3 to 12 letters, from 2 to 10 phonemes, from 1 to 5 syllables, and from 1 to 3 morphemes.

The CELEX database also provided the information regarding the (logarithm of the) words' frequency of use (LOGF) and on (the size of) their orthographic neighborhood in both the target language and the nontarget language. These variables will be referred to as NTL and NNTL, for neighborhood target language and neighborhood nontarget language, respectively. The Dutch LOGF scores ranged from 1.65 to 4.69; the English scores ranged from 1.72 to 5.31. The neighborhood sizes of the Dutch words ranged from 0 to 23 (NTL) and from 0 to 19 (NNTL); the neighborhood sizes of the English words ranged from 0 to 22 (NTL) and from 0 to 19 (NNTL).

As pointed out earlier, the first of the present two onset variables (ONS1) concerned the number of consonants in the words' spoken onset: Words beginning with a vowel received a 0; words starting with one, two, or three consonants received 1, 2, and 3, respectively. The second onset variable (ONS2) was a measure of the preferred structure of the onset as determined by both onset complexity and whether and to what extent the sonority principle was violated. The words' ratings on this variable ranged from 0 to 3. Words starting with an empty onset or a single consonant were assigned a 0 rating. All remaining onsets were scored according to two criteria: the number of consonant phonemes in the onset and the number of violations of the sonority principle. If the number of consonants in the onset was larger than 1, the initial score of 0 was increased by 1 for each additional consonant. An onset of two consonants thus received a 1, whereas an onset of three consonants received a 2. For each violation of the sonority principle the score was again increased by 1.

The variables NUMC2 and NUMC3 reflected the number of two-consonant and three-consonant phoneme clusters, respectively, in a word. The scores for both the Dutch and the English words ranged from 0 to 3 (NUMC2) and from 0 to 2 (NUMC3).

The final two variables concerned the sound intensity of the stimulus words' onset. We used the following procedure in assigning each English word a score on the first of these two variables (INT1): All phonemes occupying the first position in our English word set and assigned decibel-intensity values between 20 and 29 by Fry (1979; see Crystal, 1987) received a 0. All first-position phonemes assigned intensity values between 15 and 19 received a 1. All first-position phonemes with intensity values between 10 and 13 received a 2. Finally, all first-position phonemes with intensity values between 0 and 8 by Fry received a 3. The intensity value for /h/ is missing in Crystal (1987). We estimated it to have a low intensity and therefore assigned it a 3. Unlike Kessler et al. (in press), we performed no separate analyses for the second phonemes in the stimuli.

Lacking a similar classification scheme for Dutch, we transferred the English classification scheme to Dutch by categorizing the Dutch word-initial phonemes the way the analogous English phonemes were categorized. As the Dutch language contains some fricatives that are not included in the English phoneme inventory, we categorized them like analogous English fricatives and assigned them a 3.

As mentioned earlier, the second of our two sound-intensity variables (INT2) simply distinguished between words with stress on the first (or only) syllable and words with stress on any other syllable. The former type of words received a 1; the second received a 2.

Apparatus and Procedure

The experiment was run on an Apple Macintosh computer. In the lexical-decision condition a response box with two buttons was attached to the computer. The righthand button was to be pushed if the stimulus was a word; the lefthand button was for pseudowords. In the standard- and delayed-naming conditions a voice-operated switch which registered the participants' responses was attached to the computer. In all conditions the same Pascal program controlled the stimulus presentation and the recording of the response times. The stimuli were always presented in lowercase letters.

The experimenter was always present in the experimental room during the practice session, providing additional instructions if required. In the lexical-decision condition the experimenter withdrew from the room at the conclusion of the practice session. In the naming conditions the experimenter remained in the room throughout the experiment. Before the onset of stimulus presentation the participants received written instructions.

At the onset of every trial a fixation stimulus appeared on the screen for 1 s. Immediately following its offset the stimulus appeared in the center of the screen, slightly below the position of the earlier fixation stimulus. In the lexical-decision and standard-naming conditions the participants were to respond to the stimulus as quickly and accurately as possible. The moment the participant pushed a button in lexical decision, the word *correct*, or *wrong*, or *slow* appeared on the screen. This feedback remained on the screen for 2 s. One second after it had disappeared the next fixation stimulus appeared. The word *slow* appeared whenever the participant had taken more than 1200 ms to respond. In the standard-naming condition no feedback on performance was provided. In this condition the stimulus disappeared from the screen when the participant produced the onset of the naming response. One second later the next fixation stimulus appeared.

In the delayed-naming condition the participant was instructed not to respond immediately to the stimulus when it appeared on screen. In Dutch, 550 ms following the onset of the stimulus, this same stimulus appeared a second time, right below its first presentation, but this time enclosed by a pair of brackets. While presented for the second time, its first presentation remained in sight. In English the stimulus onset synchrony (SOA) between the stimulus' first and second presentations was 600 ms. The participant had to read the stimulus as quickly and accurately as possible when it appeared for the second time. Both these SOAs should amply suffice for recognition to have been fully completed the moment the cue appeared, because the average reading time for a word when read in a natural reading environment is about 250 ms (e.g., Balota & Chumbley, 1995). We added 300–350 ms to compensate for the fact that in the word-naming task, unlike in a natural reading environment, supportive context is not provided. The repeated-stimulus pair disappeared from the screen the moment the voice key was triggered, and 1 s later the next fixation stimulus appeared.

Though conceptually the same, this delayed-naming procedure differed from those used by other researchers (Balota & Chumbley, 1985; Forster & Chambers, 1973). Savage et al. (1990) contrasted performance in different display conditions and obtained null effects of display type. This finding suggests that these display manipulations have no differential effect on performance.

We used different SOAs in the two language conditions because our participants are somewhat more fluent in Dutch than in English and, therefore, recognize words faster in Dutch. With equally long SOAs the participants in the Dutch condition would therefore effectively have more time to process the word when it is presented for the first time and, consequently, have more time to prepare the output.

In all tasks the stimuli were presented in a random order, a different order for every participant. The stimuli were presented in blocks of 30 in the lexical-decision task and blocks of 44 in the standard- and delayed-naming tasks. In lexical decision, following each block the mean RT and number of errors for that block appeared on the screen. The lexical-decision task took about 50 min per session. The standard- and delayed-naming tasks took about 45 min.

RESULTS

For each stimulus word in each of the six language by task conditions (Dutch and English; lexical decision, standard naming, and delayed naming) we calculated a mean RT over all the responses to that word, collapsing across participants, but excluding the RTs associated with error responses. The percentage of errors was rather low in all conditions (5.6% at the most). Therefore, no analyses of the errors were performed. In the naming conditions, in addition to the error responses, trials where the voice key had responded too early or too late were also excluded in calculating the mean RT for the stimulus words. Early responses were those where the voice key was triggered by an environmental noise before the participant actually started naming the stimulus; the voice key responded too late whenever the onset of the actual naming response was spoken too softly for it to be detected by the device. Few of these voice key errors occurred, less than 2% in all naming conditions.

Table 1 presents the overall mean RTs for the lexical-decision and naming data, by task and language. Dutch shows the usual result that lexical decision takes more time than standard naming (F (1,439) = 229.04, $p < .0001$). In English a very small effect in the opposite direction occurs: RT in standard naming is 6 ms longer than that in lexical decision (F (1,439) = 5.18, $p < .05$), despite the fact that longer RTs for lexical decision than for standard naming is a common finding for English as well when native speakers of English serve as participants (e.g., Forster & Chambers, 1973; Frederiksen & Kroll, 1976). A plausible cause of this language by task interaction may be the additional requirement of the naming task that the response is produced out loud, a requirement that may be especially demanding when the response to be output involves a nonnative word (see below).

Table 1 also shows that delayed naming produced faster responses than standard naming (F (1,439) = 36474.9, $p < .0001$ in Dutch; F (1,439) = 21663.7, $p < .0001$ in English). This result is, of course, due to the fact that RT in delayed naming was measured from the onset of the second presentation of the stimulus word, when, according to the logic of the task, word recognition was already completed. Furthermore, responding in all three RT tasks was faster in Dutch than in English (t = 12.36, $p < .0001$ for lexical decision; t = 35.11, $p < .0001$ for standard naming; t = 22.61, $p < .0001$ for delayed naming). This finding may arise from our participants' higher proficiency in their native language Dutch than in English, but the difference between orthographies may also underlie the effect.

Also noteworthy is that the language effect was much larger in standard naming than in lexical decision. A plausible source of this difference is that only standard naming involves the oral production of the word and that preparing the word for production is likely to be harder in the weaker language than in the native language. This latter suggestion is supported by the language effect in the delayed-naming condition, which was hypothesized to isolate the production component of naming. A comparison of the size of the language effects in standard and delayed naming suggests that the language effect on the word-recognition component of the complete naming process was about 30 ms, a conclusion that is supported by the size of the effect in lexical decision. Conversely, this same comparison suggests that the language effect on the production component of standard naming was about 50 ms. As suggested above, this effect may be responsible for the fact that standard naming in English took as long as English lexical decision.

TABLE 1

Mean Response Times (in Milliseconds) in the Lexical-Decision and Naming Conditions with Standard Deviation in Parentheses

	Dutch	English	Difference
Lexical decision	525 (48)	560 (57)	35
Standard naming	486 (30)	566 (48)	80
Delayed naming	263 (34)	316 (43)	53

Correlation Analyses

For each of the six language by task conditions, a correlation analysis was performed on the complete data set, which contained 19 values for each of the 440 stimulus words. One of these 19 values per word was the overall RT of the correct responses to that word in that particular language by task condition. The remaining 18 values were the mean ratings for that word on the variables described earlier.[1]

[1]This data set is available on the web site of the *Journal of Memory and Language*.

Correlations between the Predictor Variables

The top part of Table 2 shows the pattern of correlations among the predictor variables for the Dutch words; the bottom part presents the corresponding data for the English words. As described earlier, the values for the predictor variables were different for the two language sets. The exception was the variable cognate status (COG), which is an interlanguage variable that expresses the form relationship between the Dutch and English words in translation pairs. We will restrict our discussion of the correlations in Table 2 to those that were high and moderate.

