Article

Processing of illegal consonant clusters: A case of perceptual assimilation?

HALLÉ, Pierre A., *et al*.

Abstract

Evidence is presented for a perceptual shift affecting consonant clusters that are phonotactically illegal, albeit pronounceable, in French. They are perceived as phonetically close legal clusters. Specifically, word-initial /dl/ and /tl/ are heard as /gl/ and /kl/, respectively. In 2 phonemic gating experiments, participants generally judged short gates—which did not yet contain information about the 2nd consonant tl—as being dental stops. However, as information for the /!/ became available in larger gates, a perceptual shift developed in which the initial stops were increasingly judged to be velars. A final phoneme monitoring test suggested that this kind of shift took place on-line during speech processing and with some extratemporal processing cost. These results provide evidence for the automatic integration of low-level phonetic information into a more abstract code determined by the native phonological system.

Reference

HALLÉ, Pierre A., *et al*. Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance*, 1998, vol. 24, no. 2, p. 592-608

DOI : 10.1037/0096-1523.24.2.592
PMID : 9554098

Journal of Experimental Psychology:
Human Perception and Performance
1998, Vol. 24, No. 2, 592–608

0096-1523/98/$3.00

Processing of Illegal Consonant Clusters: A Case of Perceptual Assimilation?

Pierre A. Hallé and Juan Segui
Centre National de la Recherche Scientifique,
Paris V University

Uli Frauenfelder and Christine Meunier
University of Geneva

Evidence is presented for a perceptual shift affecting consonant clusters that are phonotactically illegal, albeit pronounceable, in French. They are perceived as phonetically close legal clusters. Specifically, word-initial /dl/ and /tl/ are heard as /gl/ and /kl/, respectively. In 2 phonemic gating experiments, participants generally judged short gates—which did not yet contain information about the 2nd consonant /l/—as being dental stops. However, as information for the /l/ became available in larger gates, a perceptual shift developed in which the initial stops were increasingly judged to be velars. A final phoneme monitoring test suggested that this kind of shift took place on-line during speech processing and with some extratemporal processing cost. These results provide evidence for the automatic integration of low-level phonetic information into a more abstract code determined by the native phonological system.

The view that speech perception is determined by the native-language sound system is well motivated and widely shared. Ontogenetically, there is a shift from universal to language-specific perceptual capacities: Although young infants seem to be initially equipped with "universal" capacities for processing speech sounds, language-specific capacities emerge in the second half of the first year of life, by 9 months or before (Jusczyk, Cutler, & Redanz, 1993; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993; Jusczyk, Luce, & Charles-Luce, 1994; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Polka & Werker, 1994). By 10–12 months, children lose the ability to discriminate some of the nonnative consonant contrasts that are not functional in the language they are learning, at least when the consonants involved can be heard as (or assimilated to) native speech sounds (Best, 1994b; Best, McRoberts, Lafleur, & Silver-Isenstadt, 1996; Best, McRoberts, & Sithole, 1988; Werker & Tees, 1984). By the end of the first year, then, a language-specific speech processing system is well on its way toward stabilization.

The language specificity of the resulting adult system has been characterized in terms of the "phonological filter" metaphor (see Trubetzkoy, 1939, p. 57). The native language sound system acts as a phonological filter molding some nonnative phonemes into native phonemic categories provided that there is sufficient similarity. In other words, nonnative phones may assimilate to native phonemes[1] (cf. Best's, 1994b, perceptual assimilation model; also see Flege's, 1986, 1991, speech learning model). Moreover, the attuning to the native-language system also results in language-specific ways of processing speech. For example, different native languages seem to foster different segmentation procedures and, perhaps, the use of different prelexical units of representation (Cutler, Mehler, Norris, & Segui, 1986, 1992). These findings can be explained by assuming that listeners are biased toward interpreting any spoken utterance as being consistent with the organization of their native language and with its functional linguistic representations.

Most studies dedicated to the general issue of native-language bias have taken a cross-linguistic and developmen-

Pierre A. Hallé and Juan Segui, Laboratoire de Psychologie Expérimentale, Centre National de la Recherche Scientifique (CNRS), Paris V University, Paris, France; Uli Frauenfelder and Christine Meunier, Laboratoire de Psycholinguistique Expérimentale, University of Geneva, Geneva, Switzerland.

We gratefully acknowledge support from a grant from Axe de Recherche CNRS "Langage" (1994–1996) and from a Fonds National pour la Recherche Scientifique Grant 11-39553.93. We thank Catherine T. Best for her helpful comments on an earlier version of this article. We also are grateful to Catherine T. Best and Laurent Sagart, who provided phonetic and phonological data, and to Denis Lancelin, who helped in running the participants in Paris.

Correspondence concerning this article should be addressed to Pierre A. Hallé, CNRS-Paris V, Laboratoire de Psychologie Expérimentale, 28 rue Serpente, Paris, France. Electronic mail may be sent to hal@idf.ext.jussieu.fr.

[1] Certain nonnative single phonemes with a complex articulation can be assimilated to clusters of native phonemes. This is the case of the voiced versus unvoiced lateral fricatives of Zulu /ɮ/ versus /ɬ/, described by adult American listeners as sounding like the following clusters: /ʒl/ or /ðl/ versus /ʃl/ or /θl/ (Best, 1990, 1991, 1994a). In addition to the Zulu lateral fricatives, labio-velars such as [gb] or [gp], or prenasalized consonants such as [ᵐb], may be perceived as /g/ + /b/ or /g/ + /p/, or /m/ + /b/ clusters. In fact, whether certain speech sounds are clusters or singleton phonemes may sometimes be arbitrary. For example, a /dl/ : /tl/ contrast has been posited in Tlingit (a Native American language) and described as a voicing contrast of lateral affricates (Ladefoged & Maddieson, 1996, p. 207; Maddieson, 1984, p. 76). In the International Phonetic Alphabet usage, the digram "dl" denotes the same phoneme (or phonemes) as the single letter λ (*lambda*) in American usage for a "voiced alveolar laterally released affricate" (Pullum & Ladusaw, 1986, p. 94).

tal approach. At the same time, they have focused largely on the perception of single-phoneme contrasts. Although several cases of cross-linguistic phoneme assimilation have been reported, little attention has been paid to sequences of phonemes, particularly to consonant clusters. The notion of a phonological filter, however, should apply to sequences of phonemes as well as to singleton phonemes. In 1931, Polivanov already claimed that the bias to perceive foreign sounds according to one's native phonological system is not confined to singleton phonemes but that it extends to groups of phonemes. For example, Japanese listeners listening to the English word *drama* hear /do.ra.ma/ (Polivanov, 1931), consistent with the fact that Japanese does not allow syllable-initial clusters.[2] Native-language phonotactic constraints thus can bias the perception of nonnative clusters so that they fit into the framework of the native-language phonological system.

The effect of native phonotactic constraints in speech processing also can be studied within a given language. According to the phonological filter view, illegal but pronounceable clusters should tend to be assimilated to clusters that are legal, whenever there is sufficient similarity, just as nonnative singleton phones tend to be assimilated to native phonemes. Hence, examining the perception of illegal clusters within a given language is yet another approach to explore how native-language sound systems determine the perception of speech sounds.

The perception of illegal clusters has not received much attention. Brown and Hildum (1956) presented natural-speech monosyllables to native English speakers. The monosyllables had a complex onset and corresponded to (rare) English words or nonwords with either a permissible or nonpermissible cluster at their onset. Participants were required to write down what they heard. Their responses were strongly influenced by their expectation of English words and even more by their implicit knowledge of English phonotactic constraints. There was a strong bias to misperceive illegal clusters. This was observed for phonetically naive participants as well as for those trained in "linguistics" who were explicitly instructed to write phonemic transcriptions of non-English combinations of phonemes. Similar phonotactic knowledge and bias toward permissible clusters are also found in young children. Messer (1967) presented children learning English (aged 3 years 3 months to 4 years 3 months) with pairs of monosyllabic nonwords, one with an illegal (or infrequent) consonant cluster onset and the other with a legal onset. (In a few items, the onset was /ʒ/, which is illegal in English in that position.) The children were asked to say which item of each pair sounded "more like a word" by orally reproducing it. Legal-onset items were chosen by children more often than illegal-onset items. They mispronounced illegal items more often than legal ones, and the mispronunciations always turned illegal clusters into permissible clusters. As suggested by Messer (1967) these results indicate "a perceptual disposition . . . to hear sounds that are 'possible' " (p. 610). This is consistent with the notion of a phonological filter causing the perceptual assimilation of phones sequences as well as of single phones.

Another kind of evidence for perceptual assimilation of illegal clusters comes from Massaro and Cohen's (1983) study. Participants were presented with synthetic speech stimuli beginning with an obstruent + liquid cluster. The liquid belonged to a /r/ – /l/ continuum. The category boundary between /l/ and /r/ was shifted (relative to a neutral situation in which the initial obstruent was /p/), such that more /r/ phonemes than /l/ phonemes were identified after /t/ and, conversely, more /l/ phonemes than /r/ phonemes were identified after /s/. By this "phonological context effect" (in Massaro and Cohen's formulation), participants implicitly showed a bias to perceive legal clusters, namely /tr/ or /sl/, rather than */tl/ or */sr/. Participants also might have perceived the cluster-initial consonant differently as a function of the acoustic quality of the liquid. In Massaro and Cohen's (1983) Experiment 3, both the initial consonant and the liquid were ambiguous (a /b/ – /d/ continuum was combined with a /r/ – /l/ continuum). The results suggested a perceptual bias to report /bl/, /br/, and /dr/ rather than */dl/. For stimuli situated at the /l/ end point, there were fewer /dl/ judgments for the /d/ end point than /bl/ judgments for the /b/ end point. Thus, the perceptual shift also appeared to affect the perception of the initial consonant. The important point is that listeners showed a bias toward hearing legal clusters (i.e., clusters consistent with the organization of their native-language phonology).

In French as in English, /dl/ and /tl/ clusters are not permissible in syllable-initial position, although they are pronounceable. How are they perceived? One possibility is that these clusters are perceived as non-French combinations of phonemes, with the constituent phonemes correctly identified. That is, /dl/ would sound foreign but would nevertheless be perceived as the combination of /d/ and /l/. No perceptual assimilation would occur at the structural level of syllable onset. This seems to be the case for some illegal Dutch clusters that were used in an experiment by Praamstra and Levelt (1994). In a lexical-decision task, nonwords with illegal clusters in initial position led to faster "no" responses than did phonotactically regular nonwords. This difference is best understood by assuming that the illegal clusters were readily perceived as impossible sequences. In contrast to the legal cluster nonwords, further speech input was not necessary to make a lexical decision. Presumably, then, these illegal clusters were not misperceived. A second possibility, suggested by the production data in the Messer (1967) study and by the Massaro and Cohen (1983) study, is that certain illegal clusters are perceived as legal ones. This is an instance of structural perceptual assimilation: Individually, the phonemes /d/ and /l/ would not be misperceived, but their combination gives rise to the perception of a different sequence from the intended /d/ + /l/ (e.g., /d/ + /r/).

