

Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks

JOHANNES C. ZIEGLER

CNRS and Université de Provence, Marseille, France

LUDOVIC FERRAND

CNRS and Université René Descartes, Paris, France

and

MARIE MONTANT

CNRS and Université de Provence, Marseille, France

In this study, we investigated orthographic influences on spoken word recognition. The degree of spelling inconsistency was manipulated while rime phonology was held constant. Inconsistent words with subdominant spellings were processed more slowly than inconsistent words with dominant spellings. This *graded consistency effect* was obtained in three experiments. However, the effect was strongest in lexical decision, intermediate in rime detection, and weakest in auditory naming. We conclude that (1) orthographic consistency effects are not artifacts of phonological, phonetic, or phonotactic properties of the stimulus material; (2) orthographic effects can be found even when the error rate is extremely low, which rules out the possibility that they result from strategies used to reduce task difficulty; and (3) orthographic effects are not restricted to lexical decision. However, they are stronger in lexical decision than in other tasks. Overall, the study shows that learning about orthography alters the way we process spoken language.

Spoken language has primacy over written language both in the history of mankind and in the life of an individual (Liberman, 1992). We talk before we read, and many languages do not even have a writing system. Yet, there is accumulating evidence that literacy (i.e., the learning of an orthographic system) has a long-lasting influence on basic phonological processes (Olson, 1996). For example, we have known for many years that certain metaphonological skills, such as our capacity to consciously manipulate phonemes, is largely dependent on literacy (Morais, Cary, Alegria, & Bertelson, 1979). For instance, a person who has never acquired an orthographic system finds it very difficult to delete a phoneme at the beginning of a nonword. In contrast, literate people have no difficulty with this task. Once reading is acquired, orthographic influences might be strong enough to override phonological information. For example, Ehri and Wilce (1980) have shown that literate children find it dif-

ficult to count the same number of phonemes in /ritʃ/ and /pitʃ/, because the spellings of these words (“rich” vs. “pitch”) contain different numbers of letters.

The influence of orthography on spoken word recognition and production has been demonstrated not only with children and illiterates but also with completely literate college students (Damian & Bowers, 2003; Frost & Katz, 1989; Frost, Repp, & Katz, 1988; Seidenberg & Tanenhaus, 1979; Taft & Hambly, 1985; Tanenhaus, Flanagan, & Seidenberg, 1980). For example, skilled adults find it harder to judge whether two words rhyme when their rimes are spelled differently (e.g., *rye-tie*) than when their rimes are spelled the same (Seidenberg & Tanenhaus, 1979). Similarly, Frauenfelder, Segui, and Dijkstra (1990) showed that detection of the phoneme /k/ in spoken French words took longer than detection of the phoneme /p/, supposedly because the phoneme /k/ in French has more orthographic realizations (“c,” “cc,” “k,” “ck,” “qu,” etc.) than the phoneme /p/ (see Dijkstra, Roelofs, & Fieuws, 1995, for a similar finding in Dutch). Finally, Halle, Chereau, and Segui (2000) recently demonstrated that French listeners were more likely to misperceive the phoneme /p/ in /apsyrd/ than in /lapsys/, because *absurde* is spelled with the letter “b” whereas *lapsus* is spelled with the letter “p.”

Although these orthographic effects clearly suggest that orthographic information can affect basic phonolog-

We are grateful to Madeleine Leveillé, Denis Lancelin, and Fanny Meunier for technical advice and assistance. Thanks are extended to José Morais, Régine Kolinsky, and two anonymous reviewers for helpful comments on a previous version of the manuscript. Correspondence concerning this article should be addressed to J. Ziegler, Laboratoire de Psychologie Cognitive, CNRS et Université de Provence, Case 66, 3 place Victor Hugo, 13331 Marseille Cedex 3, France (e-mail: ziegler@up.univ-mrs.fr).

ical processes, it has been argued that this would be the case only in tasks with a strong metaphonological component (Ventura, Kolinsky, Brito-Mendes, & Morais, 2001). Indeed, tasks such as phoneme deletion, detection, counting, and blending all rely on people's capacity to consciously access and manipulate phoneme-sized information. Finding orthography effects in these tasks might not be so surprising given the crucial role of orthography in developing awareness of phonemes (Goswami, 2002a; Morais et al., 1979; Olson, 1996).

However, orthographic effects have also been reported in tasks that do not necessarily involve metaphonological components. For example, Jakimik, Cole, and Rudnicky (1985) studied phonological priming in a lexical decision task (LDT). They employed polysyllabic primes and monosyllabic targets that shared either initial phonology (e.g., *chocolate–chalk*), initial orthography (e.g., *fighter–fig*), or both initial phonology and initial orthography (e.g., *napkin–nap*). Facilitatory priming effects were obtained only for prime–target pairs that shared both phonology and orthography. No priming was obtained when targets and primes shared only phonology or only orthography. However, the authors themselves acknowledged that their results might reflect an orthographic strategy, because the relatedness proportion between primes and targets was rather high (30–50%). More recently, orthographic effects in phonological priming have been replicated with relatedness proportions of only 20%, making it less likely that the facilitatory effects were due to orthographic strategies (Slowiaczek, Soltano, Wieting, & Bishop, 2003).

To avoid strategies of this kind, Ziegler and Ferrand (1998) did not use the priming paradigm but a simple auditory LDT. They manipulated the orthographic consistency of spoken words. Half of the words were *inconsistent*—that is, their phonological rimes could be spelled in multiple ways (e.g., /-ip/ may be spelled “-eap” or “-eep”)—and the other half was *consistent*—that is, their rimes could be spelled in only one way (e.g., /-ʌk/ may only be spelled “-uck”). Results showed that inconsistent words produced slower correct “yes” responses and more errors than did consistent words. This effect was taken to suggest that orthographic information was automatically processed during the perception of spoken words. These data provide strong evidence of the existence of orthographic effects in spoken word recognition, especially because their study minimized the use of orthographic strategies and did not rely on a metaphonological task.