As Table 2 shows, the three semantic variables (IMA, CA, and DEF) were highly correlated in both the Dutch and the English words. In addition, in both sets the two frequency variables (LOGF and FAM) were moderately to highly correlated. Furthermore, in both sets the length variables, especially NUML, NUMP, and NUMS, were highly correlated. These variables correlated moderately with the fourth length variable, NUMM, and with the neighborhood variable NTL (and less so with the neighborhood variable NNTL). The correlations between the length variables and the neighborhood variables show that short words have more neighbors in both the target and the nontarget language than long words. This is a common finding (e.g., Andrews, 1997), which can be explained by the fact that the more letters (phonemes, syllables, and morphemes) a word has, the smaller the chance that it shares all but one of them with one or more other words in the lexicon. The moderate to high correlations between the two neighborhood variables in both language sets indicate that short words have relatively many neighbors in both of the present Dutch–English bilinguals' two languages. Furthermore, the two onset variables (ONS1 and ONS2) correlated moderately to highly in both language sets, presumably because both reflect the number of consonant phonemes in the words' onsets.

In the English set, but not in the Dutch set, one of the frequency variables (FAM) correlated moderately with two of the semantic variables (CA and DEF). This finding suggests either one of two possibilities: that our participants, when instructed to rate words in their weaker language on how easily they can be contextualized or defined, take into account the familiarity of these words in the target language or, vice versa, that they take into account word meaning when instructed to rate how often a particular English word is used.

Principal Component Analyses on the Predictor Variables

The correlation patterns in Table 2 suggest the existence of at least four underlying components among the predictor variables: a semantic component, a frequency component, a length component (which may also cover the neighborhood variables), and a component covering the two onset variables. Principal component analyses on the 18 predictor variables per language condition (with the number of factors extracted by default) confirmed this suggestion. The rotated factor matrices (Varimax Rotation Solution) of the Dutch and English sets are presented in the top and bottom parts, respectively, of Table 3.

In both language conditions six factors emerged, with the three semantic variables loading on one factor (Factor II in both languages), the two frequency variables loading on a second factor (Factor IV in Dutch and Factor V in English), the four length variables loading on a third factor (Factor I in both language sets), and the two onset variables loading on a fourth (Factor III in both language sets). In both language conditions the variable NUMC3 loaded on a fifth factor (Factor V in the Dutch set and Factor VI in the English set). The sixth factor differed between the two language conditions. In the Dutch set it can be identified as an intensity factor (Factor VI), loaded on by the two intensity variables INT1 and INT2. In contrast, the sixth factor that can be identified in the English set is a neighborhood factor (Factor IV), loaded on by the two neighborhood variables NTL and NNTL. In Dutch these two variables loaded on the length factor (Factor I). The predictor COG (cognate status) did not clearly load on any single one of the six factors. The six factors extracted in the

TABLE 2

Correlations between the Predictor Variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Dutch																	
1 IMA																	
2 CA	.82																
3 DEF	.83	.81															
4 COG	.16	.13	.11														
5 FAM	.06	.18	.03	.08													
6 LOGF	−0.05	.06	−0.01	.09	.64												
7 NUML	−0.33	−0.23	−0.28	−0.26	−0.07	−0.13											
8 NUMP	−0.31	−0.22	−0.26	−0.24	−0.07	−0.13	.94										
9 NUMS	−0.30	−0.20	−0.24	−0.18	−0.04	−0.08	.84	.86									
10 NUMM	−0.44	−0.34	−0.38	−0.34	−0.04	−0.10	.74	.71	.68								
11 NTL	.27	.19	.28	.15	.08	.10	−0.65	−0.64	−0.55	−0.41							
12 NNTL	.22	.17	.22	.17	.10	.10	−0.52	−0.47	−0.41	−0.31	.66						
13 ONS1	.11	.05	.02	.01	−0.03	−0.07	.02	.04	−0.23	−0.18	−0.06	−0.07					
14 ONS2	.03	.01	−0.05	−0.06	−0.04	−0.10	.10	.08	−0.17	−0.10	−0.16	−0.12	.80				
15 NUMC2	−0.13	−0.10	−0.11	−0.12	−0.02	−0.04	.36	.42	.13	.17	−0.37	−0.25	.32	.43			
16 NUMC3	−0.11	−0.11	−0.13	−0.14	−0.02	−0.10	.26	.31	.10	.26	−0.18	−0.14	−0.03	.09	−0.18		
17 INT1	.13	.06	.04	.05	−0.03	.01	−0.00	.05	−0.09	−0.04	.03	.07	.47	.10	.05	−0.05	
18 INT2	−0.22	−0.16	−0.21	.02	−0.02	.03	.39	.43	.51	.32	−0.30	−0.23	−0.09	−0.15	−0.05	.05	.16
English																	
1 IMA																	
2 CA	.80																
3 DEF	.74	.83															
4 COG	.19	.16	.15														
5 FAM	.20	.55	.58	.14													
6 LOGF	−0.01	.21	.29	.11	.71												
7 NUML	−0.42	−0.32	−0.24	−0.22	−0.09	−0.23											
8 NUMP	−0.40	−0.28	−0.22	−0.20	−0.09	−0.21	.89										
9 NUMS	−0.36	−0.24	−0.17	−0.22	−0.07	−0.21	.81	.87									
10 NUMM	−0.37	−0.26	−0.18	−0.23	−0.09	−0.20	.64	.58	.57								
11 NTL	.25	.15	.14	.15	.04	.21	−0.66	−0.58	−0.55	−0.31							
12 NNTL	.20	.13	.13	.10	.06	.19	−0.55	−0.43	−0.40	−0.23	.73						
13 ONS1	.10	.03	−0.01	.02	−0.02	−0.03	−0.02	.02	−0.22	−0.11	−0.02	−0.07					
14 ONS2	.02	−0.02	−0.06	−0.05	−0.04	−0.06	.06	.10	−0.11	−0.00	−0.14	−0.14	.80				
15 NUMC2	−0.24	−0.19	−0.18	−0.10	−0.07	−0.10	.46	.49	.30	.30	−0.39	−0.31	.27	.39			
16 NUMC3	−0.24	−0.23	−0.16	−0.14	−0.04	−0.04	.39	.38	.27	.26	−0.22	−0.16	−0.03	.12	−0.06		
17 INT1	.11	.01	.02	.05	−0.05	−0.05	−0.03	−0.02	−0.13	−0.07	.07	.06	.46	.20	.04	−0.08	
18 INT2	−0.37	−0.28	−0.19	−0.17	−0.11	−0.12	.51	.53	.59	.46	−0.27	−0.20	−0.22	−0.12	.17	.25	−0.09

Note. IMA, imageability; CA, context availability; DEF, definition accuracy; COG, cognate status; FAM, familiarity; LOGF, log word frequency; NUML, length in letters; NUMP, length in phonemes; NUMS, length in syllables; NUMM, length in morphemes; NTL, number of neighbors in the target language; NNTL, number of neighbors in the nontarget language; ONS1, onset complexity; ONS2, preferred onset structure; NUMC2, number of two-consonant clusters; NUMC3, number of three-consonant clusters; INT1, sound intensity 1; INT2, sound intensity 2; $p < .05$ if $r > .08$; $p < .01$ if $r > .11$.

TABLE 3

Rotated Factor Matrix for the Predictor Variables

	I	II	III	IV	V	VI
Dutch						
IMA	−0.23	.91	.05	−0.04	−0.05	.05
CA	−0.11	.93	.00	.11	−0.05	.01
DEF	−0.17	.92	−0.03	−0.02	−0.05	−0.03
COG	−0.24	.08	−0.09	.06	−0.47	.30
FAM	−0.03	.08	.00	.91	.03	−0.02
LOGF	−0.09	−0.04	−0.05	.89	−0.10	.03
NUML	.92	−0.12	.05	−0.03	.20	.06
NUMP	.91	−0.10	.06	−0.03	.21	.13
NUMS	.87	−0.10	−0.28	−0.01	.04	.11
NUMM	.69	−0.29	−0.17	.00	.33	.03
NTL	−0.77	.12	−0.18	.05	.02	.09
NNTL	−0.64	.10	−0.16	.09	.03	.16
ONS1	−0.06	.05	.89	−0.04	.02	.31
ONS2	.02	−0.09	.90	−0.03	.10	−0.00
NUMC2	.44	−0.07	.60	.02	−0.31	−0.26
NUMC3	.12	−0.05	−0.01	−0.02	.86	.06
INT1	−0.05	.07	.35	−0.02	−0.01	.78
INT2	.48	−0.15	−0.27	.03	−0.11	.58
Eigenvalues	5.32	2.43	2.05	1.64	1.24	1.13
Accounted variance (%)	29.6	13.5	11.4	9.1	6.9	6.3
English						
IMA	−0.32	.87	.06	.10	−0.14	−0.09
CA	−0.17	.91	−0.01	.02	.20	−0.11
DEF	−0.07	.89	−0.03	.06	.29	−0.06
COG	−0.22	.10	−0.03	.02	.14	−0.43
FAM	.02	.41	−0.03	−0.02	.85	−0.05
LOGF	−0.14	.01	−0.04	.14	.92	−0.04
NUML	.79	−0.15	.07	−0.45	−0.05	.17
NUMP	.84	−0.13	.12	−0.34	−0.03	.15
NUMS	.84	−0.07	−0.16	−0.31	−0.07	.10
NUMM	.75	−0.13	−0.00	−0.04	−0.06	.14
NTL	−0.37	.04	−0.04	.81	.06	−0.06
NNTL	−0.20	.04	−0.04	.85	.09	−0.04
ONS1	−0.15	.02	.93	−0.05	−0.01	.03
ONS2	−0.11	−0.05	.85	−0.22	.03	.17
NUMC2	.41	−0.21	.44	−0.38	.04	−0.36
NUMC3	.22	−0.11	.02	−0.11	.07	.85
INT1	.10	.09	.60	.33	−0.14	−0.15
INT2	.72	−0.16	−0.17	.03	−0.03	.12
Eigenvalues	5.58	2.41	2.32	1.41	1.12	0.99
Accounted variance (%)	31.0	13.4	12.9	7.8	6.2	5.5

Note. IMA, imageability; CA, context availability; DEF, definition accuracy; COG, cognate status; FAM, familiarity; LOGF, log word frequency; NUML, length in letters; NUMP, length in phonemes; NUMS, length in syllables; NUMM, length in morphemes; NTL, number of neighbors in the target language; NNTL, number of neighbors in the nontarget language; ONS1, onset complexity; ONS2, preferred onset structure; NUMC2, number of two-consonant clusters; NUMC3, number of three-consonant clusters; INT1, sound intensity 1; INT2, sound intensity 2.

Dutch set together accounted for 76.7% of the variance. The six factors extracted in the English set accounted for 76.4% of the variance.