One reason why one might expect the perceptual assimila-

[2] This finding has been recently reassessed (Dupoux et al., 1996). In nonsense words such as /ebuzo/ (uttered by a Japanese speaker: [ebɯzɔ]), Japanese participants still hear an epenthetic vowel /u/ when the vowel and its acoustic correlates have been removed.

tion of /dl/ and /tl/ to legal clusters is that the phonotactic constraint barring them from French is not motivated by a distributional regularity, such as the constraint that excludes nasal + plosive clusters. Rather, in French, /dl/ and /tl/ represent an accidental gap within the set of obstruent + liquid (OBLI) clusters (following Dell's, 1995, terminology) and may conceivably be misperceived as some legal OBLI cluster. By contrast, the onset cluster /nt/, for example, is structurally impermissible in that there is no legal cluster sharing the same structure to which it could assimilate.

Now, will /dl/ be perceived as /dr/, as Massaro and Cohen's (1983) results might suggest, or as some other cluster? We do not yet know whether one phoneme of the cluster is perceptually more stable than the other and, if so, which phoneme is most likely to be misperceived. In French, /l/ and /r/ have much different phonetic values: /l/ is a "clear" alveolar [l], whereas /r/ is an uvular [ʁ], often produced with some degree of frication. Thus, the /l/ in /dl/ or /tl/ is not likely to be perceived as (French) /r/. Instead, one would expect a shift in the perception of the initial consonant.

The first experiment was designed primarily to determine how /dl/ and /tl/ clusters in item-initial position would be perceived by French listeners: (a) literally, as /d/ + /l/ or as /t/ + /l/, or (b) assimilated to a permissible sequence. If /tl/ and /dl/ induce onset-level perceptual assimilation, the question arises as to which sequence of phonemes they will assimilate. This issue is addressed in the following open response test. If the results indicate that a perceptual shift does indeed occur, a second issue naturally arises as to the level of processing at which the assimilation occurs. The possible misperception of illegal clusters may result from a conscious reanalysis of the constituents that are perceived or from an automatic recomputation. In both cases, some additional perceptual difficulty—relative to the straightforward perception of permissible clusters—can be expected. These issues were addressed in gating experiments and in an on-line phoneme detection experiment.

Experiment 1: Open Responses to /dl/- or /tl/-Initial Nonwords

Initially, we wanted to discover what French listeners hear when they are presented with items beginning with the illegal cluster /dl/ or /tl/ pronounced by a French speaker. More specifically, if there is a systematic bias to hear another sequence than the one produced by the speaker, would it affect the perception of the initial dental plosive consonant or of the following liquid? A simple and straightforward approach to answering this question would be to let participants transcribe freely what they hear when presented with the clusters under scrutiny.

Method

Stimuli. Four nonwords with an initial /tl/ ("tlabdo," "tlabod," "tlobad," and "tlobda") and four with its voiced counterpart /dl/ ("dlapto," "dlapot," "dlopat," and "dlopta") were used as test stimuli. In addition, 24 phonotactically legal nonword filler items also were prepared; 2 with a /br/ and /gr/ cluster in initial position and the others with a single initial consonant (six were filler items with a /pl/ or a /sl/ cluster in medial position). Disyllables rather than monosyllables were chosen because, in French, most phonotactically possible monosyllables are real words or are similar in form to real words. The use of monosyllables could thus induce unwanted lexical effects. The material was recorded by a male native speaker of French on a digital audiotape tape recorder, then digitized (10-kHz, 16-bit resolution) and stored in computer files.[3] The spectrograms of the test stimuli were inspected to ensure that the onsets were not produced with an epenthetic schwa after the item-initial dental plosive, turning the illegal "clusters" into legal sequences /dəl/ or /təl/. No trace of a vocalic portion could be detected.

Participants. Eighteen French participants who were native speakers of Parisian French participated voluntarily in this experiment. They were aged 21–30 years. All of them were phonetically naive, and none reported any hearing deficit.

Procedure. The 32 nonwords were presented to the participants in random order with an interstimulus interval (ISI) of 6 s. Participants were tested individually in a sound-attenuated booth and received the speech material through Sennheiser headphones. After receiving each item, participants wrote down what they heard using a free style of transcription. These naive transcriptions were then inspected and recoded using a standard phonological transcription for French. When necessary, participants were asked to clarify their responses.

Results and Discussion

The outcome of this experiment was clear-cut: The /dl/ and /tl/ items were transcribed with a /gl/ or /kl/ cluster 85.4% of the time (range = 72.2–94.4, $SD = 7.2$), with a /dl/ or /tl/ cluster only 13.2% of the time and with a /pl/ cluster 1.4% of the time (see Table 1).[4] By contrast, the other consonantal word onsets, either singletons or clusters, were transcribed correctly 100% of the time, as were the word-medial clusters /sl/ and /pl/, except on a few occasions (6%) when participants inserted an "e" in the cluster (e.g., "jocelaire" for [ʒɔslɛʁ]). Clearly, participants transcribed the /dl/ and /tl/ clusters more often as a combination of letters corresponding to /gl/ or /kl/ than as "dl" or "tl"—by subject, $t(17) = 6.65$; by item, $t(7) = 14.3$, $ps < .0001$—and never as "dr" or "tr." This supports the prediction that the perception of the /l/ segment in the cluster is more stable than that of the initial plosive consonant. Does this outcome reflect unwanted lexical effects? None of the reported forms was close to a word. Moreover, in the context of –/a/ or –/o/ as well as other vowels, /gl/ and /kl/ are not more

[3] The talker was not a trained phonetician but was judged to have a clear pronunciation in general. He practiced the lists of nonwords once before recording.

[4] Dental-to-labial shifts were thus rare. This was not likely attributable to lexical effects: Both /pl/ and /kl/ are permissible word-initially, and /kl/ words are not more frequent than /pl/ words (see Footnote 5). A plausible reason why dental-to-velar rather than dental-to-labial shifts were observed is the closer acoustic similarity of dentals to velars than to labials, both in burst spectrum properties and in consonant–vowel formant transitions (see Figure 1 for an illustration). Also, from an articulatory point of view, both dentals and velars are linguals, whereas labials stand out as labials.

Table 1
Transcriptions of the /dl/ and /tl/ Items in Experiment 1

Item	% of transcriptions as /tl/	/dl/	/kl/	/gl/	/pl/
/tl/	9.7	1.4	81.9	4.2	2.8
/dl/	0	15.3	4.2	80.6	0

frequent word-initial clusters than /dr/, /tr/, /bl/, or /pl/.[5] The results thus suggest that /dl/ and /tl/ assimilate, respectively, to /gl/ and /kl/, both legal clusters.

The high percentages of dental-to-velar "confusions" might have been caused by phenomena other than perceptual assimilation per se. First, the /dl/ and /tl/ test items may have had an intrinsic velar quality (i.e., velar rather than dental cues to place of articulation might have been present in the speech signal). This is possibly due to some difficulty that native speakers of French have in clearly articulating /dl/ or /tl/ initial clusters. It also is possible that a velar quality unavoidably arises from the coarticulation of a dental plosive and a French /l/. We attempted to check that possibility by means of an acoustic analysis.

The spectral properties of the release burst have been found to vary as a function of the stop consonant's place of articulation, at least to a certain extent (Halle, Hughes, & Radley, 1957; Kewley-Port, 1983; Stevens & Blumstein, 1978; also see Lahiri, Gewirth, & Blumstein, 1984). For dentals, energy is usually distributed at higher frequencies than for velars or labials. We therefore compared the spectra computed at the release burst for the /dl/ and /tl/ items and for matched /gl/ and /kl/ items with identical segmental context uttered by the same speaker (e.g., "dlopat" vs. "glopat"). We then used the spectral center of gravity as a simple and objective measure to characterize these spectra.[6] The spectral center of gravity should lie at higher frequencies for dental than for velars in the same context. This was found for the dental and velar items under scrutiny: 2.62 versus 1.64 kHz, $t(7) = 7.64$, $p < .0002$. There was no overlap, except for the item *dlopta,* whose burst spectral center of gravity (1.79 kHz) was the lowest. Across the dental + /l/ items, the measured spectral center of gravity did not (negatively) correlate with the percentage of velar responses. Therefore, the measurements show that /dl/ and /tl/ items were more dental (in the sense of more fronted) than comparable /gl/ and /kl/ items. Figure 1 illustrates that difference and also shows spectra from a /pl/ item, which, as expected, exhibited a burst spectrum centered on lower frequencies. However, is this demonstration conclusive? We could just as well say that /dl/ and /tl/ items were less velar than /gl/ and /kl/ items with respect to the release burst spectrum. Moreover, perhaps more subtle cues than spectral cues signaled velar articulation. Therefore, the discussion is not definitely closed. However, the possibility of gross and extensive mispronunciations of the dental + /l/ items as velar + /l/ can be rejected. We return to the issue of dental versus velar place of articulation in Experiment 3.

There is another alternative to the perceptual assimilation account. Even if the cluster constituents were correctly perceived, listeners had time to (consciously) reconsider what they had heard and make transcriptions that were compatible with their implicit knowledge of phonotactic constraints. Such a tendency might have been reinforced in this task of transcribing the input by a bias toward producing correct spellings. Many participants were perhaps reluctant to write down a combination of letters that is not permissible in word-initial position. The overwhelming dental-to-velar shift that was obtained may thus be partly due not only to a tendency to report permissible phonemic sequences but also to participants' knowledge of permissible graphemes in word-initial position. Indeed, participants' responses often reflected the idiosyncrasies of the orthography of French words rather than being regular phonemic transcriptions (e.g., /vobral/ was often transcribed "veaubrale"). This observation points to an orthographic bias but is difficult to quantify. Although about 15% of the items presented were transcribed with the anomalous orthography "dl" or "tl," the velar responses (about 85%) were mostly due to participants who consistently avoided "dl" and "tl": There was a majority of "conservative" participants (12 of 18) whose responses always respected regular orthography, but only a few less conservative participants who produced "dl" and "tl" responses. This pattern of results is consistent with an orthographic bias explanation.

The next experiment was designed to test this interpretation further. We used a forced-choice identification task to make the participants focus their attention on the item-initial consonant. Because this task did not require listeners to write down letters that could possibly form illegal sequences, the chances that orthographic knowledge would affect the responses were reduced. If a massive dental-to-velar shift is still obtained, it reflects either the intrinsic velar quality of the item-initial plosive in the test stimuli (which was not found in the acoustic analysis) or a perceptual shift at the level of syllable onset.

[5] Before the vowels /a/ and /o/, the numbers of word types with these plosive + liquid clusters in word-initial position are 227 (velar + /l/), 166 (labial + /l/), and 242 (dental + /r/); the corresponding number of tokens are 27,066, 36,511, and 123,907, respectively (source: "Trésor de la Langue Française" [Imbs, 1971]). Hence, with regard to lexical frequencies, /gl/ and /kl/ were the least likely clusters that should be reported for the /dl/ and /tl/ items.

[6] Short-term spectra were computed for 10-ms waveform portions beginning 2 ms before the burst. In the case of /dl/ and /gl/, where the plosive was prevoiced, the voicing murmur was partly filtered out by high-pass filtering with a 500-Hz cutoff frequency. (This was done to make the measurements for prevoiced and voiceless plosives comparable.) Spectra were smoothed by means of an iterative cepstral smoothing technique, with a 2-ms cutoff quefrency. Spectral centers of gravity were computed from the smoothed spectra.