There are two potential problems with Ziegler and Ferrand's (1998) study, however. First, the consistency manipulation necessarily relies on the comparison between different types of words (i.e., inconsistent vs. consistent words). It is therefore logically possible that inconsistent words differ from consistent words in terms of their phonetic composition as well. Although Ziegler and Ferrand matched their items on a number of variables (e.g., uniqueness point, number of phonemes, frequency), subtle phonetic differences might exist between

different groups of items. For example, some of the inconsistent items in Ziegler and Ferrand's study had quite complex nasalized sounds. Indeed, Kessler, Treiman, and Mullennix (2002) have recently shown that the phonetic composition of a word has an effect on voice key measurements. Moreover, the phonetic composition of a word has been shown to affect lexical decision performance (Abramson & Goldinger, 1997). Second, Ziegler and Ferrand relied on an auditory LDT. This task has been criticized for including a number of decisional and postperceptual processes that might amplify orthographic consistency effects (Ventura, Morais, Pattamadilok, & Kolinsky, 2004).

In the present study, we attempted to deal with problems of both the stimuli and the task. To make sure that the auditory consistency effect did not result from a confound in terms of phonetic differences between consistent and inconsistent items, we manipulated the degree of inconsistency of words sharing the same phonological rime. Thus, we selected pairs of inconsistent monosyllabic words that had the same phonological rime but differed in the probability with which their phonology maps onto spelling. For example, in English the words *sign* and *wine* have the same phonological rime: /-arn/. The “-ine” spelling is *dominant*—that is, it occurs in most monosyllabic words of this particular rime family (e.g., *fine, nine, vine, dine, line, mine*). In contrast, the “-ign” spelling is *subdominant*, meaning that it occurs in few or no other monosyllabic words. Both *sign* and *wine* are inconsistent in the sense that their phonological rimes map onto multiple spellings. However, *wine* contains the dominant spelling, whereas *sign* contains the subdominant spelling. Keeping the phonological rime constant while controlling for onset complexity should guarantee that the orthographic consistency effect is indeed due to orthography rather than to phonetic complexity.

The second problem, which is related to the use of the LDT, can be addressed by investigating the orthographic consistency effect in tasks other than lexical decision. For this purpose, we used a rime detection task in Experiment 2 and an auditory naming task in Experiment 3.

EXPERIMENT 1

In this experiment, we investigated whether or not the probability with which phonology maps onto spelling can influence auditory word recognition. To this end, we compared inconsistent words with dominant spellings with inconsistent words with subdominant spellings in an auditory LDT. Words with dominant and subdominant spellings had the same rime phonology and were yoked in frequency-matched pairs (e.g., *wine–sign*). In addition, a group of completely consistent words was included in the present study to replicate our previous finding of an auditory consistency effect.

Method

Participants. Twenty-five psychology students from a private University (École des Psychologues Praticiens) in Paris participated

in the experiment. All were native speakers of French. None of them reported any hearing problems. Twenty additional students from the same population gave familiarity ratings of the stimulus material.

Stimuli and Design. The critical stimulus set consisted of 72 low-frequency monosyllabic French words (see the Appendix). They were subdivided into three groups of 24 stimuli each: (1) consistent words, (2) inconsistent words with dominant spellings, and (3) inconsistent words with subdominant spellings. Inconsistent words were yoked in frequency-matched pairs in which one member had the dominant spelling and the other had the subdominant spelling. For example, in French the words *nain* (dwarf) and *daim* (deer) have the same phonological rime: /-ɛ/. The “-ain” spelling is dominant, whereas the “-aim” spelling is *subdominant*.

Consistency and spelling probability (dominance) were determined on the basis of consistency analyses of spelling and sound in French (Ziegler, Jacobs, & Stone, 1996). Consistency ratios (token and type) for the three stimulus groups are provided in Table 1. Consistency ratios vary between 0 (inconsistent) and 1 (consistent). They are calculated by dividing the number (or frequency) of friends by the number (or frequency) of friends and enemies. In the present context, a *friend* is a word whose rime phonology is spelled the same (e.g., *gear* and *fear* are friends of *hear*). An *enemy* is a word whose rime phonology is spelled differently (e.g., *beer* and *mere* are enemies of *hear*). As can be seen in Table 1, words with subdominant spellings have very few friends and many enemies. In contrast, words with dominant spellings have approximately the same number of enemies as do words with subdominant spellings, but many more friends. As a consequence, their consistency ratios are much higher. Finally, consistent words have no enemies at all.

Because we were interested in auditory word recognition, our main focus was on spelling consistency (i.e., how consistently a phonological pattern can be spelled). This is different from the domain of reading, for which most studies manipulated pronunciation consistency (i.e., how consistently an orthographic pattern can be pronounced). Indeed, all but five of our words (*gang*, *scout*, *jus*, *jet*, *stand*) were fully consistent in terms of pronunciation. Thus, pronunciation consistency is not further taken into account in any of the following analyses.

The fact that both members of each inconsistent word pair shared the same rime phonology effectively reduces the likelihood that phonological, phonetic, or phonotactic differences between dominant and subdominant members could be responsible for the orthographic effects. However, having identical rime phonology leaves the possibility open that stimulus groups differed in terms of phonotactic complexity or consistency of the word onset (initial phoneme or phoneme cluster). We therefore calculated initial phone frequency and transitional probabilities for the onset phonemes (see, e.g., Vitevitch & Luce, 1998, 1999). In addition, we verified whether the onsets of our items differed in terms of pronunciation and spelling consistency (type counts). As can be seen in Table 1, the three groups did not differ significantly on any of these variables.

The three stimulus groups were matched on word frequency according to the BRULEX database (Content, Mousty, & Radeau, 1990). The frequency matching was double-checked using a recent database containing more up-to-date frequency norms (LEXIQUE; New, Pallier, Ferrand, & Matos, 2001). As can be seen in Table 1, the three groups were still tightly matched even when the more contemporary frequency norms were used. To complement the objective frequency measures, we also obtained subjective familiarity ratings (see, e.g., Peereman, Content, & Bonin, 1998). Twenty students who had not participated in the experiment rated familiarity using a 7-point scale on which 1 was *very unfamiliar* and 7 *very familiar*. No significant differences were obtained in terms of familiarity.