Correlations between the Predictor Variables and the Dependent Variables

One goal of the next set of analyses was to find out whether and how the language of the stimuli influences the effects of the predictors on performance. To reach this goal we compared correlation coefficients within tasks, between languages. The second goal was to find out whether and to what extent the predictor variables differentially affect performance in word naming and lexical decision. We pursued this goal by comparing corresponding correlation coefficients between tasks, within each language.

Comparison between Languages, within Tasks

Lexical decision. Table 4 presents the language effects for lexical decision. The first column shows the correlations between Dutch lexical-decision RT and the 18 predictor variables; the second column provides the analogous correlations for English lexical decision. Correlations of $r > .08$ are statistically significant ($p < .05$ if $r > .08$; $p < .01$ if $r > .11$) and suggest an effect of the predictors concerned. An asterisk following a pair of adjacent correlation coefficients indicates that the two coefficients of the pair differ significantly at the 1% level (Fischer z analysis). Here, as in all pairwise comparisons to be reported, we have chosen the 1% level of significance to reduce the risk of Type I errors, a risk that is real because of the large number of such tests that we performed.

The direction of the significant correlations in Table 4 shows that in English responding was relatively fast for words that are easy to imagine, to contextualize, and to define; for words that are perceptually similar to their translation in the nontarget language, familiar, frequent,

TABLE 4

Correlations between Dutch and English Lexical-Decision RT and the Predictor Variables

	Dutch	English
IMA	−0.16	−0.21
CA	−0.27	−0.39
DEF	−0.14	−0.40*
COG	−0.03	−0.22*
FAM	−0.46	−0.70*
LOGF	−0.49	−0.57
NUML	−0.03	.18
NUMP	−0.02	.21*
NUMS	.00	.23*
NUMM	−0.02	.18
NTL	.02	−0.11
NNTL	−0.01	−0.09
ONS1	−0.06	−0.05
ONS2	−0.00	.01
NUMC2	−0.02	.07
NUMC3	.01	.07
INT1	−0.09	−0.11
INT2	−0.00	.22*
NTL-S	.04	.01
NNTL-S	−0.01	−0.01

Note. IMA, imageability; CA, context availability; DEF, definition accuracy; COG, cognate status; FAM, familiarity; LOGF, log word frequency; NUML, length in letters; NUMP, length in phonemes; NUMS, length in syllables; NUMM, length in morphemes; NTL, number of neighbors in the target language; NNTL, number of neighbors in the nontarget language; ONS1, onset complexity; ONS2, preferred onset structure; NUMC2, number of two-consonant clusters; NUMC3, number of three-consonant clusters; INT1, sound intensity 1; INT2, sound intensity 2; NTL-S and NNTL-S concern the neighborhood of short words only.

and short (in terms of number of letters, phonemes, syllables, and morphemes); and for words that have many neighbors in the target and the nontarget language. The direction of the two correlations assumed to reflect the word onset's sound intensity (INT1 and INT2) differed between the two intensity measures. The correlations between RT and the two onset predictors and the predictors reflecting the number of two-consonant and three-consonant clusters in words were not significant.

The significant correlation between English lexical-decision RT and the second intensity variable can be understood from the rather high correlation between this variable and the length variables (see Table 2): Words with stress on the first syllable of polysyllabic words as well as all monosyllabic words had received a 1 on this variable, whereas polysyllabic words with stress on a later syllable received a 2. But of course, monosyllabic words are generally shorter than polysyllabic words. In other words, this intensity variable is correlated with word length and the significant correlation between English lexical-decision RT and INT2 may be a length effect in disguise.

In Dutch fewer correlations between lexical-decision RT and the predictors were significant than in English. The length and neighborhood variables no longer showed a significant correlation with RT, nor did the cognate variable and the second of the two intensity variables. The direction of the significant effects was the same as that in English, but the coefficients tended to be smaller. It appears that the frequency factor (familiarity and log frequency) was the most important factor in lexical decision, in both language conditions, and that the semantic factor was the next most important factor.

In previous work, the effects of neighborhood have typically been studied for short words only. To compare our neighborhood results with those reported in other studies, we selected all three-, four-, and five-letter words from the present materials ($N = 228$) and calculated the correlations between the two neighborhood variables and lexical-decision RT in both languages once again, now for these short words only. The results of this analysis are reported in Table 4 (lines NTL-S and NNTL-S). As shown, the small but significant effects of the two neighborhood variables in English disappeared. This finding suggests that the effects observed for the complete English word set should be attributed to the longer English words.

A final analysis performed on the English lexical-decision data took account of the fact that in English the variable FAM correlated moderately with the semantic variables CA and DEF (see Table 2). We suggested that the participants who had rated English words on context availability and accuracy of definition may have taken into account the words' familiarity or, vice versa, that the participants who rated English words on familiarity may have taken meaning aspects of words into account. In other words, the effects of CA and DEF in condition English may (in part) be familiarity effects in disguise, or the familiarity effect may (in part) be semantic effects in disguise.

To obtain a purer assessment of the role of CA, DEF, and FAM in English lexical decision, we performed two further sets of analyses. In one of these we partialled the correlations between FAM and RT out of the correlations between CA and DEF on the one hand and RT on the other hand. In the second set we partialled the correlations between CA and RT and DEF and RT out of the correlation between FAM and RT. Because the correlation between FAM and RT in English lexical decision was very high ($r = -0.70$, see Table 4), these new analyses had a dramatic effect on the partial correlations between RT in this condition and CA and DEF. In fact, these partial correlations dropped to about 0 ($r = -0.01$ and $r = .01$), thus becoming smaller than the corresponding correlations in condition Dutch. The partial correlations between RT and FAM (when partialling out CA and DEF) did not show similar dramatic changes in the pattern of results. The correlation between English lexical-decision RT and FAM remained substantial in both cases ($r = -0.63$, both with CA and DEF partialled out) and statistically larger than the corresponding correlation for Dutch.

In conclusion, the role of CA and DEF in English lexical decision seems as yet to be undetermined, because it is possible that these "semantic" variables were partly familiarity variables in disguise. If true, the earlier effects of CA and DEF in this condition may in fact underscore the important role of the frequency/familiarity variables in this condition. However, the opposite may also hold, namely, that familiarity judgments on the stimuli were contaminated by meaning judgments. In this case, the conclusion that semantic variables play an important role in English lexical decision is still warranted. At the same time, the earlier conclusions regarding the familiarity variable do not need to be qualified, because the partial correlations had only minimal effects on the correlation between English lexical-decision RT and FAM.

To summarize, the data of Table 4 suggest that English lexical decision is affected by frequency, length, and cognate status and that the neighborhood variables affect lexical decision to the longer English words. The role of the semantic variables in English lexical decision remains undetermined due to the confound between FAM, CA, and DEF. In contrast, Dutch lexical decision only appears to be affected by the familiarity and semantic variables.

Naming. Table 5 presents the language effects for standard naming (columns 1 and 2), delayed naming (columns 3 and 4), and for standard naming with delayed naming partialled out (columns 5 and 6). The purpose of this last analysis was to remove the effect of the predictors on the production/output component of the naming response, hypothesized to be captured by the delayed-naming task, thus getting a purer view of the role the predictors exert on the word-recognition component of standard naming. Again, an asterisk to the right of a pair of correlation coefficients indicates that the two coefficients within the pair differ significantly from one another at the 1% level (Fischer *z* analysis).

As in lexical decision, in English standard naming responding was relatively fast for words that are easy to imagine, contextualize, and define; for words that are perceptually similar to their translation, familiar, frequent, and short; and for words that have many neighbors in the target and the nontarget language. Unlike in lexical decision, effects of the variables ONS1, ONS2, NUMC2, and NUMC3 were now obtained. Words with complex and structurally not preferred onsets were responded to relatively slowly. The more consonant clusters a word contained, the more slowly it was responded to. Finally, clear effects of the two intensity variables were obtained. Responding was relatively fast for word onsets with relatively high sound-intensity levels (but recall that the effect of INT2 may be a length effect in disguise).

The direction of the effects in Dutch standard naming was generally the same as that for the corresponding effects in English, but the size of the effects differed between the two languages in a number of cases: The semantic and frequency effects were larger in English (and in fact, some of these were not significant in Dutch). Also, the effect of length in syllables was larger in English. In contrast, the effects of the onset variables were larger in Dutch. The correlation coefficients between English standard-naming RT and CA and DEF decreased substantially (to $r = -0.14$ in both cases) when FAM was partialled out of them, but they remained significant. Furthermore, the correlation between FAM and English standard-naming RT became smaller when CA and DEF were partialled out of it, but it remained significant in both cases ($r = -0.25$, and $r = -0.24$, respectively).

In standard naming the coefficients for the two neighborhood variables did not differ significantly between the two language conditions, but partialling out delayed naming decreased the coefficients for Dutch. As a result, a language difference now materialized for the NTL variable, suggesting that the recognition of English words is affected more by the target-language neighborhood than the recognition of Dutch words.