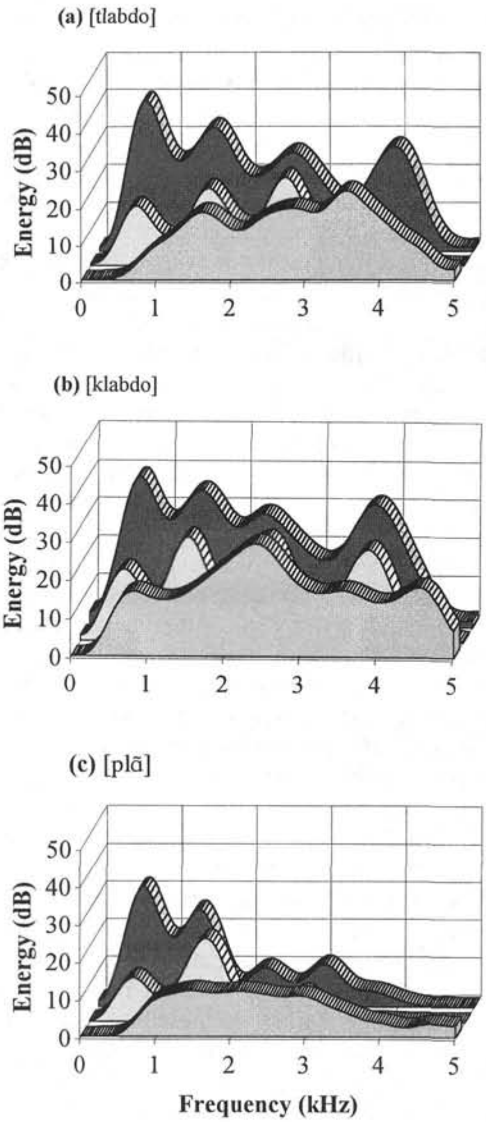

Figure 1. Short-term spectra of the initial stop release burst (medium gray), the following /l/ (light gray), and the onset of the following vowel (dark gray) in "tlabdo" (A), "klabdo" (B), and [plɑ̃] (from a filler item; C).

Experiment 2: Forced-Choice Identification of Cluster-Initial Consonants

In this experiment, listeners identified item-initial consonants in a forced-choice procedure. Two tests were conducted. In the first, participants had to choose from among three consonants with three places of articulation (i.e., dental, velar, and labial) but with voicing held constant. In the second, only dental and velar places of articulation were proposed, but either voiced or not, so that voicing confusions could be tested. With a forced-choice response, it is unlikely that orthography would interfere with the response. Moreover, because the participants' attention was focused on the initial consonant, the tendency to produce phonologically permissible responses was presumably reduced.

Method

Stimuli. The 8 /dl/ or /tl/ test items used in the previous experiment were matched with similar nonwords with legal initial clusters (/dr/ or /tr/, /gl/ or /kl/, and /gr/ or /kr/) in items such as "drapot," "trabod," "glapot," and so on. These served as control items for which few or no perceptual confusions were expected. In the first test (threefold forced choice), participants received two sets of 36 items. All the items were disyllabic nonwords. In one set, the initial consonants were voiced (/b/ – /d/ – /g/ forced choice); in the other set, they were voiceless (/p/ – /t/ – /k/ forced choice). Each set contained 4 test items (e.g., the /dl/ items), 12 control items (e.g., the /dr/, /gl/, and /gr/ items), and 20 filler items with a dental, velar, or labial word-initial stop consonant so that the three places of articulation were equiprobable. In the second test (fourfold forced choice), participants received a single set of 32 items containing the 8 /dl/ or /tl/ test items and the 8 corresponding /gl/ or /kl/ control items plus 16 filler items whose initial consonant was equiprobably /d/, /t/, /g/, or /k/. The speaker who recorded the material of the first experiment again recorded the new items using the same apparatus as before.

Participants. Twenty-four students at Université Paris V participated for course credit in the first test; 16 students participated in the second test. All were native speakers of Parisian French with no reported hearing deficit.

Procedure. The participants were tested individually in a sound-attenuated booth. Stimuli were presented through Sennheiser headphones at a comfortable level, with a 3-s ISI. Participants were instructed to circle one of the three or four letters proposed for each stimulus (first test, *P, T, K* or *B, D, G*; second test, *D, T, G, K*). In the first test, participants were assigned one of two orders of presentation.

Results and Discussion

Results can be described in terms of phonetic "confusions." Only place-of-articulation confusions could occur in the first test; in the second test, voicing confusions also could occur.

In the first test, no confusions occurred for the velar-initial control items. For this test, the description of the results is thus limited to the confusions that occurred for dental-initial test items (/dl/ or /tl/ initial cluster) and control items (/dr/ or /tr/ initial cluster). Table 2 shows the results obtained in this first test: Dental-to-labial confusions were marginal, and dental-to-velar confusions occurred mainly for the test items. The percentages of confusions were analyzed according to the order of presentation, the initial consonant (/t/ or

Table 2
Identification Test 1: Percentage of Dental-to-Velar and Dental-to-Labial Confusions According to Stimulus Type

Item-initial cluster	% of place confusions: Dental to velar	Dental to labial
Illegal		
/tl/	63.5	1.0
/dl/	47.9	1.0
Legal		
/tr/	8.3	1.0
/dr/	1.0	6.3

/d/), and the cluster type (illegal, /tl/ and /dl/; legal, /tr/ and /dr/). Both the dental-to-velar and the dental-to-labial confusions (treated as a repeated measure) were examined. Order had no significant effect, $F_1(1, 22) = 2.9$, $F_2(1, 24) = 2.8$, $ps > .1$. Dental-to-labial confusions were more rare than dental-to-velar confusions, $F_1(1, 22) = 54.4$, $F_2(1, 24) = 72.0$, $ps < .0001$. Many more dental-to-velar confusions occurred for illegal than for legal clusters (55.7% vs. 4.2% overall), $F_1(1, 22) = 75.1$, $F_2(1, 24) = 72.4$, $ps < .0001$. The percentage of these confusions was higher for /tl/ (63.5%) than for /dl/ (47.9%), but the difference was marginally significant only in the by-subject analysis, $F_1(1, 22) = 4.0$, $p = .055$.

Figure 2 shows the results obtained in the second test. In this test, both place and voicing confusions could occur. The /gl/- and /kl/-initial items served as controls. The pattern of dental-to-velar confusions for /tl/ and /dl/ items was highly similar to that observed in the first test: 50% /d/ → /g/ confusions, 71.9% /t/ → /k/ confusions. Place and voicing confusions occurred for the /tl/ items (12.5% /t/ → /g/ confusions), but not for /d/ → /k/. Place confusions (dental-to-velar for test items or velar-to-dental for control items) were examined according to the cluster type (illegal or legal) and the initial consonant "voicedness"; place confusions with and without voicing confusion were counted separately (treated as a repeated measure). Almost no place confusion occurred for control as compared to test items, $F_1(1, 15) = 65.5$, $F_2(1, 12) = 170.7$, $ps < .0001$. There were more place confusions for /tl/ than for /dl/ items when voicing confusions were ignored, $F_1(1, 15) = 15.3$, $p = .0014$, $F_2(1, 6) = 11.7$, $p = .014$. When limiting the comparison to /t/ → /k/ versus /d/ → /g/ confusions, the difference between /tl/ and /dl/ items became nonsignificant in the by-subject analysis, $F_1(1, 15) = 2.9$, $p = .11$, but it was still marginally significant in the by-item analysis, $F_2(1, 6) = 4.2$, $p = .08$.

This experiment had a more constraining design than the preceding one in order to focus participants' attention on item-initial consonants rather than on whole words and to reduce the possible effect of orthographic knowledge. The results obtained show a smaller effect than in the first experiment, but there was still a substantial dental-to-velar shift. This result suggests that some kind of perceptual assimilation occurs for /dl/- and /tl/-initial items. This identification experiment probably provides a more realistic picture of the phenomenon than the open-response experiment, in which extraperceptual factors such as orthographic knowledge conceivably exaggerated the size of the effect.

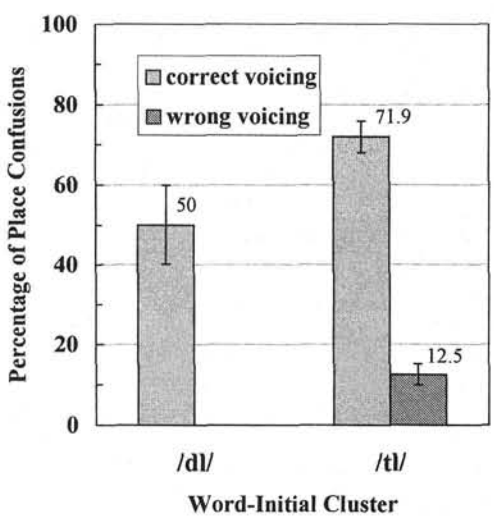

Figure 2. Dental-to-velar confusions and voicing confusions for /tl/ and /dl/ test items: fourfold forced-choice identification test.

We now return to the question of the acoustic quality of the cluster-initial plosives in the illegal clusters used. The acoustic measures already exclude gross and extensive mispronunciations of the intended dentals as velars. As pointed out earlier, however, subtle cues to velar articulation may have not been revealed by these measures. In the next experiment, we examined how the cluster-initial plosive consonants would be perceived on the basis of partial acoustic information using a gating task.

Experiment 3: Gating Task With Legal Versus Illegal Onset Clusters

The acoustic measures showed that the /dl/ and /tl/ stimuli were more fronted than the matched /gl/ and /kl/ stimuli. Altogether, this makes it less likely that velar cues were present in /dl/ and /tl/. However, we needed to obtain a more precise idea of how the cluster-initial plosives would be perceived in the absence of the following liquid context. Blumstein and Stevens (1980) have shown that brief stimuli, as short as 10–20 ms, sampled from the release burst of (synthetic) consonant–vowel (CV) syllables can be reliably identified for consonant place of articulation. Hence, participants might be able to correctly identify the initial portion of the /dl/ and /tl/ clusters for place of articulation when they are spliced out of the subsequent context.

The forward gating technique can be used to examine how brief portions of the stimuli sampled from the release burst are perceived. Moreover, if a perceptual shift does occur and develops as longer fragments are presented, the gating technique can reveal the "time course," as it were, of such a shift. Because we were investigating the phonetic perception of the illegal clusters /dl/ and /tl/, we had to use a variant of the gating paradigm in which listeners transcribed what they heard in lieu of guessing a word (Grosjean, 1980; Warren & Marslen-Wilson, 1987). Indeed, if listeners were instructed to guess a word, they could not possibly report /dl/- or /tl/-initial words, and the test would not tap into the perceptual shift at a sublexical level of processing. We therefore introduce here a novel "transcriptional" variant of the gating paradigm that we call *phonemic gating*.

How do listeners interpret the place cues as increasingly longer fragments of the /dl/- and /tl/-initial items are presented? Suppose that listeners report velar rather than dental sounds as soon as a stable perception of place has emerged. This would show that cues to velar place of articulation are actually present in the speech signal and explain the velar responses that were obtained in Experiments 1 and 2. By contrast, if listeners systematically report dental rather than velar sounds, at least for short fragments, we can reject the "velar quality" explanation. However, listeners may not consistently produce dental responses. For this reason, we needed to compare the illegal clusters /dl/ and /tl/ with matched legal clusters containing either dental- or velar-initial plosive consonants: /dr/ and /tr/ or /gl/ and /kl/. Dental-initial legal clusters provided a baseline for how dental place of articulation was normally perceived, whereas velar-initial clusters gave a baseline for how velar place was normally perceived.

In Experiment 3 we used a successive format of presentation: In a given sequence, the same speech item (a nonword) was presented incrementally. We used two tests with different participants: One group of participants listened to /dl/ and /tl/ items and the other group to matched /gl/, /kl/, /dr/, and /tr/ items. This design was used to avoid possible influences, assimilatory or contrastive, of one set of items on the other.