The stimuli were further matched on a number of standard word recognition variables, such as phonological neighborhood density, number of phonemes, and uniqueness point. These measures were taken from the LEXIQUE database (New et al., 2001). The groups did not differ on any of these variables. The only significant difference was in terms of orthographic neighborhood density. Words with subdominant spellings had on average fewer orthographic neighbors than words in the other two groups. This confound cannot be avoided, because words that are highly inconsistent in spelling have, by definition, few orthographic neighbors. However, a recent study suggests that the important factor in explaining orthographic effects on spoken word recognition is the consistency of the phonology-orthography mapping rather than the sheer num-

Table 1
Item Characteristics and Statistical Significance of the Three Stimulus Groups
(Consistent, Inconsistent/Dominant, and Inconsistent/Subdominant)

	Stimulus Group			Significance		
	Consistent (A)	Dominant (B)	Subdominant (C)	A vs. B	B vs. C	A vs. C
Frequency (Brulex)	15.1	15.8	16.0	n.s.	n.s.	n.s.
Frequency (Lexique)	13.8	15.7	13.6	n.s.	n.s.	n.s.
Subjective familiarity	4.97	4.45	4.42	n.s.	n.s.	n.s.
No. of phonemes	3.33	3.13	3.17	n.s.	n.s.	n.s.
No. of letters	4.83	4.75	4.46	n.s.	n.s.	n.s.
Uniqueness point	3.33	3.13	3.13	n.s.	n.s.	n.s.
Phonological neighbors	14.04	15.88	15.96	n.s.	n.s.	n.s.
Orthographic neighbors	6.71	6.58	3.46	n.s.	<.001	<.001
Initial phone frequency	446.82	432.26	442.29	n.s.	n.s.	n.s.
Onset transitional probability	.0076	.0079	.0065	n.s.	n.s.	n.s.
Onset consistency (pronunciation)	.96	.98	.95	n.s.	n.s.	n.s.
Onset consistency (spelling)	.94	.95	.84	n.s.	n.s.	n.s.
Variables Manipulated						
Rime consistency (token)	1.00	.50	.15	<.0001	<.0001	<.0001
Rime consistency (type)	1.00	.66	.05	<.0001	<.0001	<.0001
No. of friends	5.75	6.79	2.16	n.s.	<.0001	<.0001
Frequency of friends	122.72	1,460.62	52.69	<.0001	<.0001	<.01
No. of enemies	0.00	15.66	10.79	<.0001	<.15	<.0001
Frequency of enemies	0.00	1,824.25	3,232.19	<.02	n.s.	<.001

ber of orthographic neighbors (see Ziegler, Muneaux, & Grainger, 2003).

For the purpose of the LDT, 72 nonwords were created by randomly changing onsets and phonological rimes of a list of 72 words that were not used in the present study. No consistency manipulation was performed on the nonwords. Note that it would not have been possible to create nonwords with subdominant spellings because nonwords have no lexical spellings. Thus, the spelling that a participant might think of in an auditory task is likely to be the dominant one.

Following Goldinger's (1996) recommendation for research using the LDT, we measured auditory durations of the stimulus recordings. Both types of inconsistent words had nearly identical durations (427 and 431 msec for words with dominant and subdominant spellings, respectively). This difference was not significant [$t(23) = .10, p > .90$]. However, completely consistent words had significantly longer auditory durations than both types of inconsistent words (492 msec; both $t_s > 1.82; p_s < .04$, one tailed). This is not a problem, however, because the consistency effect appears as a latency advantage for consistent words. Thus, longer auditory durations for consistent words would work against finding a consistency effect. Nevertheless, in this and the following experiment, auditory duration is used as a covariate.

Procedure. All items were recorded by a male native French speaker in a soundproof room on a digital audiotape recorder (Teac DA-P20) using a Sennheiser MD43 microphone. Items were digitized at a sampling rate of 48 kHz with 16-bit analog-to-digital conversion using a Macintosh II FX computer and an audiomedica editor. The stimuli were transferred to the left channel of the Teac DA-P20 digital-to-analog converter. A 1000-kHz pulse on the right channel (inaudible to the participants) was time locked with the onset of the stimulus. This signal started the timer; the timer was stopped by the participants' lexical decision responses. The intertrial interval was 2 sec.

Stimulus presentation and data collection were controlled by a 386 PC connected to the digital-to-analog converter. The stimuli were presented to the participants at a comfortable level through a pair of Sony MDR-P1 headphones connected to the left channel of the digital-to-analog converter. The participants were tested individually in a soundproof room. They were instructed to decide as quickly as possible whether or not the auditorily presented stimulus was a real French word. The participants gave their responses by pressing either the "yes" or the "no" button of a button box that was placed in front of them. The dominant hand was used for the "yes" responses. At the beginning of the experiment, the participants were given 20 practice trials to familiarize themselves with the task. During the experiment, no feedback was provided. The experiment lasted approximately 15 min.

Results

Mean correct reaction times (RTs) and error rates for consistent, inconsistent/dominant, and inconsistent/subdominant words are presented in Table 2. RTs were measured from target onset to response onset. RTs longer

than 3 *SDs* beyond a participant's global mean were excluded from the analysis (accounting for less than 1.5% of the data). Analyses of variance (ANOVAs) and planned comparisons were performed on correct mean RTs and error rates with both participants (F_1) and items (F_2) as the random variables. Consistency (i.e., consistent vs. dominant vs. subdominant) was treated as a within-subjects factor in the analysis by subjects and as a between-items factor in the analysis by items. Because of differences in auditory length between consistent and inconsistent items, auditory duration was used as a covariate in all F_2 analyses. We corrected means for auditory duration by using it as a covariate.

As can be seen in Table 2, the data exhibited an overall consistency effect: Performance was fastest for consistent words, intermediate for inconsistent/dominant words, and slowest for inconsistent/subdominant words. The main effect of consistency was significant on both RTs [$F_1(2,48) = 72.49, p < .001; F_2(2,68) = 21.07, p < .001$] and errors [$F_1(2,48) = 18.81, p < .001; F_2(2,69) = 8.1, p < .001$].