However, subsequent analyses on only the three-, four-, and five-letter words in our stimulus set did not show a similar difference between the neighborhood effects for the two languages after delayed naming was partialled out: All four correlation coefficients involving the neighborhood variables varied between -0.21 and -0.23. As with the analysis of the complete word sets, the coefficients for the two neighbor-

TABLE 5

Correlations between Dutch and English Naming RT and the Predictor Variables

	Standard naming		Delayed naming		Naming	
	Dutch	English	Dutch	English	Dutch	English
IMA	−0.02	−0.23*	−0.02	−0.13	−0.01	−0.20*
CA	−0.04	−0.31*	−0.01	−0.24*	−0.05	−0.21
DEF	−0.09	−0.32*	−0.08	−0.27*	−0.05	−0.19
COG	−0.12	−0.19	−0.13	−0.06	−0.04	−0.21
FAM	−0.11	−0.37*	−0.03	−0.34*	−0.13	−0.20
LOGF	−0.15	−0.36*	−0.05	−0.26*	−0.16	−0.26
NUML	.25	.37	.25	.16	.11	.37*
NUMP	.24	.35	.25	.15	.09	.34*
NUMS	.08	.29*	.12	.05	−0.01	.36*
NUMM	.08	.19	.13	.02	−0.02	.25*
NTL	−0.25	−0.32	−0.21	−0.12	−0.15	−0.33*
NNTL	−0.19	−0.26	−0.12	−0.09	−0.15	−0.28
ONS1	.50	.23*	.44	.35	.30	−0.02*
ONS2	.60	.31*	.52	.37	.38	.08*
NUMC2	.25	.20	.22	.19	.14	.10
NUMC3	.15	.19	.12	.08	.09	.19
INT1	.31	.20	.35	.31	.09	−0.02
INT2	.16	.26	.17	.02	.06	.34*
NTL-S	−0.27	−0.21	−0.18	−0.07	−0.21	−0.23
NNTL-S	−0.18	−0.19	−0.05	−0.05	−0.21	−0.22

Note. Naming, standard naming with delayed naming partialled out. IMA, imageability; CA, context availability; DEF, definition accuracy; COG, cognate status; FAM, familiarity; LOGF, log word frequency; NUML, length in letters; NUMP, length in phonemes; NUMS, length in syllables; NUMM, length in morphemes; NTL, number of neighbors in the target language; NNTL, number of neighbors in the nontarget language; ONS1, onset complexity; ONS2, preferred onset structure; NUMC2, number of two-consonant clusters; NUMC3, number of three-consonant clusters; INT1, sound intensity 1; INT2, sound intensity 2. NTL-S and NNTL-S concern the neighborhood of short words only; $p < .05$ if $r > .08$; $p < .01$ if $r > .11$.

hood variables in standard naming did not differ significantly between the two language conditions. The only noteworthy language difference occurred in delayed naming of only the short words. Here a statistically significant correlation was obtained for the correlation between Dutch delayed naming and NTL-S ($r = -0.18$), whereas the remaining three correlations involving the neighborhood variables were statistical null effects. This outcome contrasts with the data for the complete Dutch and English word sets, where all four correlations between Dutch and English delayed-naming RT and the two neighborhood variables were significant. When combined, these data suggest an influence of the within-language neighborhood on the output component of naming Dutch short words; an effect of both the within- and between-language neighborhoods on the output component of naming the longer Dutch and English words; and an effect of both the within- and between-language neighborhoods on the recognition component of Dutch and English word naming.

As in the case of the neighborhood variables, in standard naming the coefficients for the variable cognate status did not differ significantly between the two languages, but partialling out delayed naming decreased the coefficient for cognate status in Dutch such that it was no longer significant. The fact that the corresponding coefficient remained significant in English suggests that cognate status affects the two languages differently. Specifically, it indicates that cognate status affects the recognition component of naming English but not Dutch.

Partialling out delayed naming from the length variables tended to reduce the size of these effects in Dutch but not in English. This finding suggests that a relatively large part of the length effects in Dutch is due to the production component or, vice versa, that a relatively large part of the length effects in English is localized in the recognition component of the task. The substantial decrease of the coefficients for the two onset variables and for one of the two intensity variables (INT1) indicates that the production component is an important locus of these effects. In fact, the finding that, after partialling out delayed naming, in English two of these effects were no longer significant suggests that the production component is their only locus in English. The correlations between the two onset variables and Dutch naming remained high after partialling out delayed naming. This indicates that the onset variables affect the actual recognition component of Dutch word naming.

Finally, partialling out delayed naming reduced the coefficients for the two frequency variables in English, with the effect that the pairs of coefficients no longer differed significantly between languages. This finding suggests that the frequency variables affect the word recognition component of word naming to the same extent in Dutch and English and that, in addition, they also play a role in the production component of naming English words.

Comparison between Tasks, within Languages

The next set of analyses focused on the effect of task on performance within each of the two languages. For instance, the correlations for Dutch lexical decision were compared with the correlations for Dutch standard naming and with those for Dutch standard naming with delayed naming partialled out. The analogous set of comparisons was performed for English. The relevant sets of correlation coefficients are repeated, in a reorganized way, in Tables 6 (the Dutch data) and 7 (the English data). An asterisk following a pair of adjacent correlation coefficients again indicates that the two coefficients within the pair differ significantly from one another at the 1% level (Fischer z analysis). Some columns are repeated within these tables (e.g., columns 1 and 3 in Table 6) so as to provide information on the significance of all com-

TABLE 6

Correlation Coefficients between the Criterion Variables and the Predictor Variables (Condition Dutch)

	Lexical decision	Standard naming	Lexical decision	Naming
IMA	−0.16	−0.02	−0.16	−0.01
CA	−0.27	−0.04*	−0.27	−0.05*
DEF	−0.14	−0.09	−0.14	−0.05
COG	−0.03	−0.12	−0.03	−0.04
FAM	−0.46	−0.11*	−0.46	−0.13*
LOGF	−0.49	−0.15*	−0.49	−0.16*
NUML	−0.03	.25*	−0.03	.11
NUMP	−0.02	.24*	−0.02	.09
NUMS	.00	.08	.00	−0.01
NUMM	−0.02	.08	−0.02	−0.02
NTL	.02	−0.25*	.02	−0.15
NNTL	−0.01	−0.19*	−0.01	−0.15
ONS1	−0.06	.50*	−0.06	.30*
ONS2	−0.00	.60*	−0.00	.38*
NUMC2	−0.02	.25*	−0.02	.14
NUMC3	.01	.15	.01	.09
INT1	−0.09	.31*	−0.09	.09
INT2	−0.00	.16	−0.00	.06

Note. Naming, standard naming with delayed naming partialled out; IMA, imageability; CA, context availability; DEF, definition accuracy; COG, cognate status; FAM, familiarity; LOGF, log word frequency; NUML, length in letters; NUMP, length in phonemes; NUMS, length in syllables; NUMM, length in morphemes; NTL, number of neighbors in the target language; NNTL, number of neighbors in the nontarget language; ONS1, onset complexity; ONS2, preferred onset structure; NUMC2, number of two-consonant clusters; NUMC3, number of three-consonant clusters; INT1, sound intensity 1; INT2, sound intensity 2; $p < .05$ if $r > .08$; $p < .01$ if $r > .11$.

TABLE 7

Correlation Coefficients between the Criterion Variables and the Predictor Variables (Condition English)

	Lexical decision	Standard naming	Lexical decision	Naming
IMA	−0.21	−0.23	−0.21	−0.20
CA	−0.39	−0.31	−0.39	−0.21*
DEF	−0.40	−0.32	−0.40	−0.19*
COG	−0.22	−0.19	−0.22	−0.21
FAM	−0.70	−0.37*	−0.70	−0.20*
LOGF	−0.57	−0.36*	−0.57	−0.26*
NUML	.18	.37*	.18	.37*
NUMP	.21	.35	.21	.34
NUMS	.23	.29	.23	.36
NUMM	.18	.19	.18	.25
NTL	−0.11	−0.32*	−0.11	−0.33*
NNTL	−0.09	−0.26*	−0.09	−0.28*
ONS1	−0.05	.23*	−0.05	−0.02
ONS2	.01	.31*	.01	.08
NUMC2	.07	.20	.07	.10
NUMC3	.07	.19	.07	.19
INT1	−0.11	.20	−0.11	−0.02
INT2	.22	.26	.22	.34

Note. Naming, standard naming with delayed naming partialled out; IMA, imageability; CA, context availability; DEF, definition accuracy; COG, cognate status; FAM, familiarity; LOGF, log word frequency; NUML, length in letters; NUMP, length in phonemes; NUMS, length in syllables; NUMM, length in morphemes; NTL, number of neighbors in the target language; NNTL, number of neighbors in the nontarget language; ONS1, onset complexity; ONS2, preferred onset structure; NUMC2, number of two-consonant clusters; NUMC3, number of three-consonant clusters; INT1, sound intensity 1; INT2, sound intensity 2; $p < .05$ if $r > .08$; $p < .01$ if $r > .11$.

parisons (indicated by the asterisks) within a single table.

Dutch. A comparison of the first two columns of coefficients in Table 6 (lexical decision and standard naming) shows the following results: The semantic variable CA played a larger role in lexical decision than in standard naming. The task difference for IMA was not significant at the 1% level, but the fact that the effect was significant in lexical decision but not in standard naming suggests that IMA also has a larger effect in lexical decision than in naming. The two familiarity variables affected both lexical decision and standard naming, but their role in lexical decision was substantially larger than in standard naming. In contrast, two of the length variables, the neighborhood, onset, NUMC2, and INT1 variables, as well as the variable cognate status affected performance in standard naming more than in lexical decision. In fact, the effects of these variables were generally not significant in lexical decision.

With the correlations for delayed naming partialled out the correlations between the neighborhood, length, cognate status, onset, number-of-consonant-clusters, and the intensity variables and naming RT tended to become smaller, suggesting that a large part of the effects of these variables in standard naming was due to production. However, most of these correlations remained significant, which suggests that they also play a role in the recognition component of the standard-naming task. The coefficient for cognate status was no longer statistically significant. As pointed out earlier, the latter finding suggests that cognate status affects the output component of standard Dutch naming. The drop in the size of the correlation coefficients after delayed naming was partialled out seemed especially large in the case of the onset and intensity variables, and indeed, the correlations for the two intensity variables dropped to near 0. The latter finding suggests that word production is the sole locus of the effect of the intensity variables (as it, theoretically, should be). In contrast, the coefficients for the onset variables remained relatively large. These results suggest that the onset variables have two loci, both recognition and production. On the whole, the results strongly suggest that performance in lexical decision and naming responds very differently to the present word-type manipulations and, in fact, the two tasks have rather little in common.

English. As in the Dutch set, the two familiarity variables affected both lexical decision and standard naming, but the role of these variables was larger in lexical decision. In contrast, the length variable number of letters and the neighborhood and the onset variables affected performance in standard naming more than in lexical decision. In fact, the effects of the onset variables were not significant in lexical decision. Unlike in Dutch, the semantic variables appeared to play a role not only in lexical decision, but also in standard naming. The fact that the size of the coefficients did not differ significantly between the two tasks suggests an equally large influence of the semantic variables in both tasks. The statistically equal size of the coefficients for cognate status indicates that these variables influenced RT in the two tasks to the same extent. Finally, the fact that the coefficients for NUMC2 and NUMC3 did not differ significantly between the two tasks suggests that the number of consonant clusters does not affect processing in the two tasks differently. Yet, the finding that the coefficients do not approach significance in lexical decision but *are* significant in standard naming (a pattern that also occurred in Dutch) indicates that they do play a role in standard naming but not in lexical decision.