Method

Stimuli. The experimental items in the first test consisted of the 8 /dl/- or /tl/-onset nonwords used in Experiment 1 and of 8 filler items, also nonwords. The second test used the same 8 filler items and 8 test items with /gl/, /kl/, /dr/, or /tr/ onsets, matched with those of the first test. Half of the possible matched items were used for each onset in order to limit the length of the experiment so as not to go beyond the attentional capacity of the subjects. These items were as follows: "glapot," "glopta," "klabdo," and "klobda" and "drapto," "dropat," "trabod," and "trobda." (Therefore, both vocalic contexts −/a/ and −/o/ were used for each context.) The speaker used for the previous experiments recorded the new items appearing in this experiment using the same apparatus.

The initial fragment, or gate, of each nonword contained the initial portion of the signal up to 10 ms after the stop release burst regardless of whether that stop was prevoiced. The other fragments were made increasingly longer using a 20-ms increment. The final 4 ms of each fragment was attenuated by a raised cosine function so that there were no perceivable "clicks" at the gate offset. The 10th and last fragment thus included 190 ms of signal after the release burst. This roughly corresponded to the first syllable in all items. Table 3 shows the average intervals from the release burst to the acoustic onsets of the liquid, the following vowel, and the following stop closure and release, for the three types of items: dental + /l/, velar + /l/, and dental + /r/. These intervals were measured on the basis of spectrograms.

Participants. The participants, all native speakers of Parisian French, were undergraduates aged 22–26 years; 28 students participated in the first test and 22 in the second test for course credit. The students were phonetically naive and had no hearing difficulties.

Procedure. For both tests, participants received a total of 160 stimuli (16 items × 10 gates) in a successive presentation format. For each item, 10 fragments of increasing size were presented, and participants had to transcribe exactly what they heard on an answer sheet (using a free style of transcription) and to give a confidence rating for their transcription on a 1–5 scale. Importantly, listeners were required to transcribe what they heard rather than to guess a word, as in the classical gating task. In case they could not report any speech sound at all, which often happened for the first gate, they just wrote down a question mark. Participants were tested individually. After each participant had completed the test, the experimenter checked the written responses and, whenever necessary, asked participants to clarify the phonetic value of any unclear transcriptions. In each test, the 16 items tested were presented in two different orders.

Table 3
Locations of the Main Acoustic Events in the Cluster-Initial Items Used in the Gating Experiments

Item type	Liquid	Vowel	Closure	Release
Dental + /l/	20	76	177	289
Velar + /l/	24	71	157	272
Dental + /r/	31	123	231	348

Note. Shows onsets of the liquid (/l/ or /r/) and of the following vowel (/a/ or /o/), closure and release of the following plosive (/b/ or /p/), as measured (in milliseconds) from the initial stop release burst.

Results and Discussion

Figure 3C shows the distribution of the participants' gating responses for /dl/ and /tl/ items as a function of the place of articulation. It can be seen that the percentage of dental responses was much higher than that of velar responses at each gate. The percentage of dental responses reached a maximum at the 4th gate (83%) and then gradually decreased to 67% at the last gate, whereas the velar responses symmetrically increased, reaching 32% at the last gate. The same pattern was observed for both orders of presentation.

By contrast, the identification of /gl/ and /kl/ and of /dr/ and /tr/ (see A and B in Figure 3) was increasingly accurate, reaching about 96% correct responses at the last gate. For /gl/ and /kl/, the percentage of velar responses reached about 86% at the 4th gate and then continued to increase. Intended velars were frequently judged as velars from the early gates on, and there was little trace of early dental judgments: The percentage of dental responses reached a maximum of 11% at the 2nd gate and then faded away.

This pattern of results eliminated a possible interpretation according to which the pattern obtained for /dl/ and /tl/ clusters is the result of *perceptual compensation.* This possibility is indeed suggested by the finding that ambiguous items on an English /d/ − /g/ continuum have been judged as alveolar in the nonfronting context /ar/− but as velar in the fronting context /al/− (Fowler, Best, & McRoberts, 1990; Mann, 1980, 1986). Accordingly, an initially high percentage of dental judgments should be obtained with fronted variants of velars, as expected in the context of a following alveolar liquid: Velar plosives should be sufficiently close to dentals to be judged as such before the subsequent segmental context is available and can be used to compensate for the initially perceived frontedness. However, in our data, few initial dental judgments and shifts to velar judgments were observed for the /gl/ and /kl/ clusters, whereas this pattern was frequently present for the /dl/ and /tl/ clusters. Therefore, the many early dental judgments, later shifting to velar judgments when /l/ was detected, had to reflect a perceptual shift rather than a perceptual compensation phenomenon. For /dr/ and /tr/ items, the convergence toward correct (i.e.,

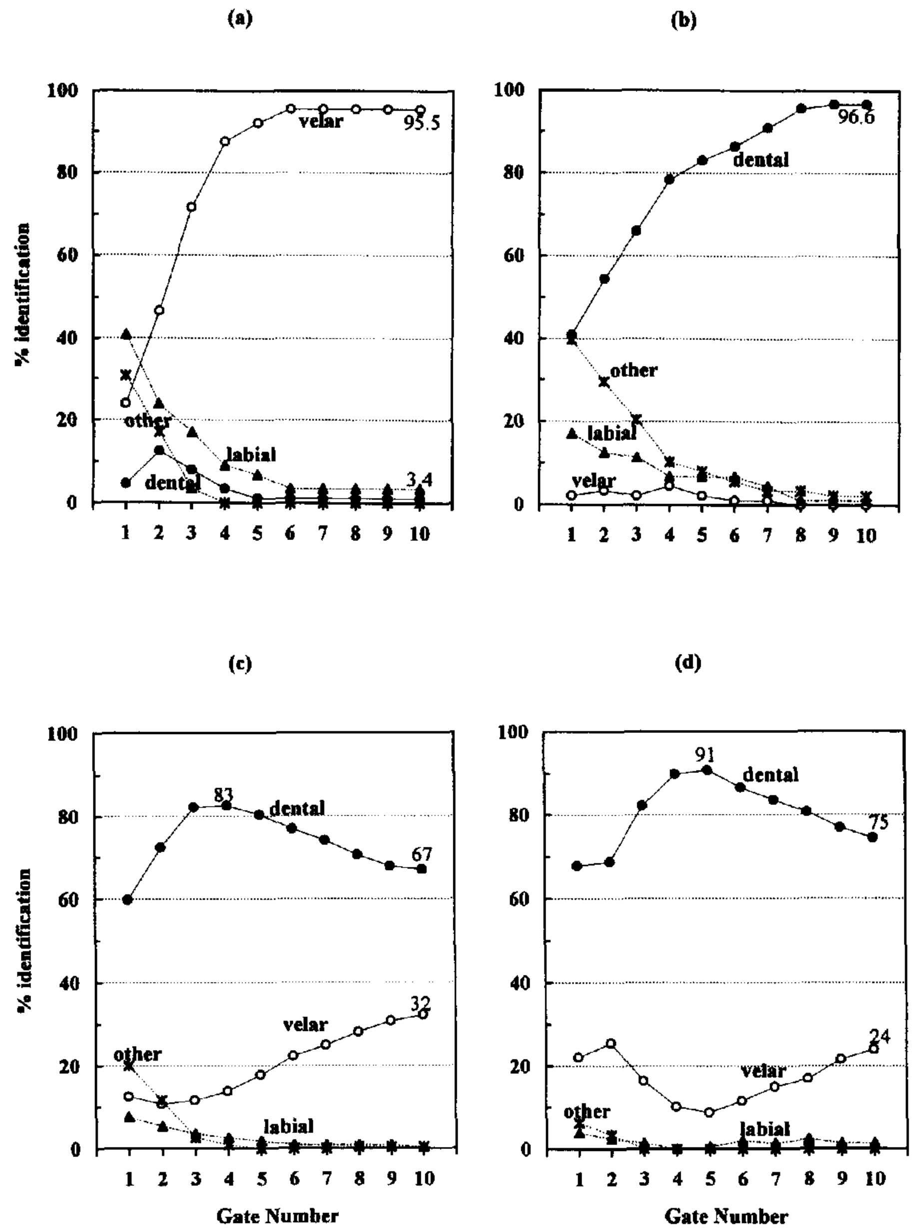

Figure 3. Judgments of initial consonant place of articulation according to gate number: /kl/ and /gl/ control items (successive format; A), /tr/ and /dr/ control items (successive format; B), /tl/ and /dl/ test items (successive format; C), and /tl/ and /dl/ test items (duration-blocked format; D). A–C show the data from Experiment 3, and D shows the data from Experiment 4.

dental) responses was slower than for /gl/ and /kl/ items, reaching the 85% level at the 6th gate. This was perhaps related to the longer duration of /r/ than /l/ (see Table 3). Still, the important point is that there was no decrease in the percentage of dental responses for dental + /r/ items, just as there was no decrease of velar responses for velar + /l/ items. This difference with the dental + /l/ items suggests that a perceptual shift was at work for illegal clusters, not for legal clusters. Because the control situations provided by the matched /gl/, /kl/, /dr/, and /tr/ items yielded sufficiently clear results—/gl/ and /kl/ were perceived as velar onsets and /dr/ and /tr/ as dental onsets—the analyses reported next are limited to the items of interest: the /dl/- and /tl/-onset items.

The mean gate number at which /l/ was identified was 3.6 (SD = 0.7); the mean gate number at which the vowel following /l/ was reported (correctly or not) was 4.6 (SD = 0.6); finally, although the last gate did not reach the upcoming /b/ or /p/ consonant release burst (see Table 3), there were sufficient cues in the speech signal for a majority of participants (an average 17.6 participants out of 28) to detect a labial plosive (mean gate number = 7.3, SD = 1.0).

Participants' responses were analyzed according to place of articulation (dental, velar, or labial), the order of presentation, and the gate number. Order had no significant effect: by subject, $F_1(1, 26) < 1$; by item, $F_2(1, 42) < 1$. Place was highly significant, $F_1(2, 52) = 92.2$, $F_2(2, 28) = 98.2$, $ps < .0001$. Overall, there were many more dental than velar responses: 73.6% against 20.4%, $F_1(1, 26) = 49.8$, $F_2(1, 14) = 55.4$, $ps < .0001$, whereas labial responses were negligible (2.1%). Gate was significant, $F_1(9, 234) = 14.2$, $F_2(9, 126) = 27.3$, $ps < .0001$. Indeed, there was substantial variation in the proportion of dental and velar responses across gates (see Figure 3C). The largest variation, however, seemed to occur in the first few gates, where the percentage of dental responses increased from an initial 60% to 83% at Gate 4.

Was the decline in dental responses after the 4th gate significant? This point was examined by limiting the analyses to Gates 4–10. The effect of gate on dental responses was significant in this portion: $F_1(6, 156) = 12.5$, $F_2(6, 84) = 13.3$, $ps < .0001$. Moreover, the correlation between gate number and percentage of dental responses from the 4th until the 10th gate was significant: by item, $r(54) = -.38$, $p = .004$; by subject, $r(194) = -.21$, $p = .004$. The decrease in dental responses between Gates 4 and 10 was thus real. It was larger for /tl/- than for /dl/-items: by-item analysis, 25% against 6.3%, $t(6) = 3.99$, $p = .007$.