More importantly, a planned comparison between items with dominant and those with subdominant spellings confirmed that there was a significant spelling probability effect. That is, there was a significant (41-msec) difference between inconsistent words with dominant spellings and those with subdominant spellings [$F_1(1,24) = 55.99, p < .001; F_2(1,45) = 13.29, p < .001$]. In the error data, the 2.7% difference between inconsistent words with dominant spellings and those with subdominant spellings failed to reach significance [$F_1(1,24) = 3.69; F_2(1,46) = 1.05$]. Planned comparisons also confirmed that performance on consistent words was better than performance on either of the inconsistent word types (all $F_1s > 33.79, p_s < .001$; all $F_2s > 4.2, p_s < .05$).

Discussion

The present data can be summarized as follows. First, we replicated the auditory consistency effect previously obtained by Ziegler and Ferrand (1998). That is, both types of inconsistent words obtained longer RTs and more errors than did consistent words. More importantly, inconsistent words with subdominant spellings produced longer RTs and more errors than did inconsistent words with dominant spellings. Given that these two groups of inconsistent words had the same phonological rime and did not differ in terms of onset complexity or consistency, it is unlikely that subtle phonological or phonotactic differences were responsible for the observed orthographic effects.

One potential problem with holding rime phonology constant is that rimes were necessarily repeated. Previous researchers (Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Taraban & McClelland, 1987) noticed that the consistency effect in reading was amplified when the experiment contained the same orthographic rime pronounced in different ways (e.g., *give-five*; *have-cave*). In analogy, one could argue that repeating the rimes in

Table 2
Mean Latencies (RTs, in Milliseconds) and Percentages of Errors (%Error) for the Three Groups of Items in the Lexical Decision Task of Experiment 1

Group	RT		%Error
	<i>M</i>	<i>SE</i>	
Consistent	741	7.8	12.8
Inconsistent/dominant	770	7.6	19.6
Inconsistent/subdominant	811	7.6	22.3

the present experiment might have amplified the auditory consistency effect. However, we do not think that such intralist priming can explain the present consistency effect, for two reasons. First, in our original study (Ziegler & Ferrand, 1998), rimes were not repeated and yet the size of the observed consistency effect (62 msec) was almost identical to the size of the consistency effect in the present experiment (61 msec). Second, in the area of reading, subsequent research by Jared, McRae, and Seidenberg (1990) suggested that unreliable consistency effects in previous studies were not due to the absence of intralist priming but rather to weak consistency manipulations (i.e., many friends and few enemies). In summary, we believe that the graded nature of the consistency effect obtained in the present study reflects the statistical mapping between phonology and orthography rather than intralist priming or phonetic confounds.

Before these conclusions can be accepted, however, one further potential criticism needs to be taken into account. In this and the earlier study (Ziegler & Ferrand, 1998), the auditory LDT was employed to investigate the orthographic consistency effect. One could argue that the LDT is so difficult that participants might try to "visualize" the spoken words in order to improve task performance. Indeed, the LDT produced error rates on the order of 20%, which clearly suggests that it is not an easy task. Thus, it cannot be ruled out that participants use an orthographic checking mechanism in a strategic way in order to improve overall performance. To address this potential problem, we attempted to replicate the present effects in a less difficult auditory task.

EXPERIMENT 2

A rime detection task was used in this experiment to investigate whether the orthographic consistency effect would persist in a task that produced a much lower error rate than the LDT. The rime detection task seemed to be a good choice because (1) it is a purely phonological task (i.e., a participant does not have to be literate to do the task), (2) it should be quite easy because the rime, unlike the phoneme, is an easily accessible unit in speech perception (Goswami, 1999; Kirtley, Bryant, MacLean, & Bradley, 1989), and (3) in previous studies (e.g., Seidenberg & Tanenhaus, 1979) orthographic effects on rime judgments have been found. Thus, on each trial, participants were presented auditorily with a target rime followed by the target word. On half of the trials, the target rime was present in the word, and on the other half the target rime was absent. We predicted that detecting a rime in an inconsistent word should take longer than doing so in a consistent word, especially if the inconsistent word contained the less frequent (subdominant) spelling.

Method

Participants. Thirty-seven psychology students from the same population as those in Experiment 1 participated in Experiment 2.

None of them had participated in Experiment 1. All were native speakers of French. None of them reported hearing problems.

Stimuli and Design. The stimulus set and design were identical to those of Experiment 1 except that two items (*hausse* and *hall*) had to be replaced because their onsets were silent. They were replaced by *grosse* and *goal*, respectively. In addition, 72 new words were selected for the target-absent trials. They were matched to the target words in terms of frequency, onset, and number of phonemes (all $F_s < 1$). New recordings were needed for the target-absent trials and the rime targets; we also obtained new recordings for the critical target-present trials. In these new recordings, auditory length was again perfectly matched for the critical comparison between items with dominant and subdominant spellings (422 and 418 msec, respectively). The consistent items, however, were again much longer than the inconsistent items (559 msec). Thus, as in the previous experiment, auditory length was used as a covariate in all item analyses.

Procedure. Items were recorded by a male native French speaker using the same procedure as in Experiment 1. The stimuli were sampled at 44 kHz and then transferred to a computer at a sampling rate of 22.05 kHz and 16-bit conversion using SoundEdit 16 (MacroMedia, San Francisco). Stimulus presentation and data collection were controlled by a Macintosh computer and PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). The stimuli were presented to the participants at a comfortable decibel level through a pair of Sony MDR-P1 headphones. At the end of the auditory rime and after a delay of 50 msec, the target was presented. The participants were instructed to judge as quickly and as accurately as possible whether the auditorily presented rime was present or absent in the following French word. The participants gave their responses by pressing either the "yes" or the "no" button of a button box that was placed in front of them. The dominant hand was used for the "yes" responses. The participants were tested individually in a soundproof room. As in Experiment 1, the participants were first given 20 practice trials. No feedback was provided during the experiment. The experiment lasted approximately 20 min.

