HAL
archives-ouvertes.fr

Dental-to-velar perceptual assimilation:A cross-linguistic study of the perception of dental stop+/l/ clusters

Pierre Hallé, Catherine Best

▶ To cite this version:

Pierre Hallé, Catherine Best. Dental-to-velar perceptual assimilation:A cross-linguistic study of the perception of dental stop+/l/ clusters. pre-print: à paraître dans Journal of the Acoustical Society of Americaapril 2007, .. 2007. <halshs-00129735>

HAL Id: halshs-00129735
https://halshs.archives-ouvertes.fr/halshs-00129735

Submitted on 8 Feb 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Dental-to-velar perceptual assimilation:

A cross-linguistic study of the perception of dental stop+/l/ clusters

Pierre A. Hallé [a)]

Laboratoire de Phonétique et Phonologie, 19 rue des Bernardins, 75005 Paris,

Laboratoire de Psychologie et Neurosciences Cognitives, CNRS-Paris 5,

and Haskins Laboratories

and

Catherine T. Best [b)]

MARCS Auditory Laboratories, University of Western Sydney,

and Haskins Laboratories

Final version: December 2006

Running title: *Dental-to-velar perceptual assimilation*

[a)] Electronic mail: pierre.halle@univ-paris5.fr

[b)] Electronic mail: c.best@uws.edu.au

ABSTRACT

French listeners perceive illegal /tl/ and /dl/ clusters as legal /kl/ and /gl/, suggesting that /dl, tl/ undergo "phonotactic perceptual assimilation" to the phonetically most similar permissible clusters (Hallé, Segui, Frauenfelder, and Meunier, 1998). However, without a comparison to native speakers of a language allowing initial /tl, dl/, other explanations remain open (e.g., universal phonetic biases). Experiment 1 compared native French and Hebrew listeners on perception of Hebrew /tl/-/kl/ and /dl/-/gl/. On a language-specific phonotactics account, these contrasts should be difficult for listeners whose language disallows initial /tl, dl/ while allowing /kl, gl/ (French), but not for listeners whose language permits all four clusters (Hebrew). Indeed, French but not Hebrew listeners showed difficulty discriminating /tl/-/kl/, and tended to categorize the initial consonant of /tl/ as /k/; analogous effects for /dl/-/gl/ were weaker. Experiment 2 tested speakers of American English, which also disallows initial /tl, dl/ but realizes stop-voicing differently than French or Hebrew, to examine possible contributions of language-specific phonetic settings. Their performance was similar to that of French listeners, though they had significantly greater difficulty with /dl-/gl/. The results support the proposal of language-specific phonotactic perceptual assimilation, with modest contributions from language-specific phonetic settings.

PACS numbers: 43.71.Hw [ARB]

I. INTRODUCTION

In the early thirties, Polivanov (1931) suggested that people's native language phonological system must bias the way they perceive foreign sounds. Thereafter, metaphors such as "phonological deafness" and "phonological filter" coined in the 1930s (cf. Trubetzkoy, 1939) have been widely used and tacitly accepted as correct characterizations of how listeners perceive nonnative sounds. They reflect Polivanov's (1931) proposition that the native phonological system molds perception of nonnative speech patterns to follow its rules.

In that proposition, Polivanov included the phoneme "grouping laws" (ibid., p. 80) that govern permissible sound sequences. For instance, he noted the case of Japanese listeners perceiving the word "drama" as *dorama*, or alternatively as *zurama*, illustrating how the specific phonotactics of a language –here, the general ban against consonant clusters in Japanese– may constrain the way nonnative sequences are perceived. This example can be viewed as an early illustration of what was later called "phonological repair" in both loanword phonology and cross-language speech perception research.[1] The subsequent findings in both lines of research are largely consistent with Polivanov's general intuition and informal observations. Indeed, it seems that the knowledge of which speech sounds and sound combinations occur in real-life utterances does guide and facilitate speech perception, at the predictable expense of precision in the phonetic analysis of ill-formed utterances.