The individual contribution of the different items was estimated by collapsing all 10 gates: In a by-subject analysis, an item effect could thus be examined in addition to place and order. Overall, item was marginally significant, $F_1(7, 182) = 2.0$, $p = .062$. As can be seen in Table 4, three items led to fewer dental responses (thus more velar responses) than the others. These were "dlopat," "dlopta," and "tlobda." (Recall that the burst of [d] in "dlopta" was spectrally the least dental of all dental + /l/ items.) Item variability was indeed to be expected because we used natural speech. However, even the "least dental" test item collected an average 53% of dental judgments across the 10 gates.

In summary, the results show that, overall, the initial stops of /dl/- and /tl/-items were primarily identified as dental rather than velar. However, participants revised their dental judgments and reported more and more velars from the 4th gate on. Still, the percentage of dental responses did not drop below 67%. Therefore, it is unlikely that there were cues to velar articulation in the speech signal for participants to report, for example, /glapɔt/ for "dlapot" in the preceding experiments.

The shift toward velar identification began to occur by the 4th gate, where /l/ and the following vowel were first identified. From this point on, the dental judgments declined and were revised into velar judgments. This points to a perceptual shift that was contextual rather than simply to a misperception of the illegal cluster's first constituent. Although the trend for a dental-to-velar shift was significant, it led to many fewer velar responses than in the preceding experiments. At the last gate, however, the entire consonant–consonant–vowel (CCV) portion of the test item was presented (e.g., /tla/ for "tlabod" or "tlabdo"), that is, sufficient information for a contextual velar shift to occur in principle. So why was the rate of velar responses so low compared with the first two experiments? One possibility is that the assimilation process at work, whatever its nature, takes time to unfold completely and that the responses given to the gated stimuli reflected an intermediary state of the process, not a final postperceptual categorization decision. Another possibility is that the particular format of forward gating used in Experiment 3, the successive format of presentation, induced perseveration in participants' responses (Walley, Michela, & Wood, 1995; however, see Cotton & Grosjean, 1984). Perseveration in reporting dental sounds might have masked the strength of the dental-to-velar shift effect. Walley et al. (1995) found that a duration-blocked format of presentation avoids perseveration effects because, in this format, successive stimuli are fragments of different (and unpredictable) items. We thus replicated the gating task of Experiment 3 using the same stimuli but with a duration-blocked format of presentation to test for the perseveration explanation of the low rate of velar shifts that was found in this experiment.

Table 4
Average Percentages of Dental Judgments Across All 10 Gates According to Test Items

Item	Type of presentation: Successive	Duration blocked	*M*
dlapot	90.7	91.2	90.9
dlapto	82.9	80.0	81.4
tlabod	79.3	80.0	79.7
tlabdo	80.4	83.8	82.1
tlobad	78.9	81.9	80.4
tlobda	68.9	75.0	71.9
dlopat	55.0	72.7	63.8
dlopta	52.9	76.9	64.9
M	73.63	80.19	76.89

Experiment 4: Gating Task With Stimuli Blocked by Duration

This experiment differed from Experiment 3 only in the format of presentation. Gated items were blocked by their duration so that response perseveration could not possibly bias the results. In particular, the percentage of dental responses at the last gate, which might have been inflated in Experiment 3, could now decrease further. A control test using matched /gl/, /kl/, /dr/, and /tr/ items was no longer necessary since the dental nature of the /dl/ and /tl/ items was clearly established in Experiment 3. The issue at stake was now whether or not a stronger shift from dental to velar judgments would appear in the duration-blocked presentation format.

Method

Stimuli. The stimuli were the same as in the first test of Experiment 3.

Participants. Twenty-six undergraduates, aged 22–28 years, who were native speakers of Geneva French participated for course credit in the experiment. We did not expect significant differences with Parisian French in this experiment because the two dialects differ mainly with respect to prosodic dimensions. None of the participants reported hearing problems.

Procedure. The stimuli presented were blocked by duration. Each block contained 16 stimuli corresponding to the same gate number for each item; item order was randomly changed from one block to the next, and successive blocks corresponded to increasingly larger gates.

Results and Discussion

The results obtained were essentially the same as in the previous gating experiment. The percentages of dental, velar, and other transcriptions for the /dl/ and /tl/ items are shown in Figure 3D. The percentage of dental responses reached a maximum at the 5th gate (91%), then gradually decreased (to 74.5% at the last gate) while velar responses increased symmetrically, reaching 24% at the last gate.

The mean gate number in which /l/ was identified was 3.6 (SD = 0.8). The mean gate number in which the vowel following /l/ was reported (correct or not) was 4.5 (SD = 0.8). Finally, a few participants (an average 6.6 out of 26) were able to detect a labial plosive in final position (mean gate number = 8.5, SD = 0.9).

Analyses similar to those for Experiment 3 were performed, except that there was no order variable. In the analysis of participants' responses according to place and gate, the same effects as in Experiment 3 were found: place, $F_1(2, 50) = 126.7$, $F_2(2, 14) = 382.5$, $ps < .0001$; gate, $F_1(9, 225) = 2.9$, $p = .0033$, $F_2(9, 63) = 9.6$, $p < .0001$. Dental responses were much more frequent than velar responses (80.2% vs. 17.2%): $F_1(1, 25) = 76.2$, $F_2(1, 7) = 232.9$, $ps < .0001$. Labial responses were negligible (1.7%). The decline following the 5th gate was significant, as shown by analyses restricted to Gates 5–10. The effect of gate on dental responses was significant in this portion: $F_1(5, 125) = 7.4$, $F_2(5, 35) = 14.7$, $ps < .0001$. The correlation between gate number and percentage of dental responses, from the 5th until the 10th gate, was significant: by item, $r(46) = -.68$, $p < .0001$; by subject, $r(154) = -.21$, $p = .008$. The decrease in dental responses was significantly larger for /tl/− than for /dl/-items (23% vs. 9.6%), $t(6) = 2.94$, $p = .026$. As shown in Table 4, between-items differences paralleled those found in Experiment 3. Again, "dlopat," "dlopta," and "tlobda" led to somewhat fewer dental responses.

In summary, the results were similar to those obtained in the first gating task. The percentage of dental responses at the 10th gate was not smaller in Experiment 4 (74.5%) than in Experiment 3 (67%). Perseveration effects thus could not explain the high percentage of dental responses in the gating tasks relative to those obtained in Experiments 1 and 2.

The results of Experiments 3 and 4 clearly demonstrate that the initial consonants in the /dl/- and /tl/-initial items did not have an intrinsic velar quality. Once a stable perception of place of articulation emerged (by the 4th gate), it was frequently "dental" rather than "velar." In fact, even when most of the first syllable was presented, the rate of dental-to-velar shifts did not exceed 33%, whereas an overwhelming dental-to-velar shift was found in the first two experiments. Aside from the difference in the size of the assimilation effect, the identification and gating results were strikingly similar in detail. In particular, the higher rate of confusions for /tl/ than for /dl/ in the identification tasks was paralleled by more frequent revisions from dental to velar judgments for /tl/ than for /dl/ in the gating tasks.

Was the modest size of the velar shift observed in the gating tasks caused by the limited amount of information presented? We observed that the velar shift began to appear around the 4th or 5th gate, just after the liquid context was identified, and continued to increase in strength up and through the last gate, when the entire word-initial CCV portion (i.e., a full acoustic syllable) was presented. Because the following syllable presumably does not contribute to the perception of the item-initial cluster, it is unlikely that more acoustic information was needed for the assimilation process to build up fully; rather, more time was needed. We assume that the gating responses provide a snapshot picture of the state of processing at the point in time that corresponds to the end point of each gated fragment. In other words, phonemic gating appears to tap into the partial product of the assimilation process before its completion.[7] We thus assume that the proportions of dental and velar responses in the gating experiments reflect the time course of the assimilation process. We turn now to the nature of this process.

One possible account for the "contextual" perceptual assimilation assumed here is in terms of a competition process between the activations of units at two levels (at least) of prelexical representation: single phonemes and syllable onsets. Dental plosives are initially (i.e., for the shortest fragments presented) more strongly activated than velars and velars more activated than labials (for acoustic and articulatory reasons). However, as soon as the bottom-up information begins to activate the following /l/, the activation of velar plosives begins to increase at the expense of the dental plosives, presumably because of top-down feedback from the onset-level representations favoring velars over dentals: Whereas /tl/ or /dl/ syllable onsets are, in practical terms, ruled out by their low phonotactic probability, /kl/ and /gl/ are not and can in turn activate the velar consonants. Labial plosives might conceivably benefit from a similar feedback activation but cannot overcome their initial handicap relative to velars. According to this account, the decline of dental responses and the increase of velar responses that develops as longer fragments are presented reflect the time course of phoneme-level activations.

Another account of the data is provided by noninteractive

[7] The responses could have reflected a later state of the process at work because the phonemic gating task imposed only a limited time pressure. One possible reason why they did not is that the gating task induced a kind of a minima strategy: The stimuli presented earliest in a session were so short, that is, so poor in phonetic information that they encouraged—throughout a session—responses based on low-level, straightforward, phonetic perception rather than on phonological knowledge at a decisional stage.

models such as the fuzzy logical model of perception (FLMP) (Massaro & Cohen, 1993; Oden & Massaro, 1978). The FLMP would claim that individuals determine the dental as opposed to the velar or labial place of articulation of the item-initial plosives on the basis of the evaluation and integration of two independent sources of information: the plosive itself and its phonological context. The first source provides support for "dentalness" to the extent that the acoustic features of the plosive agree with a prototype /d/ or /t/; for the /dl/ and /tl/ test items, this source provides weaker support for "velarness" and a very weak support for "labialness." The second source (the #_/l/ context) provides independent evidence for velarness or labialness but not for dentalness, supporting only the alternatives that are phonotactically compatible with the following /l/ context (as well as with the word-initial position). A relative-goodness rule applies at the decision stage, yielding quantitative predictions for the various alternatives. Also, both sources of information are processed over time. Conceivably, the contextual information occurs later in items such as "dlapto." The FLMP could account for this by weighting the various amounts of support by a negatively accelerated growth function of processing time and introducing a fixed delay for the arrival of right context information (see Massaro & Oden, 1980, 1995). Appropriate parameters (6 degrees of "static" support for three place alternatives × two sources of information, rates of information processing, and time delays) could probably produce accurate predictions of the dental, velar, and labial responses over time. Note that evaluating the support provided by phonological context (being in word-initial position and followed by /l/) for each type of response is not basically different from evaluating the plausibilities of the corresponding clusters as syllable onsets or from assigning them initial levels of activation. The latter account is perhaps more appealing in that prelexical representational units are explicitly posited, and it avoids the underspecified notion of contextual source of information.

In summary, the experimental data do not allow choosing from among the main competing theories. According to both the interactive and independent views, though, the perceptual shift observed here seems to involve automatic processes rather than postperceptual decisions. However, more evidence is needed to conclude that this perceptual shift results from an automatic process. An on-line task, such as phoneme monitoring, could well provide such evidence.

Experiment 5: Detection of Cluster-Initial Consonants

Suppose that perceptual assimilation occurs in the on-line processing of speech as reflected by a task such as phoneme monitoring. Participants should often miss dental targets in the /dl/ and /tl/ carrier items but should incorrectly detect velar targets in these same items. Such a pattern of on-line misses and false positives would be analogous to the dental-to-velar confusion data that were obtained in off-line tasks. Detection latencies of the velar targets in the dental + /l/ items could be longer than those for the same targets in the matched velar + /l/ items. Such a result would suggest an additional processing cost attributable either to a reanalysis of the phonetic input or, in terms of competing activations, to a delay required for the velar activation to overcome the initial dental activation.