Results

Mean latencies and error rates are presented in Table 3. Only target-present trials were considered for statistical analyses. Two items and their matched controls had to be excluded because of systematic misperceptions, probably due to problems with the recordings. RTs that were more than 3 SDs beyond a participant's global mean were excluded (accounting for less than 1% of the data). The remaining data were submitted to a global (consistent vs. dominant vs. subdominant) ANOVA and to planned comparisons with both participants (F_1) and items (F_2) as the random variables. Because of differences in auditory length between the items, auditory duration was used as a covariate in all F_2 latency analyses. We corrected means for auditory duration by using it as a covariate.

Table 3
Mean Latencies (RTs, in Milliseconds) and Percentages of Errors (%Error) for the Three Groups of Items in the Rime Detection Task of Experiment 2

Group	RT		%Error
	<i>M</i>	<i>SE</i>	
Consistent	879	14.0	1.8
Inconsistent/dominant	920	12.6	3.1
Inconsistent/subdominant	947	12.7	3.4

The global ANOVA exhibited a significant effect of consistency on both RTs [$F_1(2,72) = 5.04, p < .01$; $F_2(2,62) = 5.77, p < .01$] and errors [$F_1(2,72) = 3.59, p < .05$; $F_2(2,63) = 2.52, p < .10$]. More importantly, planned comparisons showed a significant difference between words with dominant spellings and those with subdominant spellings in the latency data [$F_1(2,72) = 6.12, p < .02$; $F_2(1,41) = 2.58, p < .12$]. Moreover, consistent items were processed significantly faster than both types of inconsistent items ($p < .004$). No significant differences were obtained in the error data.

Discussion

The data from the rime detection task replicated the major effects found in the LDT—that is, the existence of a graded consistency effect and a significant difference between words with dominant and subdominant spellings (e.g., *wine–sign*). These effects were obtained despite the fact that the overall error rate was very low (about 3%). This provides evidence against the hypothesis according to which orthographic information would become activated only when the task is relatively difficult, making it useful for participants to encode spoken words in different formats (orthographic and phonological). Alternatively, the present data show that orthographic information is activated even in a task that produces a low error rate.

One potential problem with the rime detection task, however, is that it might involve metaphonological skills to some extent. Metaphonological skills are typically assessed by having people delete, add, or substitute phonemes (Morais et al., 1979). It is quite well documented that orthographic information affects such metaphonological skills, possibly by providing an abstract model of how to analyze speech into its constituents (Olson, 1996). Although rime detection does not involve the conscious manipulation of phonemes, the aim of the following experiment was to replicate the orthographic consistency effect in a task that clearly does not involve metaphonological components.

EXPERIMENT 3

The goal of this experiment was to replicate the graded consistency effect in an auditory naming task (see Liu, Bates, Powell, & Wulfeck, 1997). In this task, participants are presented with spoken words and are asked to repeat them as quickly as possible. This task was chosen because it still produces a low error rate, does not involve a decision component, and does not involve metaphonological judgments. It is obviously risky to use such a shallow task, because classic word recognition effects tend to be reduced in word repetition (Radeau & Morais, 1990). However, given that a number of effects (e.g., neighborhood effects) are not fully eliminated in repetition tasks (Vitevitch, 2002; Vitevitch & Luce, 1998), we still expected to find signatures of consistency effects, at least in the strongest comparison (i.e., consistent vs. inconsistent subdominant words).

Note also that the task contains a naming component. Ideally, in naming tasks onsets have to be matched across experimental conditions because voice keys trigger differently depending on the nature of the onset (e.g., much later for words with initial fricatives than for words with initial plosives; see Rastle & Davis, 2002). Because we did not match for onsets in this experiment, we added a delayed naming task after each trial (see, e.g., McRae, Jared, & Seidenberg, 1990). Because articulatory differences would affect delayed naming as much as immediate naming, delayed naming latencies can be used to factor out potential articulatory differences if necessary.

Method

Participants. Twenty-five psychology students participated in the experiment. They were recruited from the same population as those in the previous experiments. None had participated in either of the previous two experiments. All were native speakers of French. None of them reported hearing problems.

Stimuli and Design. The stimulus set and design were identical to those of Experiment 1, except that only word trials were used in the present experiment.

Procedure. As in the previous experiments, the stimuli were presented binaurally via headphones. The participants were instructed to listen carefully to each word and to say it aloud as quickly and accurately as possible. The delayed naming task was added after each trial. That is, after the immediate naming response was given, the participants were asked to say the same word aloud again. However, they were also instructed to delay their responses until a response signal (a 400-Hz beep) occurred. To avoid anticipations, the interval between the offset of the target and the onset of the beep was varied (1,400, 1,600, 1,800, or 2,000 msec). The participants were told to keep the word “on the tip of their tongues” during that interval, ready to name it as quickly as possible and as soon as the response signal occurred. Response latencies in the delayed naming trials were measured from the onset of the beep until the participant’s response triggered the voice key. The experiment was controlled using DMDX software (Forster & Forster, 2003). This software provides RT measurements and records a digital version of the participants’ responses.

Results

The results are presented in Table 4. Voice key errors were excluded from the data analyses (accounting for 1.7% of the data). Overall, the error rate was extremely low (<2%) and did not show any differences between conditions. However, one item from the inconsistent/dominant group (*lobe*) produced an error rate of 44% and was therefore excluded from all further analyses. RTs that were 3 *SDs* beyond a participant’s global mean were excluded from the analysis (accounting for less than 1.5% of the data).

With regard to immediate naming, the main effect of consistency was significant by subjects [$F_1(2,48) = 12.7, p < .001$] but failed to reach significance by items [$F_2(2,67) = 1.28, p > .20$]. The critical comparison between inconsistent dominant and inconsistent subdominant members of the pairs was significant by subjects [$F_1(1,24) = 5.0, p < .05$] but failed to reach significance by items [$F_2(1,44) = 1.1, p > .20$]. In the strongest planned comparison, the difference between inconsistent subdominant and fully consistent items was significant by

Table 4
Mean Latencies (RTs, in Milliseconds) and Percentages of Error (%Error) for the Three
Groups of Items in the Immediate and Delayed Auditory Naming Tasks of Experiment 3

Group	Immediate Naming			Delayed Naming		
	RT		%Error	RT		%Error
	<i>M</i>	<i>SE</i>		<i>M</i>	<i>SE</i>	
Consistent	827	8.7	1.7	346	11.4	1.0
Inconsistent/dominant	836	8.3	1.9	347	11.7	1.2
Inconsistent/subdominant	847	8.1	1.7	348	12.0	1.3

subjects [$F_1(1,24) = 7.9, p < .01$] and approached significance by items [$F_2(1,45) = 2.0, p < .20$]. The delayed naming condition produced no significant effects (all $F_s < 1$), suggesting that there were no major articulatory onset differences between the different item groups.