When the coefficients for the delayed-naming task had been partialled out of those for the standard-naming task, the coefficients for the semantic variables tended to become smaller. The resulting coefficients for context availability and definition accuracy were now significantly smaller than those for lexical decision, suggesting, as in Dutch, a larger role of semantic variables in lexical decision than in the word-recognition component of naming. The coefficients for the two frequency variables remained significant but became smaller, suggesting that not only the recognition component but also the output component of the standard-naming task is a locus of the frequency effects in the naming of nonnative English. Partialling out delayed-naming RT dramatically decreased the size of the coefficients for the two onset variables, causing the task effect to disappear. This finding suggests that, unlike in Dutch, the production component of naming is the only locus of these effects in English naming.

Summarizing, lexical decision and standard naming again produced very different patterns of results, although the differences between the two tasks were less extreme than those in Dutch. But wherever significant differences between tasks occurred, they were analogous to those observed in Dutch.

Correlations between Tasks

The preceding analyses strongly suggest that performance in lexical decision and standard naming responds very differently to the present word-type manipulations. This conclusion receives additional support from the fact that the between-task correlations were rather low. For Dutch the correlation between lexical decision and standard naming was $r = .10$. The analogous correlation for English was $r = .45$. suggesting more similar processing across the two tasks in English than in Dutch. The highest between-task correlations occurred between the standard- and delayed-naming tasks for the separate languages. For Dutch this correlation was $r = .72$; for English it was $r = .70$.

Multiple Regression Analyses

To find out how much of the variance in each of the six language by task conditions was accounted for by the predictor variables, one multiple regression analysis was performed for each of the language by task conditions. Each of these analyses included all 18 predictor variables and standard-naming RT, delayed-naming RT, or lexical-decision RT in either Dutch or English as the criterion variable.

The last column of Table 8 (labeled “all”) presents the percentages of accounted variance (R^2) in these six analyses. It shows that the amount of accounted variance depends on both

TABLE 8

Percentages of Variance Accounted for by the Separate Clusters of Variables and the Clusters Combined, by Language and by Task

	Length[a]	Semantic[b]	Onset[c]	Frequency[d]	Nc3[e]	Intensity[f]	Neighborhood[g]	Cognate status	All
Dutch									
Standard naming	14	2	36	2	2	11	6	2	50
Delayed naming	11	2	27	0	2	13	4	2	43
Lexical decision	0	9	1	27	0	1	0	0	36
English									
Standard naming	15	11	10	16	4	12	10	3	40
Delayed naming	7	9	14	12	1	10	1	0	33
Lexical decision	6	21	1	50	0	6	1	4	57

[a]Length in letters, length in phonemes, length in syllables, length in morphemes.
[b]Imageability, context availability, definition accuracy.
[c]ONS1, ONS2.
[d]Familiarity, frequency.
[e]Number of three-consonant clusters.
[f]INT1, INT2.
[g]NTL, NNTL.

task and language. In Dutch more variance was accounted for in the naming tasks than in lexical decision, whereas in English the opposite pattern of results was obtained.

Effects of the separate clusters of variables. A further set of multiple regression analyses was run to provide additional support for the different roles of the various clusters of variables in the different language by task conditions. In this series of analyses, RT in each of the six language by task conditions served as the criterion variable in a cluster of eight multiple regression analyses: In one of them only the four length variables (in other words, the four common Factor I variables in the two sets of language materials; see Table 3) were entered into the regression analyses. In the second analysis only the variables loading on the semantic factor (imageability, definition accuracy, and context availability; Factor II in both language sets) were entered. In the third analysis the two onset variables were entered (that is, the variables that loaded on Factor III in both language sets). In the fourth analysis the two variables that loaded on the frequency factor, familiarity and log frequency, were entered (Factor IV in the Dutch set and Factor V in the English). In the fifth analysis only NUMC3 was entered (the variable that loaded on Factor V in the Dutch set and on Factor VI in the English set). In the sixth analysis the two sound-intensity variables, which in the analysis of the Dutch materials (but not in the analysis of the English materials) had loaded on one factor (Factor VI), were entered. In the seventh analysis the two neighborhood variables were entered, which in the analysis of the English materials (but not the Dutch materials) had loaded on one factor (Factor IV). Finally, in the eighth analysis only cognate status was entered, because some of the earlier analyses suggested a differential role of this variable in the two language conditions. The results of these analyses are summarized in the columns labeled Length through Cognate status in Table 8. Each of these columns shows the percentage of accounted variance (R^2) for one of the above clusters of variables for each of the three tasks and separately for the two language conditions. (The summed percentages for a given language by task condition exceed the percentage accounted for by all 18 variables combined because the clusters of variables are not completely uncorrelated.)

Table 8 succinctly shows the striking dissociation between lexical decision and word naming, especially in Dutch: In Dutch lexical decision only the frequency variables and, to a lesser extent, the semantic variables appear to play important roles, but in Dutch naming especially onset is important and, to a lesser extent, length and intensity. The fact that the values for the onset, length, and intensity variables in Dutch delayed naming are also all substantial suggests that the production component is the most important locus of these effects.

In English lexical decision, as in Dutch lexical decision, the frequency variables appear to determine performance most, followed again by the semantic variables; unlike in Dutch lexical decision, length seems to be important as well. As pointed out earlier, the apparent effect of intensity is presumably a length effect in disguise, and the semantic effects may be a frequency effect in disguise. In English naming more clusters of variables determine performance than in Dutch naming. Of these, the length, meaning, onset, frequency, and intensity variables also affect delayed naming substantially, suggesting an important role of these variables in actual word production. A particularly interesting finding is that neighborhood hardly affects English delayed naming, which indicates that the neighborhood effect in standard naming is to be attributed to the actual word-recognition process.

These results suggest that the onset and intensity variables play important roles in all naming conditions (delayed and standard; Dutch and English) and that word frequency is the most important predictor in lexical decision. In addition, semantic variables have a substantial effect on Dutch lexical decision, and word frequency has a clear effect on English standard and delayed naming. The role of all other clusters of variables appears modest at the most.

Onset: Number of consonants or preferred onset structure? The variables ONS1 and ONS2 were highly correlated due to the fact that ONS1, the number of consonants in the word's onset, was one of the components encompassed in ONS2. The ONS2 variable was more general in that it categorized words in terms of how well their onsets agreed with universally preferred onset structures. In addition to the number of consonants in the onset, it took account of the fact that not all word onsets conform with the sonority principle. Because onset turns out to be an extremely important variable in word naming, especially in shallow Dutch, an interesting further question is which of the two onset variables has the stronger impact on performance. To answer this question, we performed two additional sets of regression analyses. In one of them RT in the four naming conditions served as criterion variables, and either ONS1 or ONS2 was entered as the only predictor variable. When only ONS1 was entered, the percentages of accounted variance in Dutch standard naming, Dutch delayed naming, English standard naming, and English delayed naming were 25, 19, 5, and 12%, respectively. However, when only ONS2 was entered these percentages were 36, 27, 10, and 14%, respectively. These latter percentages were equivalent in all cases to the accounted variance when both onset variables were entered (see Table 8). These results suggest that ONS2 is the more important variable.

In the second set of analyses all 18 variables except either ONS1 or ONS2 were entered as predictors. The question was how much the percentage of accounted variance dropped when either ONS1 or ONS2 was excluded. When all variables except ONS1 were entered, the percentages of accounted variance in the above four conditions were 50, 42, 40, and 33%, respectively (which in all but one case are equivalent to the percentages accounted variance when all 18 variables are entered). However, when all variables except ONS2 were entered, a substantial decrease in the percentages of accounted variance occurs: 40, 33, 36, and 30%, respectively. The results again suggest that ONS2 is the more important variable.

Multiple regression analyses on the factor scores of the Principal Component Analyses (PCAs). The preceding analyses that looked at the effect of the separate clusters of variables on RT involved a large number of multiple regression analyses. This implies a large risk of Type I errors. To substantiate the above conclusions in an analysis that does not suffer from this problem we ran a final set of six multiple regression analyses, now with the factor scores that emerged from the PCAs as predictors. This set of analyses at the same time provides a statistical test of the differential effects of the various principal components (PCs) within each of the language by task conditions and of the differential effect of a particular PC across the three tasks within a language.

As was shown in Table 3, the PCAs on both the English and the Dutch sets had extracted 6 PCs from the 18 original variables. Five of these could be labeled the same way in both language sets (length, meaning, onset, frequency, and NUMC3). The sixth concerned a different component in the two language sets (intensity in Dutch and neighborhood in English). Table 9 shows the results of the multiple regression analyses in which the factors extracted from the PCAs served as predictors and where RT in each of the six language by task conditions served as criterion variables. For each of the PCs in each of the language by task conditions a β coefficient (at the same time indicating the correlation) is given, as well as (in parentheses) the corresponding confidence interval (95%). An asterisk attached to a β coefficient indicates that the corresponding factor contributes significantly to the regression equation. The relative sizes of the β coefficients within each task by language condition indicate the relative importance of the factors. Non-overlapping confidence intervals of the β coefficients for each individual factor within a language condition (the "vertical" comparisons in Table 9) indicate that these coefficients differ significantly from one another. The squared β coefficients equal the amount of variance accounted for by each factor. Because the factors in these analyses are uncorrelated (unlike the clusters of variables on which the results in Table 8 are based), the squared β coefficients for each task by language condition add up to the total percentage of variance accounted for by the six factors. Table 9, last column, presents this percentage for each of the six conditions.

Dutch. The results generally confirm those of the analyses reported above. The most important factor in Dutch standard and delayed naming is the onset component. The other three factors that contribute significantly to prediction in both naming conditions are intensity, length, and NUMC3, in that order of importance. The frequency component affects only standard naming. The semantic component is significant in neither of the two Dutch naming tasks.