In this experiment, we used the same /dl/ and /tl/ items as in the gating and identification tasks, but this time we used them in a generalized phoneme monitoring task (cf. Frauenfelder & Segui, 1989, for a discussion of this technique). In addition to these items, participants were also presented with the corresponding legal velar + /l/ as well as dental + /r/ sequences.

Method

Stimuli and design. The same 32 nonwords items that were used in the preceding experiments served as target-bearing carriers in this experiment, that is, 8 legal (/tr/, /dr/) and 8 illegal (/tl/, /dl/) dental + liquid clusters and 8 legal (/kl/, /gl) and 8 legal (/kr/, /gr/) velar + liquid clusters. These items were associated with four different lists, one for each phoneme target (/t/, /d/, /k/, and /g/). Each list for velar targets contained four disyllabic "foil" items with the corresponding illegal dental + /l/ clusters to test for possible incorrect detections resulting from perceptual assimilation. Similarly, the dental lists contained corresponding velar + /l/ foils to assess the probability of misperceptions in these cases.

Each list also contained items with targets in other positions and items without targets. Four sets of 16 disyllabic nonwords containing the respective targets in noninitial positions (medial and final) and two sets of 48 similar nonwords not containing the targets were constructed with a phonological structure similar to that of the target-bearing test items. Thus, each list contained the following items: 8 item-initial target test items, 16 noninitial (medial and final) target-bearing fillers, four foils with the other place of articulation (velar or dental), and 48 non–target-bearing items. Finally, one distractor list containing the same distribution of test and filler items also was constructed for a labial plosive target.

Four groups of participants were each assigned to two test lists according to the phoneme targets (/t/ and /k/ targets vs. /d/ and /g/ targets) and to the list presentation order (dental target first vs. velar target first). All participants first received a practice list with 36 stimuli and a phoneme target not used in the test lists but appearing equiprobably in word initial, medial, or final position. The same distractor list with labial targets was inserted between the two test lists for all four groups. This distractor list served primarily to determine a common miss rate criterion across the phoneme monitoring lists. (Targets were unambiguous in this list, and a low miss rate was expected.) It also served to make less apparent the presence of phonotactically illegal items during the experimental sessions. The new stimuli were produced by the speaker who had recorded the material of the previous experiments.

Participants. Fifty-two undergraduates at the Université Paris V, aged 21–30 years, participated in the experiment for course credit. Four students did not reach the acceptance criterion (less than 50% misses) in the distractor list. No one reported any hearing difficulties.

Procedure and apparatus. Participants were tested individually in a sound-attenuated booth. They received the stimuli via Sennheiser headphones at a comfortable listening level. They were told that they would hear nonwords and would have to make a speeded detection response to target phonemes by pushing a Morse key with their preferred hand. They were instructed that targets could occur anywhere in the carrier nonwords. Targets were specified in a booklet of cards on which they were written in capital letters. The phonetic value of each letter had been explained and illustrated before the experiment. In particular, it was emphasized that the letter *G* stood for /g/ as in "gare" /gar/, not for /ʒ/ as in "gilet" /ʒilɛ/. Participants were orally warned (through the headphones) when the list changes occurred and were instructed at this point to switch to the next card for the subsequent target. The presentation of stimuli and of oral instructions was controlled by a 486/66 microcomputer interfaced with an AD/DA OROS board (using 10-kHz sampling frequency and 16-bit resolution). The computer also collected reaction times (RTs), measured from the release burst of the targets. (All the targets were plosive consonants.) RTs were computed on the basis of the number of samples that had been run through the digital-to-analog converter at the moment the key was pressed.

Results and Discussion

The percentages of misses and false positives are shown in Figure 4. As can be seen, there were many false-positive responses to velar targets and misses of dental targets for the illegal cluster dental + /l/ items. By contrast, false-positive responses to dental targets as well as misses of velar targets were rare for the legal velar + /l/ items. This pattern of results provides strong support for the view that /dl/ and /tl/ are heard as or assimilated to /gl/ and /kl/ in on-line speech processing.

The detection rates of dental and velar targets were analyzed according to cluster type (illegal /dl/ or /tl/ vs. legal /gl/ or /kl/) and experimental group (target voicing and order of presentation). Overall, neither order nor voicing reached significance: order, $F_1(1, 44) < 1$, $F_2(1, 24) < 1$; voicing, $F_1(1, 44) = 1.9, p = .17$, $F_2(1, 24) = 1.4, p = .24$. However, one group (voiced targets and dental target first) produced a comparatively high hit rate for the dental targets (68.8%) and low false-positive rate for the velar target (66.7%) in the illegal dental + /l/ items. This group differed significantly

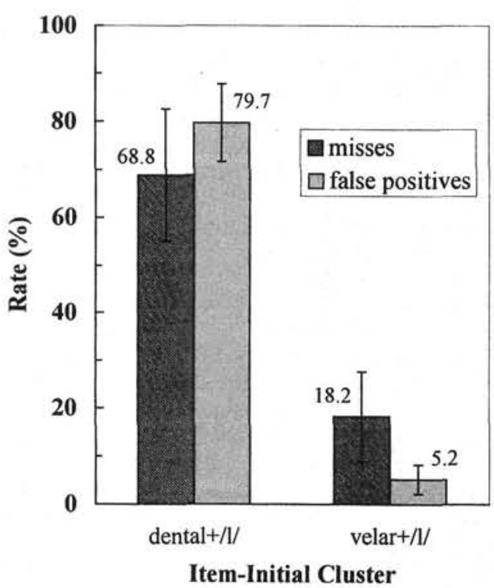

Figure 4. Miss and false-positive rates for dental and velar targets in dental + /l/ versus velar + /l/ items.

from the other "voiced target" group in the illegal cluster condition, $F_1(1, 44) = 5.9, p = .018$; $F_2(1, 24) = 5.1, p = .032$. Despite the variability introduced by this group, the overall pattern of results was significant. In the illegal /dl/ or /tl/ items, false-positive responses to velar targets were more frequent than hits to dental targets (79.7% vs. 33.7%), $F_1(1, 44) = 78.3$, $F_2(1, 12) = 51.6$, $ps < .0001$. Moreover, these false-positive responses were about as frequent as the hits to velar targets in legal /gl/ or /kl/ items (79.7% vs. 81.8%), $F_1(1, 44)$ and $F_2(1, 24) < 1$. False positives to dental targets were rare in /gl/ and /kl/ items (5.2%). There was a nonsignificant trend for the false-positive rate to be higher for /k/ than /g/ in dental + /l/ items (83.3% vs. 76.1%), $F_1(1, 44) = 1.1, p = .30$; $F_2(1, 12) = 1.21, p = .29$.

To judge the effect of cluster type on the detection of dentals, we compared directly the miss rates for dentals in illegal (/dl/ or /tl/) and legal (/dr/ or /tr/) clusters. The miss rate was higher in illegal than in legal clusters (68.8% vs. 25%), $F_1(1, 44) = 155.9$, $F_2(1, 24) = 43.2$, $ps < .0001$. There were more misses for /t/ than for /d/ target (81.2% vs. 68%) in illegal clusters, $F_1(1, 44) = 11.9, p = .0013$, $F_2(1, 12) = 7.9, p = .015$, but this was partly due to the group of participants assigned to the detection of /d/ and list order dental–velar, in which the miss rate of /d/ in /dl/ was unusually low (31.2%). Overall, the detection rate of dentals in the illegal clusters was so low that an RT analysis was meaningless. However, the prediction that dental targets would be difficult to detect in illegal dental + /l/ clusters already received strong support from the miss-rate data.

The second issue of interest was whether velar targets would be more difficult to detect in the dental + /l/ than in the velar + /l/ clusters. Such difficulty was not apparent in the detection rates (79.7% false positives in dental clusters vs. 81.8% hits in velar clusters). It was, however, reflected in the RT data, which could be compared directly given the relatively high and roughly comparable detection rates. Participants were slower to detect, for example, /g/ in /dl/ (i.e., to false positive because of perceptual assimilation) than in /gl/ by about 34 ms (647 vs. 613 ms), as can be seen in Figure 5. This difference was marginally significant in the by-item analysis, $F_2(1, 24) = 3.24$, $p = .081$, but was reliable in the by-subject analysis, $F_1(1, 46) = 6.4, p = .014$. The target variable (i.e., /g/ vs. /k/) was not significant (F_1 and $F_2 < 1$).

Between-items differences tended to parallel those observed in the gating and off-line identification tasks: The perceptual assimilation effect was somewhat stronger with voiceless than with voiced dental + /l/ clusters. In the speeded detection task, this was apparent in the trend for the false-positive rate to velars to be higher in /tl/ than in /dl/ items and in the higher miss rate for dental detection in /tl/ than in /dl/.

A striking outcome of Experiment 5 was that the perceptual assimilation effect was larger than in the off-line identification task (80% false-positive rate in detecting velars in dental + /l/ clusters vs. an average 61.5% dental-to-velar identification confusions). Because the assimilation effect was so great in an on-line task, we can conclude that it is automatic and irrepressible, taking place at an uncon-

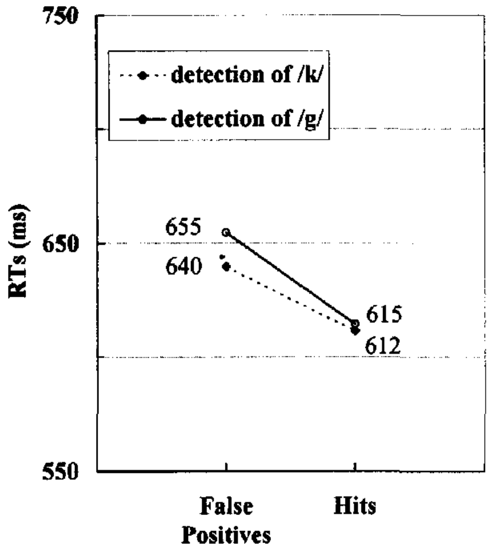

Figure 5. Detection latencies of velar targets in dental + /l/ items (false positives) versus velar + /l/ items (hits). RTs = reaction times.

scious level of processing. However, the obvious question concerns the exact level of processing at which this perceptual assimilation takes place. We used homogeneous lists of nonwords to discourage a lexical strategy (see Frauenfelder, Segui, & Dijkstra, 1990, for a discussion of the list structure effects obtained, e.g., in Marslen-Wilson, 1984), and consequently participants presumably did not rely on lexical information. However, it also appears that participants did not respond solely on the basis of phonetic-acoustic information. If this had been the case, they would presumably have identified the dentals in the dental + /l/ clusters as they did in the gated stimuli, in which dental perception in the early gates evolved toward velar perception. By analogy, participants would have been faster to detect dentals than velars in illegal clusters on the rare occasions that they did so. However, they were not faster (642 vs. 647 ms for 31% vs. 80% detection rates, respectively). Conceivably, then, the on-line detection of velar targets was made after the perceptual integration of the segments making up the onset cluster was completed. This suggests that perceptual assimilation occurred at a moment in the time course of perceptual integration at which a more abstract code than a phonetic code had been built and where chunks larger than individual phonemes had been assembled. At the same time, this is an indication that phoneme detection is based on the perception of prelexical units larger than the phoneme, such as onsets. By contrast, as we proposed in the discussion of Experiment 4, the data for early gates seem to reflect phonetic perception before the perceptual integration of clusters has taken place.