Discussion

In the present experiment, we found that orthographic consistency had some effect on immediate naming latencies. Note that this effect was neither masked nor amplified by potential articulatory differences between items because the delayed naming task did not show a difference between conditions. The effect in immediate naming was much smaller than that obtained in the previous experiments. Also, the critical comparison between items with dominant spellings and those with subdominant spellings was statistically less robust—that is, it failed to reach significance in the item analysis. However, note that the item analysis tends to underestimate effect sizes in research designs, in which items are tightly matched on a large number of dimensions (Raaijmakers, Schrijnemakers, & Gremmen, 1999). This possibility would certainly apply to the present research, because items were selected as a pair and matched on a large number of variables (see Table 1).

Also, as was mentioned earlier, auditory naming is a fairly low-level articulatory task that tends to minimize lexical influences (Radeau & Morais, 1990). This certainly limits our chance of finding a strong consistency effect because, in the consistency effect, it is the *lexically based* spelling of an inconsistent word (e.g., *sign*) that is in conflict with the more dominant, phonologically based spelling (e.g., “-ine”).

GENERAL DISCUSSION

The present experiments provide a strong test of the hypothesis that orthographic information influences auditory word recognition. Whereas in previous studies fully consistent words were compared with maximally inconsistent words (Ventura et al., 2004; Ziegler & Ferrand, 1998), in this study we compared inconsistent words, which had either a dominant or a subdominant sound-spelling mapping (i.e., a graded consistency effect). The advantage of such a research strategy was that items in the critical (dominant-vs.-subdominant) comparison had identical rime phonology. Keeping rime phonology constant while matching onset characteristics reduces po-

tential phonological or phonotactic differences between different groups of items.

The main results can be summarized as follows. In all three experiments, consistent items were processed faster than inconsistent items (the consistency effect). Moreover, all the experiments showed that inconsistent dominant items were processed faster than inconsistent subdominant items (the spelling probability effect). The relative effect sizes of the consistency and spelling probability effects across the three tasks are presented in Figure 1. Because the three tasks differ in overall latency, *z* scores were derived from the group differences expressed in *SD* units.

Figure 1 clearly illustrates that the size of the consistency and the spelling probability effects varies as a function of lexical involvement across tasks. In the LDT, lexical access is required to discriminate words from pseudowords. Indeed, both effects were strongest and most robust in the LDT. In the rime detection task, access to the lexicon might be beneficial for segmenting words into onsets and rimes, but it is not strictly required for performing the task. In accordance with this idea, the data exhibited smaller consistency and spelling probability effects in rime detection than in lexical decision. Finally, auditory naming does not require lexical processes at all, since it can be easily accomplished on a purely sublexical level. Not surprisingly, auditory naming showed the smallest effects, reaching significance only by participants. One might have expected that slower tasks, such as rime detection, would produce the largest effects because consistency effects might need time to develop. However, this was not the case. As the *z* score data clearly show, lexical involvement rather than overall processing speed predicts the size of the consistency effects.

Although there is strong evidence for the existence of orthographic effects in a number of auditory tasks (e.g., Damian & Bowers, 2003; Frost & Katz, 1989; Halle et al., 2000; Slowiaczek et al., 2003; Taft & Hambly, 1985), the question remains as to why orthographic effects should be there in the first place, given the primacy of speech over reading and writing. The most straightforward answer is provided by the dynamical systems approach to word recognition (Rueckl, 2002; Stone, Vanhoy, & Van Orden, 1997; Stone & Van Orden, 1994; Van Orden, 2002; Van Orden & Goldinger, 1994). According to this approach, word recognition results from the recurrent flow of activation within a highly interactive network representing

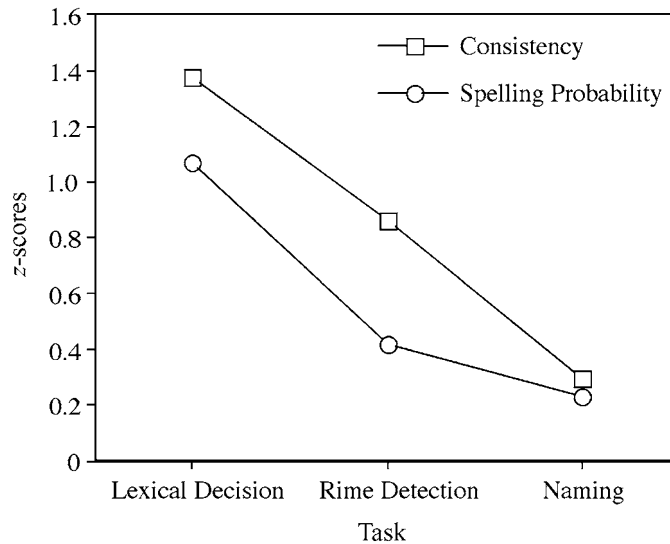


Figure 1. Cross-task comparisons of the consistency effect (i.e., inconsistent items minus consistent items) and the spelling probability effect (i.e., items with subdominant spelling minus items with dominant spelling). Effect sizes are expressed in *z* scores (i.e., group differences divided by pooled *SDs*).

different kinds of linguistic information (e.g., orthographic, phonological, semantic). Inconsistencies between these different units increase the time it takes for the system to settle into a stable state. Because learning in these networks is sensitive to the statistical properties of the spelling-to-sound mapping, these networks are well prepared to account for the graded nature of the consistency effect.