TABLE 9

The β Coefficients and Corresponding Confidence Intervals of Multiple Regression Analyses on the Factor Scores from the Principal Component Analyses

	I	II	III	IV	V	VI	AV
Dutch							
Standard naming	.20*	−0.01	.53*	−0.11*	.14*	.25*	41
	(3.8, 8.2)	(−2.6, 1.8)	(13.7, 18.1)	(−5.6, −1.2)	(2.2, 6.5)	(5.3, 9.6)	
Delayed naming	.21*	−0.00	.46*	−0.01	.15*	.27*	34
	(4.4, 9.6)	(−2.7, 2.5)	(13.1, 18.3)	(−3.0, 2.2)	(2.5, 7.7)	(6.7, 11.9)	
Lexical decision	−0.06	−0.20*	−0.06	−0.52*	−0.02	−0.05	31
	(−6.5, 0.9)	(−13.5, −6.1)	(−6.3, 1.1)	(−28.6, −21.2)	(−4.5, 2.9)	(−6.2, 1.2)	
English							
Standard naming	.25*	−0.20*	.29*	−0.17*	−0.33*	.16*	35
	(8.4, 15.8)	(−13.2, −5.8)	(10.2, 17.6)	(−11.8, −4.4)	(−19.4, −12.0)	(3.9, 11.3)	
Delayed naming	.05	−0.19*	.41*	−0.04	−0.27*	.05	28
	(−1.2, 5.6)	(−11.4, −4.5)	(14.0, 20.8)	(−5.0, 1.9)	(−15.0, −8.2)	(−1.4, 5.5)	
Lexical decision	.13*	−0.27*	−0.06	−0.04	−0.62*	.10*	48
	(3.6, 11.3)	(−19.3, −11.5)	(−7.3, 4.0)	(−6.3, 1.4)	(−39.2, −31.4)	(2.1, 9.8)	

Note. Dutch: Factor I, length/neighborhood; Factor II, meaning; Factor III, onset; Factor IV, frequency; Factor V, NUMC3; Factor VI, intensity. English: Factor I, length/intensity 2; Factor II, meaning; Factor III, onset; Factor IV, neighborhood; Factor V, frequency; Factor VI, NUMC3; AV, accounted variance.

Whereas the pattern of results is rather similar for the two Dutch naming conditions, Dutch lexical decision shows a completely different pattern of results. The most substantial effect in Dutch lexical decision is that of the frequency component, which plays a much more modest role in Dutch standard naming and none whatsoever in Dutch delayed naming. The only remaining factor that contributed to prediction was the semantic component.

How very different lexical-decision performance is from naming performance is also obvious from a comparison of the confidence intervals of the β coefficients associated with each of the factors across the three tasks. Table 9 shows that the confidence intervals of the β coefficients for five factors in lexical decision have no overlap with the corresponding confidence intervals in the two naming tasks. In contrast, the overlap is generally large for the two naming tasks.

English. Whereas in Dutch standard naming the semantic factor did not contribute to prediction, in English standard naming all factors contributed. The relative order of importance of the four factors that contribute significantly to prediction in both Dutch and English standard naming (frequency, onset, length, and NUMC3) differs between the two languages: Whereas frequency was relatively unimportant in Dutch standard naming, it is the most important factor in English standard naming. The most important factor in Dutch standard naming, onset, comes second in English standard naming. As in Dutch lexical decision, frequency is the most important factor in English lexical decision, followed by the semantic factor. Also, as in Dutch lexical decision, the onset factor plays no role in lexical decision. Unlike in Dutch lexical decision, in English lexical decision other factors than the frequency and semantic factors also play a role, namely, length and NUMC3.

A comparison of the β coefficients and confidence intervals for each of the factors across the three tasks shows that English lexical decision and standard naming have more in common than Dutch lexical decision and standard naming. In contrast, English standard and delayed naming have less in common than Dutch standard and delayed naming.[2]

DISCUSSION

We have focused on two main questions: (1) Does performance in the two tasks that are most commonly used to study word recognition respond in the same way to the present word manipulations? (2) Do the present word manipulations affect the processing of Dutch and English words by Dutch–English bilinguals in similar ways or do differences between the two languages occur? We found processing similarities as well as differences across tasks and across languages. The similarities can be summarized by saying that whenever a statistically significant effect occurred, the direction of this effect was generally the same in the two tasks and the two languages. Whenever an effect occurred, it was always the case that performance was better for words that are easy to imagine, to contextualize, and to define than for words hard to imagine, contextualize, and to define, respectively; performance was always better for cognates than for noncognates; it was always better for familiar and frequent words than for unfamiliar and infrequent words, respectively; it was always better for short words and words with many neighbors in the target and nontarget lan-

[2]A variable not included in our set but that has been shown to influence word recognition is spelling-to-phonology consistency (e.g., Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995). To see whether consistency also played a role in our study we selected all monosyllabic words ($N = 209$) from the English set, divided each of them in its onset and rime, and determined whether each rime was consistent or inconsistent in feedforward (from spelling to phonology) and feedback (from phonology to spelling) directions (Ziegler, Stone, & Jacobs, 1997). We assigned a word a 1 if its rime was consistent. Words with inconsistent rimes were assigned a 2. We then looked at the correlations between feedforward and feedback consistency as predictors and RT in each of the three tasks as criterion variables. In none of the tasks did a significant correlation between feedforward consistency and RT occur ($r = -0.04$, $r = .07$, and $r = .02$ for standard naming, delayed naming, and lexical decision, respectively). The feedback analyses showed small effects in the naming tasks ($r = .08$, $r = .10$, and $r = .00$, in the same order). The reason the effects, if they occur at all, are so small may be that the binary scores on consistency that we used may not be fine-grained enough to detect robust effects.

guage than for long words and words with few neighbors; performance was better for words with simple and structurally preferred onsets than for words with complex and structurally nonpreferred onsets; it was better for words that do not contain consonant clusters than for words that do contain them. But more salient than these similarities are the differences that can be detected when we scrutinize the data at a more detailed level.

Task Effects

The most salient difference between tasks occurred in Dutch, where in lexical decision frequency had a major impact on performance and where of all the remaining variables only the semantic variables played some role in that task. In contrast, in Dutch standard word naming the onset variables had by far the largest impact, and all the other clusters of variables except the semantic variables affected performance to some extent. In terms of the converging-evidence approach set forth earlier, these results indicate that only the small frequency effect in Dutch word naming reflects a word-recognition component that is shared by the two tasks. In other words, the “functional overlap” (Grainger & Jacobs, 1996) between the two tasks in Dutch is very small indeed.

In contrast, the functional overlap between Dutch standard and delayed naming is large: The patterns of results in Dutch delayed and standard naming were similar, except that the frequency effect in standard naming did not materialize in delayed naming. These findings suggest that the frequency effect in Dutch standard naming has the recognition component of the task as its only locus and that large parts of the remaining effects are localized in the output component of the task. Importantly, after delayed naming is partialled out, significant effects remain for the frequency variables, (some of) the length variables, the neighborhood, onset, and number-of-consonant-clusters variables, and (surprisingly) one of the intensity variables (surprisingly, because one would not expect the sound intensity of the onset to affect the actual recognition process). The effect of cognate status, however, disappeared, suggesting that this variable affects the production component of Dutch word naming. Finally, analyses that looked at the relative importance of the two onset variables showed that the preferred-onset-structure variable was especially important.

The difference between tasks are less extreme in English, where both in lexical decision and in standard naming effects of the length variables, the semantic variables, the frequency variables, and cognate status occur. Here again, frequency is by far the most important factor in lexical decision, and the onset factor only plays a role in standard naming (and, again, it was the preferred-onset-structure variable that was the most important of the two onset variables). The standard- and delayed-naming tasks again produced similar patterns of results, but the overlap between these two tasks was somewhat smaller than that in Dutch. The onset effect in English naming seems to have the production component of the tasks as its only locus (whereas in Dutch naming recognition also appeared to be affected by the onset structure) and frequency affects both the recognition and the output components of English naming (in Dutch it only appeared to affect recognition).

A clear task effect that occurred in both languages concerns the neighborhood variables: When the analyses are restricted to the short (three-, four, and five-letter) words in the stimulus sets, in both language conditions neighborhood effects occurred in standard naming but not in lexical decision.

To summarize, the two tasks that are most commonly used to study word recognition respond very differently to the present word manipulations and thus appear to reflect very different processes.

Language Effects

Perhaps the most salient language effect is that standard naming and lexical decision had much more in common in English than in Dutch. This conclusion can be drawn both from the fact that the correlation between these tasks was much larger in English than in Dutch and from the regression analyses on the factor scores, which showed a larger overlap between standard naming and lexical decision in English than in Dutch.

At a more detailed level, the following language effects were obtained: (1) frequency appeared to affect both the recognition and the production component in English naming, but only the recognition component in Dutch naming, (2) Semantic effects occurred in English naming but not in Dutch naming, (3) A cognate effect occurred in English lexical decision and naming. In Dutch a small cognate effect only occurred in naming, but—unlike in English—it appeared to have production as its only locus, (4) In Dutch but not in English responding was faster in standard naming than in lexical decision. The two tasks elicited equally long response times in English, (5) The onset effects in naming were larger in Dutch than in English. The difference appeared to be caused by the recognition component of naming, which was affected by onset structure in Dutch but not in English. The production component of the task was affected equally by onset structure in the two languages, (6) Length affected lexical decision in English but not in Dutch; similarly, length affected (the recognition component of) English naming more than Dutch naming. Length thus appears to affect the processing of English words more than it affects Dutch word processing.

Explaining the Data: Task Requirements, Orthographic Depth, and Language Proficiency

The task and language effects can be explained by taking into account that lexical decision and word naming require different processing operations, that Dutch and English differ in orthographic depth, and that our participants are somewhat more fluent in Dutch than in English. But before explaining the data that way, we must acknowledge that the validity of our conclusions regarding the locus of specific effects obtained in word naming depends on the appropriateness of our choice of SOAs between the stimulus and the cue to respond in the delayed-naming task, 550 ms in Dutch and 600 ms in English. This choice was based on two assumptions: That word recognition has been fully completed when the cue appears, while at the same time the response has not yet been prepared fully and stored in the output buffer. Nevertheless, others authors (e.g., Monsell et al., 1989) have argued that longer SOAs should be used. If they were right (but see Balota and Chumbley, 1990, for a retort) some of the present conclusions regarding the locus of a particular effect in word naming may be premature.

Task requirements. Balota and Chumbley (1984, 1990; Chumbley & Balota, 1984) argued that lexical decision is a word discrimination task rather than a word identification task and that participants performing the task exploit information about the familiarity (and meaningfulness) of words to discriminate words from nonwords. This exaggerates the effect of the frequency variables in lexical decision (see Balota & Chumbley, 1990, p. 232, for details) and is presumably the reason why in the present study the frequency variables stand out as by far the most important determinants of lexical-decision performance (see also Whaley, 1978). If the lexical-decision task is essentially a discrimination task, a lexical-decision response does not always require the identification of the presented word (see also Grainger & Jacobs, 1996) and, consequently, effects observed in the task do not necessarily provide information on the process of word identification.

Balota and Chumbley (1985) argue that the frequency effects observed in word naming also exaggerate the role of frequency in actual word identification, because the output component of the naming task may be affected by frequency as well. The results from the present study support that idea. However, the task comparisons reported here also suggest that the impact of frequency on the discrimination process in lexical decision is especially large.