General Discussion

Experiments using a variety of different tasks have provided consistent evidence for the perceptual assimilation of illegal clusters (/tl/ and /dl/) into legal ones (/kl/ and /gl/) *in both the on-line and off-line processing of nonwords. In* the open-response task, on which participants could use their phonological and orthographic knowledge postperceptually, the effect was strong and perhaps artificially inflated. The gating results confirmed that the phonetic quality of the cluster-initial consonants in the /tl/ and /dl/ items was dental rather than velar. The gating data also revealed the time course of the perceptual assimilation phenomenon. It began to unfold after the /l/ segment was identified and developed gradually thereafter. Finally, perceptual assimilation was found to be strongest in on-line processing, as reflected by the phoneme monitoring task, and there was a significant trend for velar targets to be detected more slowly in illegal rather than legal clusters.

We have discovered a robust phenomenon that is not attributable to the misperception of the cluster initial constituent (the dental stop) but is induced by the combination of the two constituents making up the syllable onset. We call it "contextual perceptual assimilation." It is consistent with the notion of a language-specific phonological filter in which nonnative phones tend to be assimilated to native categories. Our findings show that this notion can be extended to explain perceptual assimilation within the native language of the listeners for sequences of phonemes such as consonant clusters. Certain phonotactically illegal but possible clusters assimilate to legal clusters. There is no reason to believe, however, that this phenomenon can be generalized to all illegal clusters. What we found for /dl/ and /tl/ in French does not necessarily apply to such illegal clusters as /nt/ or /ms/. As was pointed out earlier, /dl/ and /tl/ sequences are predicted to induce assimilation because they represent an accidental gap in the otherwise regular distribution of OBLI clusters in French. Among the languages of the world, /dl/ and /tl/ clusters do exist but seem to be infrequent and perhaps unstable combinations.[8] Although a systematic study of clusters and their evolution in various languages is not yet available, there is anecdotal evidence that dental and back articulated plosive + /l/ clusters do not coexist in certain languages, with these two types alternating in the different dialects of a same family.[9] We may surmise that

[8] In both Slavic (Vaillant, 1950, pp. 88–90) and Latin (Bliville, 1990, p. 318) languages, a common trend is the evolution (from old Slavic, Greek, or archaic Latin) from initial /tl−/ to /l−/ (e.g., *dlingua > lingua, *tlatus > latus, etc.) and from medial /−tl−/ to /−kl−/ (e.g., *−tlo− suffix > −c(u)lo−, vet(u)lus > ueclus). Although word-initial /tl/ and /dl/ reappeared in modern Slavic languages, they are not frequent and again, tend to change. For example, Polish has a /l/ : /ł/ liquid contrast (still denoted by the written forms "l" and "ł"). After /d/ and /t/, /ł/ evolved to /w/: The written forms "dł" and "tł" are actually pronounced /dw/ and /tw/ in the standard variety of Polish; /dl/ and /tl/ have not undergone a similar diachronic change but are rare in this language.

[9] Such is the case in the Miao-Yao group (Niederer, 1995). The various dialects of this group may have a series of dental stop + /l/ clusters or a series of uvular stop + /l/ clusters, such as /tl/ or /ql/, but not both. Wang (1979) proposed that these clusters all originate from velar–uvular clusters in proto-Miao. Certain words of proto-Miao with a back stop + /l/ cluster onset are borrowed from archaic Chinese words whose onset was most probably dental (e.g., N-lwat

such clusters are intrinsically confusable and that this is why certain languages do not include /dl/ and /tl/ in their cluster repertoire. According to this view, the perceptual assimilation we observed in modern French is just another facet of the same intrinsic confusability. Conceivably, then, the dental stop + /l/ clusters could produce perceptual assimilations in other languages than French.

A consistent finding across the different experiments was that /dl/ was more resistant to assimilation than /tl/. We may ask why. A first explanation appeals to some perceptual differences. The data on (American) English phoneme confusability under various signal-to-noise ratios (Miller & Nicely, 1955) show that English /d/ is perceptually more robust than /t/: 47% confusions for /d/ versus 66% for /t/ (for signal-to-noise ratio = −12 dB). Interestingly, both alveolar plosives are misperceived as velars more often than as labials: 33% versus 18% (computed from Miller & Nicely, 1955, Table 18). Of course, the French contrast /d/ : /t/ is different. Voiced plosives are prevoiced and voiceless plosives have a short voicing lag. Some cues to place of articulation may be present in the prevoiced portion. In addition, the prevoiced portion by itself sounds like a nasal murmur and may be viewed as a weakly resonant vowellike sound, which somehow enhances the perceptibility of the /dl/ cluster.[10]

An alternative explanation for the difference in perceptual assimilation between the two illegal clusters refers to their relative frequency of usage. Speaking of "frequency of usage" may seem strange in the case of illegal sequences, but both illegal sequences /dl/ and /tl/ are sometimes found utterance-initially in contracted forms such as /dlo/ for "De l'eau" ([some] water) or /tlavy/ for "Tu l'as vu?" (did you see him/it?). The former construction (partitive article + noun), however, is probably more common and more productive than the second one.[11] That French listeners are more often confronted with utterance-initial /dl/ than with /tl/ sequences in running speech might partly explain why they are less likely to assimilate /dl/ than /tl/. We believe, however, that this can only be part of the story and is perhaps a consequence rather than a cause of the greater robustness of /dl/ in French, which may be largely explained by basic acoustic and perceptual characteristics.

As discussed earlier, the FLMP model can accommodate the contextual perceptual assimilation phenomenon, including its time course, without making any explicit reference to phonological levels of representation. (However, this approach implicitly requires that phonotactic knowledge be represented somehow.) Contextual perceptual assimilation can also be understood within an interactive activation framework. Here, an initially strongly activated phoneme—a dental stop—is inhibited and eventually overcome by a different—initially weakly activated—phoneme whose increase in activation is attributable to top-down feedback received from representations at a higher level. The TRACE model (McClelland & Elman, 1986) would explain the perceptual shift in terms of lexical-to-phoneme feedback because it does not propose any intermediary prelexical levels of representation higher than the phoneme. However, there is no reason to exclude the possibility of other prelexical levels of representation and perceptual integration. For instance, there is evidence in French that the syllable and its constituent parts may be involved in on-line speech processing (Mehler, Dommergues, Frauenfelder, & Segui, 1981). Consequently, we are inclined to posit a higher level of sublexical representation (i.e., of perceptual integration). More specifically, we appeal to a syllabic or syllabic onset level of representation to explain how phoneme activation levels can be modified differentially in a top-down fashion when the phoneme sequences are or are not compatible with these higher level representations. In such models, phonotactic knowledge can be represented either by differential resting levels of activation or by specific activation thresholds for syllabic or onset nodes. Models incorporating prelexical levels of representation in this fashion would predict a delay in the activation of phoneme-level representations when a contextual perceptual shift has occurred. Indeed, the top-down activation (causing the shift) takes time to overcome the initial bottom-up activation. This is consistent with the slower increase in velar responses in the /dl/ or /tl/ clusters than those obtained in /kl/ or /gl/ sequences in the gating experiment and with the longer detection latencies found in Experiment 5 for velar detection.

With the TRACE model, top-down feedback is assumed to be lexical. In the case of the dental + /l/ items, even though no word is ultimately recognized, all the words sharing a velar + /l/ onset are activated to some degree and are the source of the feedback activation of velar stops. With models incorporating, for example, onset-level representa-

---

[nᵈlwat] > dlwat > tʰuo; from archaic to modern standard Chinese). Hence, assimilation probably occurred at the time of the borrowing. In some modern Miao-Yao dialects, such as Shi Men, uvulars changed back to dentals in stop + /l/ clusters; in others they did not, but in either case uvular + /l/ and dental + /l/ series are now mutually exclusive.

[10] Indeed, /tl/ and /dl/ are permissible in intervocalic position (i.e., word-medially) as in "atlas," "atlantique," and so on. (Examples for /dl/ are infrequent words or loan words: "iodler" /jɔdle/, "landlord" /lɑ̃dlɔrd/, or found in family names: "Adler" /adlɛr/.) Word-medial schwa-deletion results in a somewhat larger number of word-medial /dl/ and /tl/ (as in "châtelain" /ʃɑtlɛ̃/, "godelureau" /gɔdlyro/). Hence, /tl/ and /dl/ are probably "correctly" perceived in word-medial position (i.e., after a vowel). If the prevoicing murmur plays a similar role as a faint vowel, this could explain why /dl/ is less often assimilated to /gl/ than /tl/ is to /kl/.

[11] Phonetic counts based on a corpus of spoken language (Tubach & Boë, 1990), including contracted forms, indicate that the /tl/ sequence is about three to four times more frequent than /dl/ regardless of word position (raw counts: 134 vs. 39 out of about 300,000 diphones). On the other hand, lexical counts based on written language (excluding possible contractions in real speech) indicate that /tl/ is about 10 times more frequent than /dl/ word-medially ("Trésor de la Langue Française"; Imbs, 1971). The discrepancy between spoken and written data (more /dl/ than expected in spoken language) therefore suggests that /dl/ is a more frequent contracted form than /tl/ in real speech. (Word-medially, /dl/ and /tl/ contracted forms have similar frequencies of usage.)

tions, the source of the feedback is prelexical. In other words, the latter models leave open the possibility that phonotactic knowledge and constraints are not exclusively coded in the mental lexicon. A similar view was held by Pitt and Samuel (1995), who proposed to distinguish lexical and sublexical interactive feedback mechanisms: "The hypothesized sublexical mechanism . . . produces top-down effects similar to a lexical feedback mechanism" (p. 184). In other words, prelexical effects are able to "mimic" lexical effects.

Our results also can be viewed from a more phonological perspective. Normal phonetic perception of fully articulated speech items (as opposed to gated fragments) may be based on syllabic onset processing units rather than on the component singleton phonemes. In other words, a syllable-initial cluster may be perceived as a whole. This claim is supported by the widespread uncertainty about the single phoneme versus cluster status of certain phones with a complex articulation (see Footnote 1). In this perspective, word-initial /tl/ may be processed as an atypical, or bad, exemplar of the /kl/ category with an associated processing cost. In a recent model of perceptual assimilation (Best, 1990, 1994b), the focus is on the difficulty in perceiving nonnative contrasts when one or both members are assimilated as bad exemplars of a native category ("category goodness" and "single category" contrasts). We suggest that the difficulty is basically attributable to that of assimilating atypical exemplars. The assimilation of illegal clusters to their legal counterparts must introduce some additional processing difficulty.