Orthography and phonology become intricately related during learning to read and to spell. One can describe the amalgamation of orthography and phonology in terms of the development of a cell assembly (see, e.g., Hebb, 1988). When different events co-occur frequently (as do orthography and phonology during reading), neurons coding those events become associated. Associated neurons develop into a functional unit, or cell assembly. An intriguing property of such cell assemblies is that if a sufficiently large number of the assembly neurons are stimulated by external input, activity will spread to additional assembly members and, finally, the entire assembly will be active. This explosion-like process has been called *ignition* of the assembly (Braitenberg, 1989). For example, full ignition of a cell assembly after stimulation of only some of its neurons can explain why perceivers have no problem identifying a partially hidden object (Hebb, 1988). As far as word recognition is concerned, the acquisition of reading causes strong associations between orthography and phonology to exist. As a consequence, word form assemblies include both orthographic and phonological information. Full ignition of the entire word form assembly after stimulation (by speech input) of a sufficiently large number of phono-

logical neurons may thus be the reason why orthography is automatically activated during speech recognition.

A second possible explanation for the existence of orthographic effects in speech perception is that they go back to a developmental phase in which orthographic information alters the nature of the phonological representations themselves (Goswami, 2002b). Such a possibility is best described in the context of the phonological restructuring model (Metsala, 1997; Metsala & Walley, 1998). Although this model does not take orthographic influences into account, there is evidence to suggest that as children learn to read and spell, orthographic information is used in order to restructure, specify, and organize lexical phonological representations. According to this idea, words with inconsistent spellings would never acquire fully specified phonological representations, which would explain the latency costs for accessing inconsistent words. This account would naturally predict that the size of the consistency effect decreases as the task relies less and less on accessing lexical representations, as indeed is the case in the present study. Note also that the restructuring account does not have to rely on on-line feedback to explain the existence of an orthographic effect (for discussions, see Norris, McQueen, & Cutler, 2000; Peereman et al., 1998; but see also Luce, Goldinger, & Vitevitch, 2000). Instead, orthographic information would have had its major impact during a developmental restructuring phase of lexical representations (i.e., off line).

While we await further data to decide between these challenging theoretical explanations, the present study clarifies a number of outstanding empirical issues. First,

orthographic consistency effects are not artifacts of phonological or phonotactic properties of the stimulus material. Second, they can be obtained in tasks that produce low error rates, which makes it unlikely that the effects are due to orthographic strategies used to reduce the difficulty of the task. Finally, orthographic consistency effects are not restricted to the LDT. However, their size seems to vary as a function of the lexical involvement of the task. Empirical specifications such as these are certainly necessary for a deeper understanding of the effect. However, they should not distract from the bigger picture showing that, in one way or another, learning about orthography seems to permanently alter the way we perceive spoken language.

REFERENCES

- ABRAMSON, M., & GOLDINGER, S. D. (1997). What the reader's eye tells the mind's ear: Silent reading activates inner speech. *Perception & Psychophysics*, *59*, 1059-1068.
- BRAITENBERG, V. (1989). Some arguments for a theory of cell assemblies in the cerebral cortex. In L. Nadel, L. A. Cooper, P. Culicover, & R. M. Harnish (Eds.), *Neural connections, mental computation* (pp. 137-145). Cambridge, MA: MIT Press.
- COHEN, J., MACWHINNEY, B., FLATT, M., & PROVOST, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers*, *25*, 257-271.
- CONTENT, A., MOUSTY, P., & RADEAU, M. (1990). BRULEX: Une base de données lexicales informatisée pour le français écrit et parlé. [BRULEX: A computerized lexical data base for the French language]. *L'Année Psychologique*, *90*, 551-566.
- DAMIAN, M. F., & BOWERS, J. S. (2003). Effects of orthography on speech production in a form-preparation paradigm. *Journal of Memory & Language*, *49*, 119-132.
- DIJKSTRA, T., ROELOFS, A., & FIEUWS, S. (1995). Orthographic effects on phoneme monitoring. *Canadian Journal of Experimental Psychology*, *49*, 264-271.
- EHRI, L. C., & WILCE, L. S. (1980). The influence of orthography on readers' conceptualization of the phonemic structure of words. *Applied Psycholinguistics*, *1*, 371-385.
- FORSTER, K. I., & FORSTER, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116-124.
- FRAUENFELDER, U. H., SEGUI, J., & DIJKSTRA, T. (1990). Lexical effects in phonemic processing: Facilitatory or inhibitory? *Journal of Experimental Psychology: Human Perception & Performance*, *16*, 77-91.
- FROST, R., & KATZ, L. (1989). Orthographic depth and the interaction of visual and auditory processing in word recognition. *Memory & Cognition*, *17*, 302-310.
- FROST, R., REPP, B. H., & KATZ, L. (1988). Can speech perception be influenced by simultaneous presentation of print? *Journal of Memory & Language*, *27*, 741-755.
- GOLDINGER, S. D. (1996). Auditory lexical decision. *Language & Cognitive Processes*, *11*, 559-567.
- GOSWAMI, U. (1999). Causal connections in beginning reading: The importance of rhyme. *Journal of Research in Reading*, *22*, 217-240.
- GOSWAMI, U. (2002a). In the beginning was the rhyme? A reflection on Hulme, Hatcher, Nation, Brown, Adams, and Stuart (2002). *Journal of Experimental Child Psychology*, *82*, 47-57.
- GOSWAMI, U. (2002b). Phonology, reading development and dyslexia: A cross-linguistic perspective. *Annals of Dyslexia*, *52*, 141-163.
- HALLE, P. A., CHEREAU, C., & SEGUI, J. (2000). Where is the /b/ in "absurde" [apsyrd]? It is in French listeners' minds. *Journal of Memory & Language*, *43*, 618-639.
- HEBB, D. O. (1988). The first stage of perception: Growth of the assembly. In J. A. Anderson & E. Rosenfeld (Eds.), *Neurocomputing: Foundations of research* (pp. 45-56). Cambridge, MA: MIT Press.
- JAKIMIK, J. A., COLE, R. A., & RUDNICKY, A. I. (1985). Sound and spelling in spoken word recognition. *Journal of Verbal Learning & Verbal Behavior*, *24*, 165-178.
- JARED, D., MCRAE, K., & SEIDENBERG, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory & Language*, *29*, 687-715.
- KESSLER, B., TREIMAN, R., & MULLENNIX, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory & Language*, *47*, 145-171.
- KIRTLEY, C., BRYANT, P., MACLEAN, M., & BRADLEY, L. (1989). Rhyme, rime, and the onset of reading. *Journal of Experimental Child Psychology*, *48*, 224-245.
- LIBERMAN, A. M. (1992). The relation of speech to reading and writing. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning: Advances in psychology* (Vol. 94, pp. 167-178): Oxford: Elsevier, North-Holland.
- LIU, H., BATES, E., POWELL, T., & WULFECK, B. (1997). Single-word shadowing and the study of lexical access. *Applied Psycholinguistics*, *18*, 157-180.
- LUCE, P. A., GOLDINGER, S. D., & VITEVITCH, M. S. (2000). It's good . . . but is it ART? *Behavioral & Brain Sciences*, *23*, 336.
- MCRAE, K., JARED, D., & SEIDENBERG, M. S. (1990). On the roles of frequency and lexical access in word naming. *Journal of Memory & Language*, *29*, 43-65.
- METSALA, J. L. (1997). An examination of word frequency and neighborhood density in the development of spoken-word recognition. *Memory & Cognition*, *25*, 47-56.
- METSALA, J. L., & WALLEY, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89-120). Mahwah, NJ: Erlbaum.
- MORAIS, J., CARY, L., ALEGRIA, J., & BERTELSON, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, *7*, 323-331.
- NEW, B., PALLIER, C., FERRAND, L., & MATOS, R. (2001). Une base de données lexicales du français contemporain sur l'internet: LEXIQUE. [A lexical database of contemporary French on the Internet: LEXIQUE]. *L'Année Psychologique*, *101*, 447-462.
- NORRIS, D., MCQUEEN, J. M., & CUTLER, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral & Brain Sciences*, *23*, 299-325.
- OLSON, D. R. (1996). Towards a psychology of literacy: On the relations between speech and writing. *Cognition*, *60*, 83-104.
- PEEREMAN, R., CONTENT, A., & BONIN, P. (1998). Is perception a two-way street? The case of feedback consistency in visual word recognition. *Journal of Memory & Language*, *39*, 151-174.
- RAAIJMAKERS, J. G. W., SCHRIJNEMAKERS, J. M. C., & GREMMEN, F. (1999). How to deal with "The language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory & Language*, *41*, 416-426.
- RADEAU, M., & MORAIS, J. (1990). The uniqueness point effect in the shadowing of spoken words. *Speech Communication*, *9*, 155-164.
- RATTLE, K., & DAVIS, M. H. (2002). On the complexities of measuring naming. *Journal of Experimental Psychology: Human Perception & Performance*, *28*, 307-314.
- RUECKL, J. G. (2002). The dynamics of visual word recognition. *Ecological Psychology*, *14*, 5-19.
- SEIDENBERG, M. S., & TANENHAUS, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning & Memory*, *5*, 546-554.
- SEIDENBERG, M. S., WATERS, G. S., BARNES, M. A., & TANENHAUS, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning & Verbal Behavior*, *23*, 383-404.
- SLOWIACZEK, L. M., SOLTANO, E. G., WIETING, S. J., & BISHOP, K. L. (2003). An investigation of phonology and orthography in spoken-word recognition. *Quarterly Journal of Experimental Psychology*, *56A*, 233-262.
- STONE, G. O., VANHOY, M., & VAN ORDEN, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory & Language*, *36*, 337-359.