There still are a number of unresolved questions about *perceptual* phonological repairs (as opposed to repairs in *produced* loanword adaptations). The first question concerns the relative roles of low (phonetic/acoustic) and higher (phonological) levels of speech analysis. This question is the subject of an ongoing debate in loanword phonology between proponents of exclusively phonological motivations for loanword adaptations (Paradis and LaCharité, 1997, 2001) and proponents of the "phonetic approximation" view (see, among others, Shinohara, 2006; Silverman, 1992; Vendelin and Peperkamp, 2004; Yip, 1993). In the

broader field of speech perception, the phonological versus phonetic tension is directly related to the issue of language-specificity versus universality, respectively. Are language-specific phonological rules and constraints so firmly imprinted in listeners' knowledge of the native system that they supersede physical evidence? Are certain physical characteristics of speech sounds so salient that potential repairs are universally ruled out?

Another issue is that of the repair of ill-formed structures above the segmental level, for sequences of segments. It has often been discussed in the loanword phonology literature (e.g., Paradis and LaCharité, 1997) but has seldom been addressed in the speech perception literature, with the notable exception of the well documented case of the epenthetic "illusory" vowel heard by Japanese listeners presented with consonant clusters. Dupoux and colleagues examined the perception of utterance-medial consonant clusters by Japanese vs. French listeners (Dehaene-Lambertz, Dupoux, and Gout, 2000; Dupoux, Kakehi, Hirose, Pallier, and Mehler, 1999; Dupoux, Pallier, Kakehi, and Mehler, 2001). They found that Japanese, but not French listeners, hear the same epenthetic /u/ in nonwords such as *ebza*, perceived as /ebuza/, and in pseudowords such as *sokdo* or *mikdo* (from *sokudo* 'speed' or *mikado* 'emperor'). This is precisely the type of phonological repair Polivanov had observed Japanese listeners to apply to words such as *drama*.[2] In several studies, Dupoux and colleagues used systematic cross-language comparisons to reveal the language-specific nature of the repair. Their findings show that language-specific phonological constraints can supersede physical evidence. Presence/absence of a vowel is prima facie a large phonetic difference, yet because of their phonological system, Japanese listeners perceive a vowel that is physically absent. This failure to perceive the large phonetic difference between presence/absence of a vowel is all the more remarkable in that vowel epenthesis introduces a change in syllabic structure, from the perspective of languages which allow word-medial consonant clusters.

Hallé, Segui, Frauenfelder, and Meunier (1998) reported another case of perceptual

phonological repair associated with native phonotactic constraints but involving consonant substitution rather than vowel insertion. In this case, the phonotactic repair leaves syllabic structure unchanged. The reported data showed that French listeners strongly tend to perceive utterance-initial /tl/ and /dl/, which are not permissible in French, respectively as /kl/ and /gl/, which are permissible. This "dental-to-velar shift" effect (in Hallé et al.'s formulation) was found to be stronger for /tl/ than /dl/. It was demonstrated using mainly identification tasks on nonwords such as *dlopta* or *tlabod*, pronounced by a French speaker, as compared to nonwords such as *dropta* or *trabod*, whose word-initial /dr, tr/ clusters are permissible in French. The latter nonwords did not induce any dental-to-velar shift. However, this study used a within-language rather than a cross-language comparison, leaving unresolved the question of whether the repair is language-specific rather than universal. Dupoux et al. (2001) commented that the dental-to-velar shift in Hallé et al. (1998) could be "due to universal effects of compensation for coarticulation" and that "it could be that /dl/ is universally harder to perceive than /gl/." In a nutshell, the compensation for coarticulation argument holds that listeners expect dorsal stops to be fronted in the context of the coronal contact for /l/ (Mann, 1980; Mann and Repp, 1981). Thus, a stop with an articulatory and hence acoustic fronted quality can be heard as a fronted dorsal stop in /l/ context, resulting in a dental-to-velar perceptual shift. To evaluate whether the /dl, tl/→/gl, kl/ perceptual shift reflects such universal perceptual processing or, rather, a language-specific phonotactic repair, a cross-language comparison is needed. If the effect is language-specific, the dental-to-velar phonotactic perceptual repair should occur only in listeners whose language disallows dental stop + /l/ clusters, and not in listeners whose language permits these clusters.