The phenomenon observed here is reminiscent of the effect of "subcategorical mismatches" produced by cross-splicing speech manipulations (Marslen-Wilson & Warren, 1994; Streeter & Nigro, 1979; Whalen, 1984, 1991). The distortions resulting from cross-splicing are not consciously perceived but generally introduce perceptual integration difficulties depending on the stimulus structure, the lexical status of the stimuli, and its competitor environment. The observed patterns of RT costs provide new insights into lexical access and lexical representations (Marslen-Wilson & Warren, 1994). The case of assimilated illegal clusters is similar, in that a perturbation of the speech input goes unnoticed but nevertheless entails an increase in processing difficulty. The general implication is that listeners are biased toward perceiving events and objects that fit with their internal representations rather than with the actual stimuli. The bias may be even stronger when the perceptual object is a meaningful event, as is the case for words as opposed to nonwords. Here, a considerable mismatch between percept and stimulus can be tolerated, but at the expense of some additional processing. Experiments involving subcategorical mismatches in which such accommodating processes are at work are thus of particular interest. Phonotactic violations constitute yet another form of mismatch, with the presumed advantage of being ecologically more relevant. Illegal clusters still are natural speech, whereas subcategorical manipulations result in "speech" items that may sound relatively natural but that do not correspond to any human articulation. Contextual perceptual assimilation therefore has some potential methodological applications. Perceptually assimilated illegal clusters (inducing extraprocessing cost) could serve as distortions of the speech input at the phonetic level and help investigate further stages of speech perception: How, for example, do such distortions influence perceptual integration at various sublexical and lexical levels? Addressing such issues could shed light on the mechanisms whereby people perceive internalized plausible or meaningful events rather than their objective physical manifestations.

Finally, we return to the issue of the origin of the contextual perceptual assimilation phenomenon. Lexical mediation as predicted by the TRACE model cannot be excluded, although the results of the gating experiments point to an assimilation process emerging at a sublexical level within prelexical units such as CCV. However, it can be argued that lexical activation is possible even on the basis of partial information (see McClelland & Elman's, 1986, pp. 33–35, critique of Massaro & Cohen's, 1983, notion of phonological context effect). According to a different view, the implicit knowledge of phonotactic constraints may originate at a more basic level. Listeners may know that certain combinations of sounds are legal or not, not only by reference to their mental lexicon but also on the basis of a deeper intuition: "Speakers must have tacit knowledge about how phonemes are strung together in their language" (Pinker, 1994, p. 173). Such knowledge is more primitive, as it were, than lexical knowledge in the sense that it developed from early infancy. Data from infants are more consistent with this second view. Nine-month-olds, that is, infants with a small or no receptive lexicon (see Hallé & Boysson-Bardies, 1994), are sensitive to phonotactic violations (Jusczyk, Friederici, et al., 1993) and even to the frequency of occurrence in the language of particular CVC sequences (Jusczyk et al., 1994), which probably correlates with the degree of familiarity that infants have with such sound combinations. Hence, before a lexicon has developed, young listeners are already "attuned to" the sounds and combinations of sounds heard in the ambient language. We surmise that this primitive (i.e., protolexical) knowledge of phonotactic constraints may survive in adults and partly explain the modified perception of illegal clusters such as /dl/ and /tl/ in French.

References

Best, C. (1990). Adult perception of nonnative contrasts differing in assimilation to native categories [abstract]. *Journal of the Acoustical Society of America, 88,* S177(A).

Best, C. (1991, April). *Phonetic influences on the perception of nonnative speech contrasts by 6–8 and 10–12 month olds.* Paper presented at the Society for Research in Child Development, Seattle, WA.

Best, C. (1994a). Learning to perceive the sound pattern of English. *Advances in Infancy Research, 9,* 217–304.

Best, C. (1994b). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). Cambridge, MA: MIT Press.

Best, C., McRoberts, G., Lafleur, R., & Silver-Isenstadt, J. (1996).

Divergent developmental patterns for infants' perception of two nonnative consonant contrasts. *Infant Behavior and Development, 18,* 339–350.

Best, C., McRoberts, G., & Sithole, N. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 45–60.

Bliville, F. (1990). *Les emprunts du Latin au Grec (French): Approche phonétique* [Borrowings from Greek to Latin: A phonetic approach]. Louvain-Paris: Peeters.

Blumstein, S. E., & Stevens, K. N. (1980). Perceptual invariance and onset spectra for stop consonants in different vowel environments. *Journal of the Acoustical Society of America, 67,* 648–662.

Brown, R. W., & Hildum, D. C. (1956). Expectancy and the perception of syllables. *Language, 32,* 411–419.

Cotton, S., & Grosjean, F. (1984). The gating paradigm: A comparison of successive and individual presentation formats. *Perception & Psychophysics, 35,* 41–48.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language, 25,* 385–400.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology, 24,* 381–410.

Dell, F. (1995). Consonant clusters and phonological syllables in French. *Lingua, 95,* 5–26.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1996). *Epenthetic vowels in Japanese: A perceptual illusion?* Manuscript in preparation.

Flege, J. (1986). The production and perception of foreign language speech sounds. In H. Winitz (Ed.), *Human communication and its disorders* (pp. 224–401). Norwood, NJ: Ablex.

Flege, J. (1991). Perception and production: The relevance of phonetic input to L2 language learning. In C. Ferguson & T. Heubner (Eds.), *Crosscurrents in second language acquisition and linguistic theories* (pp. 249–289). Philadelphia: Benjamins.

Fowler, C. A., Best, C. T., & McRoberts, G. W. (1990). Young infants' perception of liquid coarticulatory influences on following stop consonants. *Perception & Psychophysics, 48,* 559–570.

Frauenfelder, U., & Segui, J. (1989). Phoneme monitoring and lexical processing: Evidence for associative context effects. *Memory & Cognition, 17,* 134–140.

Frauenfelder, U., Segui, J., & Dijkstra, T. (1990). Lexical effects in phoneme processing: Facilitatory or inhibitory? *Journal of Experimental Psychology: Human Perception and Performance, 16,* 77–91.

Grosjean, F. (1980). Spoken word-recognition processes and the gating paradigm. *Perception & Psychophysics, 28,* 267–283.

Hallé, P. A., & Boysson-Bardies, B. (1994). The emergence of an early lexicon: Infants' recognition of words. *Infant Behavior and Development, 17,* 119–129.

Halle, M., Hughes, G. W., & Radley, J.-P. A. (1957). Acoustic properties of stop consonants. *Journal of the Acoustical Society of America, 29,* 107–116.

Imbs, P. (1971). *Études statistiques sur le vocabulaire français (French): Dictionnaire des fréquences—Vocabulaire littéraire des XIXe et XXe siècles* [Statistical studies on French vocabulary: Dictionary of frequencies—Literary vocabulary of the 19th and 20th centuries]. Paris: Librairie Marcel Didier.

Jusczyk, P. W., Cutler, A., & Redanz, N. (1993). Preference for the predominant stress patterns of English words. *Child Development, 64,* 675–687.

Jusczyk, P. W., Friederici, A. D., Wessels, J., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' recognition of foreign versus native language words. *Journal of Memory and Language, 32,* 402–420.

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language, 33,* 630–645.

Kewley-Port, D. (1983). Time-varying features as correlates of place of articulation in stop consonants. *Journal of the Acoustical Society of America, 73,* 322–335.

Kuhl, P., Williams, K., Lacerda, F., Stevens, K., & Lindblom, B. (1992). Linguistic experiences alter phonetic perception in infants by 6 months of age. *Science, 255,* 606–608.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages.* Oxford, England: Blackwell.

Lahiri, A., Gewirth, L., & Blumstein, S. E. (1984). A reconsideration of acoustic invariance for place of articulation in stop consonants: Evidence from a cross-language study. *Journal of the Acoustical Society of America, 76,* 391–404.

Maddieson, I. (1984). *Patterns of sounds.* Cambridge, England: Cambridge University Press.

Mann, V. A. (1980). Influence of preceding liquid on stop-consonant perception. *Perception & Psychophysics, 28,* 407–412.

Mann, V. A. (1986). Distinguishing universal and language-dependent levels of speech perception: Evidence from Japanese listeners' perception of English "l" and "r." *Cognition, 24,* 169–196.

Marslen-Wilson, W. D. (1984). Function and process in spoken-word recognition. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 125–130). Hillsdale, NJ: Erlbaum.

Marslen-Wilson, W. D., & Warren, P. (1994). Levels of perceptual representations and process in lexical access: Words, phonemes, and features. *Psychological Review, 101,* 653–675.

Massaro, D., & Cohen, M. (1983). Phonological context in speech perception. *Perception & Psychophysics, 34,* 338–348.

Massaro, D., & Cohen, M. (1993). The paradigm and the fuzzy logical model of perception are alive and well. *Journal of Experimental Psychology: General, 122,* 115–124.

Massaro, D., & Oden, G. (1980). Speech perception: A framework for research and theory. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 3, pp. 129–165). New York: Academic Press.

Massaro, D., & Oden, G. (1995). Independence of lexical context and phonological information in speech perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1053–1064.

McClelland, J., & Elman, J. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1–86.

Mehler, J., Dommergues, J. Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior, 20,* 298–305.

Messer, S. (1967). Implicit phonology in children. *Journal of Verbal Learning and Verbal Behavior, 6,* 609–613.

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America, 27,* 338–352.

Niederer, B. (1995). *Les langues Hmong-Mjen (Miao-Yao): Problèmes de phonologie historique* [Hmong-Mjen (Miao-Yao) languages: Issues of diachronical phonology]. Unpublished doctoral dissertation, Paris III University.

Oden, G., & Massaro, D. (1978). Integration of featural information in speech perception. *Psychological Review, 85,* 172–191.

Pinker, S. (1994). *The language instinct (how the mind creates language).* New York: Morrow.

Pitt, M., & Samuel, A. (1995). Lexical and sublexical feedback in auditory word recognition. *Cognitive Psychology, 29,* 149–188.

Polivanov, E. (1931). La perception des sons d'une langue étrangère. *Travaux du Cercle Linguistique de Prague, 4,* 79–96.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 421–435.

Praamstra, P., & Levelt, W. J. M. (1994). Neurophysiological manifestations of phonological processing: Latency variation of a negative ERP component timelocked to phonological mismatch. *Journal of Cognitive Neuroscience, 6,* 204–219.

Pullum, G., & Ladusaw, W. (1986). *Phonetic symbol guide.* Chicago: University of Chicago Press.

Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America, 64,* 1358–1368.

Streeter, L. A., & Nigro, G. N. (1979). The role of medial consonant transitions in word perception. *Journal of the Acoustical Society of America, 65,* 1533–1541.

Trubetzkoy, N. S. (1939). Grundzüge der Phonologie [Principles of phonology]. *Travaux du Cercle Linguistique de Prague, 7,* 1–271.

Tubach, J. P., & Boë, L.-J. (1990). *Un corpus de transcription phonétique (300,000 phones): Constitution et exploitation statistique* (Paris France Télécom Rep. No. 90 D 002). Paris: France Télécom.

Vaillant, A. (1950). *Grammaire comparée des langues Slaves: Tome 1: Phonétique* [Comparative grammar of Slavic languages: Vol. 1. Phonetics]. Paris: IAC, Klincksiek.

Walley, A. C., Michela, V. L., & Wood, D. R. (1995). The gating paradigm: Effects of presentation format on spoken word recognition by children and adults. *Perception & Psychophysics, 57,* 343–351.

Wang, F. S. (1979, September). *Miao-yu fangyan shengyunmu bijiao* [A comparative study of the onsets and rimes of Miao languages]. Paper circulated at the 10th Sino-Tibetan Conference, Paris.

Warren, P., & Marslen-Wilson, W. D. (1987). Continuous uptake of acoustic cues in spoken word recognition. *Perception & Psychophysics, 41,* 262–275.

Werker, J., & Tees, R. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7,* 49–63.

Whalen, D. H. (1984). Subcategorical phonetic mismatches slow phonetic judgments. *Perception & Psychophysics, 35,* 49–64.

Whalen, D. H. (1991). Subcategorical phonetic mismatches and lexical access. *Perception & Psychophysics, 50,* 351–360.

Received July 29, 1996
Revision received December 20, 1996
Accepted January 31, 1997 ■