- STONE, G. O., & VAN ORDEN, G. C. (1994). Building a resonance framework for word recognition using design and system principles. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 1248-1268.
- TAFT, M., & HAMBLY, G. (1985). The influence of orthography on phonological representations in the lexicon. *Journal of Memory & Language*, **24**, 320-335.
- TANENHAUS, M. K., FLANIGAN, H. P., & SEIDENBERG, M. S. (1980). Orthographic and phonological activation in auditory and visual word recognition. *Memory & Cognition*, **8**, 513-520.
- TARABAN, R., & McCLELLAND, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory & Language*, **26**, 608-631.
- VAN ORDEN, G. C. (2002). Nonlinear dynamics and psycholinguistics. *Ecological Psychology*, **14**, 1-4.
- VAN ORDEN, G. C., & GOLDINGER, S. D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 1269-1291.
- VENTURA, P., KOLINSKY, R., BRITO-MENDES, C., & MORAIS, J. (2001). Mental representations of the syllable internal structure are influenced by orthography. *Language & Cognitive Processes*, **16**, 393-418.
- VENTURA, P., MORAIS, J., PATTAMADILOK, C., & KOLINSKY, R. (2004). The locus of the orthographic consistency effect in auditory word recognition. *Language & Cognitive Processes*, **19**, 57-95.
- VITEVITCH, M. S. (2002). Influence of onset density on spoken-word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 270-278.
- VITEVITCH, M. S., & LUCE, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, **9**, 325-329.
- VITEVITCH, M. S., & LUCE, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory & Language*, **40**, 374-408.
- ZIEGLER, J. C., & FERRAND, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, **5**, 683-689.
- ZIEGLER, J. C., JACOBS, A. M., & STONE, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments, & Computers*, **28**, 504-515.
- ZIEGLER, J. C., MUNEAUX, M., & GRAINGER, J. (2003). Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory & Language*, **48**, 779-793.

APPENDIX

Inconsistent Words With Identical Rimes (Dominant Spelling/Subdominant Spelling)

cran/paon, glande/stand, faille/bail, mangue/gang,
 claque/vrac, crampe/trempe, pente/plante, vase/gaz,
 creux/noeud, joute/scout, frais/jet, nain/daim, hausse/sauce,
 fève/glaive, fraise/thèse, cru/flux, blond/plomb, lobe/snob,
 tôle/hall, gomme/rhum, mou/loup, four/bourg, dru/jus,
 fruit/pluie

Consistent Words

moche, stage, figue, louche, crabe, match, plume, huile, cave,
 luge, brune, jade, fiche, gaffe, ruche, singe, blague, biche,
 cube, grade, plage, tige, digue, lune

(Manuscript received June 5, 2003;
 revision accepted for publication January 6, 2004.)