Whereas the specific requirements of the lexical-decision task can account for the exceptionally large effects of frequency in lexical decision, the unique demands of the naming task provide an explanation for the exclusive role of the onset variables and the number-of-consonant-clusters variables in word naming. Recall that both onset variables reflected the number of consonants in the word's onset and that, in addition, ONS2 classified onsets in terms of their degree of agreement with the sonority principle. A plausible reason these variables are so important in naming is that complex onsets—and consonant clusters in general—and onsets that violate the sonority principle may be harder to pronounce than simple onsets and onsets that conform to the sonority principle. Of course, of the present two tasks, only the naming task (both its standard and its delayed versions) requires a pronunciation response. Hence the strong and exclusive impact of the onset variables in this task (the differential effects of the onset variables in English and Dutch will be discussed in the next section). Likewise, the fact that consonant clusters are relatively hard to pronounce provides an explanation for the exclusive effect of the number-of-consonant-clusters variables in word naming. If consonant clusters are especially hard to pronounce, the more of them in the word, the more word naming will be hindered.

A further finding that can be explained in terms of task-specific requirements concerns the effect of sound intensity of the onset. We used two measures to assess the onset's sound intensity, one in terms of average sound intensity in decibels of the initial phoneme (based on Fry, 1979) and a second in terms of word-stress assignment. Inadvertently, the latter measure was confounded with word length and will therefore be ignored. The effect of the remaining sound-intensity variable confirms the finding of other authors (Kessler et al., in press; Sakuma, Fushimi, & Tatsumi, 1997) that voice keys are not equally sensitive to all speech sounds. Only word naming, the only task that uses a voice key to register the response, showed clear effects of this variable. A comparison between the standard- and delayed-naming tasks suggests that the output component of word naming is the only locus of the effect, a finding that agrees with the commonsense prediction that a voice key does not respond to the mental word-recognition process.

Finally, the finding that word naming responds to the neighborhood variables but lexical decision does not may result from the specific requirements of the two tasks. As mentioned earlier, it is currently a topic of debate whether neighborhood effects in naming have a lexical source or result from an effect of neighborhood characteristics on nonlexical naming processes (Andrews, 1997, p. 449; Peereman & Content, 1995). The null effects of neighborhood size in lexical decision may suggest a nonlexical source of the effect in naming. This conclusion, however, does not tally with the suggestion in the naming data that the actual recognition component of the task is affected by neighborhood.

Orthographic depth. The finding that the semantic variables affected naming in English but not in Dutch replicates in bilinguals the results of the cross-language study by Katz and Feldman (1983). In that study a semantic-priming effect was obtained when English native speakers named English words, but not when native speakers of Serbo-Croatian named Serbo-Croatian words. Katz and Feldman took this finding as support for the orthographic-depth hypothesis, which states that the more orthographically shallow a language is, the larger the role of assembled, nonlexical phonology in generating naming responses. The differential effect of semantic variables in Dutch and English thus suggests that in Dutch, a relatively shallow language, but not in deeper English, the lexicon is bypassed in the naming process.

However, the conclusion that the lexicon is not involved at all in Dutch naming does not tally with the finding that the small frequency effect in this condition appeared to have the recognition component of the task as its locus. Apparently, lexical access does take place on at least a subset of the trials. The null effects of semantic variables combined with a small frequency effect can be reconciled by adapting the orthographic-depth hypothesis. Instead of assuming that the phonological assembly process bypasses the lexicon altogether, it may be assumed that, as the addressed route, it provides access into the mental lexicon (and as such is sensitive to the frequency manipulation), but that it stops short of semantic access. Naming words in a deeper language, such as English, necessarily exploits the addressed route relatively often, because an access attempt via the indirect assembly route will often fail. The semantic effects in this language indicate that word naming via the direct route, unlike naming via the indirect route, does pick up a word's semantic information in the lexicon. This specific view of the orthographic-depth hypothesis is a special case of the well-known dual-route model of reading (Colheart, Davelaar, Jonasson, & Besner, 1977; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), which assumes both of the above direct and indirect routes into the lexicon. The present data suggest that in a shallow language only one of these, the indirect route, is effective, even though both may be operative. A horse race version of the model where the indirect route delivers an output first could account for the data.

Two other language effects that might be explained in terms of this specific version of the orthographic-depth hypothesis are the different magnitudes of the onset effects in Dutch and English and the fact that the onset effects in Dutch naming appear to have two loci (both recognition and output), whereas in English naming they appear to be localized exclusively in the output component of the task. Above, we explained the onset effects in terms of the relative difficulty of pronouncing consonant clusters and phoneme sequences that violate the sonority principle. It is plausible that not only the process of preparing the articulatory code during the output component of the naming task is negatively affected by complex onsets and onsets that violate the sonority principle, but that the process of assembling the phonological code is likewise affected. The effect would be that under circumstances where the naming responses are based primarily or exclusively on the assembly route (Dutch) larger onset effects occur than when the direct-access route is also effective (English). If this analysis is correct, the finding that by far the largest part of the onset effects in English is localized in production, suggests that in this study the vast majority of all responses in English naming were based on the direct-access route.

A final result that can be explained in terms of a difference in orthographic depth between Dutch and English is that in all three tasks the responses were faster in Dutch than in English. We pointed out earlier that there are two potential sources of this effect, namely our participants' higher proficiency in their native language than in their nonnative language and the orthographic-depth difference between these two languages. A study by Paulesu et al. (2000) supports the latter explanation. In a cross-language study, these authors demonstrated faster word and nonword reading as well as the involvement of different brain areas in Italian students reading their native language Italian (an orthographically shallow language) than in English students reading their native language English (a deep language). They attributed this differential pattern of behavioral and brain responses to the difference in orthographic depth of Italian and English and concluded that "reading in a complex and inconsistent orthography comes at a considerable cost" (Paulesu et al., 2000, p. 93). In other words, the present finding replicates in bilinguals what Paulesu et al. (2000) observed for monolinguals, for whom an account in terms of language-proficiency differences does not hold.

One effect appears to refute an interpretation in terms of a difference in orthographic depth between Dutch and English: We predicted a larger effect of length (in number of letters) in Dutch than in English, assuming that a shallow language exploits the grapheme-to-phoneme assembly process (which we regarded the source of this length effect) more than a deeper language does. Contrary to prediction, length had a larger influence on the recognition component of English standard naming than of Dutch standard naming. Similarly, English lexical decision but not Dutch lexical decision responded to the length variable. If lexical access in the lexical-decision task, as in word naming, also exploits both a direct and an indirect lexical-access route (as the dual-route model holds), one would also have expected the opposite result in that task.

An analysis of the Dutch and English high-frequency and low-frequency words separately[3]

[3]We performed this analysis in an attempt to dissociate the language effects that arise from a language-proficiency difference between Dutch and English from those caused by orthography differences between Dutch and English. If some of the language effects are in fact proficiency effects, the pattern of results may be expected to differ between the sets of high-frequency words and low-frequency words within each of the two languages. The reason is that word

showed that length interacted with word frequency in the English word set, but not in the Dutch set. In fact, in lexical decision and standard naming the length effects for high-frequency English words did not differ from the length effects for both high- and low-frequency Dutch words. In other words, the language effects—larger (in naming) or exclusive (in lexical decision) length effects in English—only occurred for low-frequency words, that is, for the words our participants have practiced the least. This suggests that these language effects were in fact proficiency effects. Other proficiency effects are discussed in the next section.

What remains to be explained is the source of the length effects in naming high-frequency words. The fact that with these words the effect of length in letters does not vary with language depth suggests that maybe the assembly process is not the effect's locus. If access in lexical decision also exploits the assembly process at least some of the time, especially in shallow languages, the null effect of length in letters in that task suggests the same. An alternative candidate locus for the length effect in naming is the production component of the task, which may require more time to prepare longer words for output than shorter words. Because lexical decision does not involve an oral output, no effect of length is obtained in that task. If this account is correct, the number of phonemes and not the number of letters should be the relevant variable, because it is a sound pattern that needs to be produced.

Language proficiency. A number of results can be explained in terms of proficiency differences between Dutch and English in our participants. One of these is the finding that in Dutch but not in English, lexical decision took longer than standard naming. When native English speakers are tested on English materials, lexical decision does take longer than standard naming (Forster & Chambers, 1973; Frederiksen & Kroll, 1976), as we observed for Dutch. We argued that the task effect may not have materialized in English because of the requirement of the naming task to prepare the response for production, a processing component that we hypothesized to be relatively time-consuming in a weaker language. Our delayed naming data provided support for this view, as does the finding that English native speakers show the common response time difference between lexical decision and naming.

Another language effect that can be explained in terms of (and qualifies) the view that preparing a naming output was especially hard in nonnative English is the effect the two frequency variables exerted in English delayed naming but not in Dutch delayed naming. According to the logic of the delayed-naming task this finding indicates that the output component of English word naming (but not of Dutch naming) is affected by frequency. Apparently for our participants it is easier to prepare a frequent English word for production than an infrequent English word, whereas infrequent Dutch words are not especially hard to prepare for output. This result thus qualifies the above conclusion that for our not completely balanced bilinguals it is harder to prepare nonnative English words for output than native Dutch words: Particularly the infrequent English words cause production difficulties. This may be because these are, of course, the words that our participants have least practiced.

The finding that our participants took longer to prepare English words (especially infrequent ones) for production than Dutch words may impact on other bilingual studies, especially studies that have used tasks that require the response word to be output verbally. These include the picture-naming and word-translation tasks that have received a substantial amount of attention in the study of bilingualism. A number of studies have shown that translating words from the dominant to the weaker language takes more time than vice versa (but see De Groot et al., 1994; De Groot & Poot, 1997; La Heij, Hooglander, Kerling, & Van der Velden, 1996, for

frequency may be expected to mimic language proficiency within a language (a language user is relatively proficient in processing frequent words). Apart from the length effects in English, the effects of the remaining variables did not differ significantly between the sets of high-frequency and low-frequency words. The reason may be that the proficiency difference between Dutch and English in our participants is rather modest.

counterexamples), an effect that is one of the cornerstones of the “asymmetry” model of bilingual lexical representation (Kroll & Stewart, 1994; Sholl, Sankaranarayanan, & Kroll, 1995). The present findings suggest that such an effect of translation direction on translation time when it occurs at all may not only be caused by directional asymmetries in bilingual memory representations, as the asymmetry model holds, but that it results partly from the fact that output preparation is relatively hard in the weaker language. The common finding of slower response times when pictures are named in the weaker language than when named in the stronger language may also partly be caused by a difference in the ease of output preparation for the two languages.

A further language effect that may have resulted from our participants’ not completely balanced bilingualism concerns the language-dependent cognate effect that occurs in lexical decision and also, after delayed naming is partialled out, in naming: The cognate effect occurs for English, the weaker language, only. This finding of a language-dependent cognate effect replicates similar results in lexical-decision studies by Caramazza and Brones (1979) and Van Hell (1998, Chap. 4). As pointed out in the introduction, a cognate effect supports the view that access to the bilingual lexicon is nonselective. Apparently, the coactivated representation of the cognate in the nontarget language system facilitates the response to the target. The impact of this coactivated representation on the response plausibly depends on the relative degree of activation of the coactivated nontarget representation and the target representation. Especially when the nontarget representation is activated more than the target representation, the former may be expected to exert a clear influence on responding. This situation will generally hold when the nontarget language is the stronger of the bilingual’s two languages. The ground for this assumption is that the activation level of the stronger language (here Dutch) will be larger overall than that of the weaker language (English). For the present study the implication would be that the cognate effects should be larger in English (with stronger Dutch as the nontarget language) than in Dutch (with weaker English as the nontarget language), as indeed they are. In fact, the cognate effect does not occur at all in Dutch lexical decision, nor in Dutch naming after delayed naming was partialled out of standard naming.

In addition to the language-dependent cognate effect just discussed, the effect of a stimulus’ cross-language neighborhood size that we observed in word naming also supports the nonselective-access view (see also Grainger & Dijkstra, 1992; Van Heuven et al., 1998). Other support for nonselective access comes from studies that have shown that a language stimulus gives rise to automatic phonological activation of both of a bilingual’s two languages (or both of a bialphabetic or bidialectal language user’s two alphabets or dialects), even when the specifics of the task invite language-specific performance (e.g., Lam, Perfetti, & Bell, 1991; Lukatela, Savic, Gligorijevic, Ognjenovic, & Turvey, 1978; Tzelgov, Henik, Sneg, & Baruch, 1996).

Conclusion

We began this investigation with the goal of learning more about the word-recognition process than is currently available from the dominant type of studies in this research field, that is, studies that have tested English-speaking monolingual language users. Like the majority of other researchers in the field, we chose the lexical-decision and word-naming tasks as tools, on the assumption that similar effects in the two tasks inform us about the process of interest, word recognition. Instead of monolingual English language users, we tested Dutch–English bilinguals in both of their languages. The many task and language effects that we obtained provide a detailed picture of the differential sensitivity of each of the two tasks and of each of the two languages to the present variables. A strikingly large number of task differences occurred, especially in Dutch, where the data patterns obtained with the two tasks converged on nothing more than a small, shared frequency effect. In English the overlap between tasks was larger than in Dutch, but still many differences between tasks occurred. These findings are extremely interesting from a task analysis point of view, but they are also disquieting because they suggest that many research efforts in this field may not shed light on the process of word recognition after all. Instead, they may inform us about task performance in the specific task that is being used and that, in many cases, will not be of interest in its own right. This will particularly be the case when the rather unnatural lexical-decision task is being used. Similarly, the language effects that we obtained may shed light on task-specific processing more than on actual word recognition. In conclusion, it seems that the tools we have developed to study word recognition are not optimally suited for their job (see also, e.g., Balota & Chumbley, 1990) and may even be unsuited to studying word-recognition in a shallow language.

In addition to this worrisome conclusion, the present study provides the basis for the intrinsically interesting conclusion that bilingual language users flexibly adapt to the specific characteristics of the language they are currently processing. A number of the language effects that we observed were presumably caused by our participants' different fluency levels in Dutch and English, but other language effects were likely caused by the different orthographic depths of Dutch and English. Because forms of bilingualism where the orthographies of the two languages do not differ at all in terms of depth or type (alphabetic, syllabic, or logographic) are likely to be relatively rare, we believe that the present conclusion that bilinguals process their two languages in different ways holds for many other bilingual populations, and maybe for all. Apparently, the maxim "that most languages get the orthography they deserve" (Katz & Frost, 1992, p. 67) can be expanded with the statement that orthographies get the processing they invite.

REFERENCES

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review*, **4**, 439–461.

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

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, **10**, 340–357.

Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language*, **24**, 89–106.

Balota, D. A., & Chumbley, J. I. (1990). Where are the effects of frequency in visual word recognition tasks? Right where we said they were! Comment on Monsell, Doyle, and Haggard, 1989. *Journal of Experimental Psychology: General*, **119**, 231–237.

Baluch, B., & Besner, D. (1991). Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **17**, 644–652.

Caramazza, A., & Brones, I. (1979). Lexical access in bilinguals. *Bulletin of the Psychonomic Society*, **13**, 212–214.

Chumbley, J. I., & Balota, D. A. (1984). A word's meaning affects the decision in lexical decision. *Memory & Cognition*, **12**, 590–606.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, **108**, 204–256.

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). London: Academic Press.

Crystal, D. (1987). *The Cambridge encyclopedia of language*. Cambridge UK: Cambridge Univ. Press.

De Groot, A. M. B. (1989). Representational aspects of word imageability and word frequency as assessed through word association. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 824–845.

De Groot, A. M. B., Dannenburg, L., & Van Hell, J. G. (1994). Forward and backward word translation by bilinguals. *Journal of Memory and Language*, **33**, 600–629.

De Groot, A. M. B., Delmaar, Ph., & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: Support for nonselective access to bilingual memory. *Quarterly Journal of Experimental Psychology*, **53A**, 397–428.

De Groot, A. M. B., & Poot, R. (1997). Word translation at three levels of proficiency in a second language: The ubiquitous involvement of conceptual memory. *Language Learning*, **47**, 215–264.

Dijkstra, A., & Van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189–225). Mahwah, NJ: Erlbaum.

Dijkstra, T., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, **1**, 51–66.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, **12**, 627–635.

Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 361–379.

Fry, D. B. (1979). *The physics of speech*. Cambridge: Cambridge Univ. Press.

Fudge, E., & Shockey, L. (2000) *The Reading syllable database* [on-line]. Available: http://www.rdg.ac.uk/AcaDepts/cl/slals/sylls.htm.

Grainger, J., & Dijkstra, T. (1992). On the representation and use of language information in bilinguals. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 207–220). Amsterdam: North-Holland.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, **103**, 518–565.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, **36**, 3–15.

Henderson, L. (1982). *Orthography and word recognition in reading*. London: Academic Press.

Jakobson, R., & Halle, M. (1956). *Fundamentals of language*. The Hague: Mouton.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, **104**, 130–136.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, **21**, 60–99.

Katz, L., & Feldman, L. B. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **9**, 157–166.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 67–84). Amsterdam: North-Holland.

Kawamoto, A. H., & Kello, C. T. (1999). Effect of onset cluster complexity in speeded naming: A test of rule-based approaches. *Journal of Experimental Psychology: Human Perception and Performance*, **25**, 361–375.

Kessler, B., Treiman, R., & Mullennix, J. (in press). Phonetic biases in voice key response time measurements. *Journal of Memory and Language*.

Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **12**, 92–107.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, **33**, 149–174.

La Heij, W., Hooglander, A., Kerling, R., & Van der Velden, E. (1996). Nonverbal context effects in forward and backward word translation: Evidence for concept mediation. *Journal of Memory and Language*, **35**, 648–665.

Lam, A. S. L., Perfetti, C. A., & Bell, L. (1991). Automatic phonetic transfer in bidialectal reading. *Applied Psycholinguistics*, **12**, 299–311.

Lukatela, G., Savic, M., Gligorijevic, B., Ognjenovic, P., & Turvey, M. T. (1978). Bi-alphabetical lexical decision. *Language and Speech*, **21**, 142–166.

McRae, K., Jared, D., & Seidenberg, M. S. (1990). On the roles of frequency and lexical access in word naming. *Journal of Memory and Language*, **29**, 43–65.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General*, **118**, 43–71.

Paulesu, E., et al. (2000). A cultural effect on brain function. *Nature Neuroscience*, **3**, 91–96.

Peereman, R., & Content, A. (1995). Neighborhood size effect in naming: Lexical activation or sublexical correspondences? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 409–421.

Sakuma, N., Fushimi, T., & Tatsumi, I. (1997). Measurement of naming latency of Kana characters and words based on speech analysis: Manner of articulation of a word-initial phoneme considerably affects naming latency. *Japanese Journal of Neuropsychology*, **13**, 126–136.

Savage, G. R., Bradley, D. C., & Forster, K. I. (1990). Word frequency and the pronunciation task: The contribution of articulatory fluency. *Language and Cognitive Processes*, **5**, 203–236.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language*, **27**, 499–520.

Selkirk, E. (1985). The major class features and syllable theory. In M. Aronoff & R. Oehrle (Eds.), *Language sound structure* (pp. 105–136). Cambridge, MA: MIT Press.

Sholl, A., Sankaranarayanan, A., & Kroll, J. F. (1995). Transfer between picture naming and translation: A test of asymmetries in bilingual memory. *Psychological Science*, **6**, 45–49.

Smith, N. V. (1973). *The acquisition of phonology: A case study*. Cambridge: Cambridge Univ. Press.

Tabossi, P., & Laghi, L. (1992). Semantic priming in the pronunciation of words in two writing systems: Italian and English. *Memory & Cognition*, **20**, 315–328.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General*, **124**, 107–136.

Tzelgov, J., Henik, A., Sneg, R., & Baruch, O. (1996). Unintentional word reading via the phonological route: The Stroop effect with cross-script homophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **22**, 336–349.

Van Hell, J. G. (1998). *Cross-language processing and bilingual memory organization*. Ph.D. thesis, University of Amsterdam. ISBN 90-5470-075-0.

Van Hell, J. G., & De Groot, A. M. B. (1998). Disentangling context availability and concreteness in lexical decision and word translation. *Quarterly Journal of Experimental Psychology*, **51**, 41–63.

Van Heuven, W. J. B., Dijkstra, A., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, **39**, 458–483.

Whaley, C. P. (1978). Word-nonword classification time. *Journal of Verbal Learning and Verbal Behavior*, **17**, 143–154.

Ziegler, J. C., Stone, G. O., & Jacobs, A. M. (1997). What is the pronunciation for *-ough* and the spelling for /u/? A database for computing feedforward and feedback consistency in English. *Behavior Research Methods, Instruments, & Computers*, **29**, 600–618.

(Received August 10, 2000)
(Revision received July 11, 2001)