It is the need for a cross-language examination of the dental-to-velar phonotactic repair that is the primary motivation for the present study. A comparison language that permits /dl, tl/ is needed to determine whether the repair is specific to the phonotactic

constraints of the listener's language, or rather is associated with universal perceptual processes (e.g., compensation for coarticulation). French listeners' tendency to identify /dl/ as /gl/ and /tl/ as /kl/ clearly reflects perceptual confusability between /tl/ and /kl/ and between /dl/ and /gl/; this should entail poor discrimination of /dl/-/gl/ and /tl/-/kl/ contrasts. We therefore evaluated discrimination of these clusters by French listeners. For the critical cross-language examinations, the comparison language had to allow /gl, kl/ as well as /dl, tl/ onset clusters. If language-specific phonotactic repair is responsible for the French perceptual bias, then, the native listeners of the comparison language should have little or no difficulty discriminating /dl/-/gl/ and /tl/-/kl/. But if universal perceptual constraints are instead responsible, a possibility suggested in Dupoux et al. (2001), then even the listeners of the comparison language should display substantial difficulty perceiving these contrasts.

Only a few languages have coronal-dorsal contrasts such as /tl/-/kl/, which, in itself, might reflect a universal trend toward avoiding these contrasts. One language that does have these contrasts is Modern Hebrew. It allows all possible stop+liquid clusters word-initially (Rosen, 1962; Téné, 1972), including /dl, tl/ and /gl, kl/, as well as /dr, tr/ and /gr, kr/. The voiced stops of Hebrew have substantial voicing lead (Laufer, 1998; Raphael, Tobin, Faber, Most, Kollia, and Milstein, 1995). The voiceless stops have been reported as intermediate between phonetically short lag unaspirated and long lag aspirated (Obler, 1982; Raphael, Tobin, and Most, 1983; Raphael et al., 1995). The phonological voicing distinction for stops is thus similar though not identical in Hebrew and in French, which contrasts pre-voiced and short lag unaspirated stops (see, for example, Nearey and Rochet, 1994). French and Hebrew have phonetically similar "light" /l/s (French: Chafcouloff, 1979; Simon, 1967; Hebrew: Chayen, 1973). Their /r/s have also been described as similar in place (uvular), although they might differ somewhat in manner of articulation: approximant or fricative /r/s in the French of our Parisian listener group (cf. Hallé, Best, and Levitt, 1999) vs. approximant or trilled /r/s in

Hebrew[3] (Devens, 1978, 1980; Laufer, 1990; Rosen, 1962). Therefore, in Experiment 1, we compared native French and Hebrew listeners on their discrimination of the Hebrew dental-velar contrast in the context of a following /l/ or /r/; Experiment 2 extended these cross-language comparisons to American English, which also disallows /dl/ and /tl/ but differs from French in the phonetic realizations of its coronal stops and its liquids. Would these phonetic differences have substantive impact on American, relative to French, listeners' perception of /dl, tl/? Given that all stop+/r/ clusters are permissible in the three languages (French, English, Hebrew), the /r/ context was used for comparison with Hallé et al.'s (1998) study, in which stop+/r/ clusters served as a within language baseline comparison.

While the cross-language comparison is the main motivation for this study, there are additional reasons to reconsider the earlier findings. First, there may be reason for concern about the phonetic and/or articulatory quality of the illegal /dl, tl/ speech tokens used in the earlier study. They were produced by a native speaker of French, who had no familiarity with any language allowing /dl, tl/ word-initially and had no training in phonetics. The critical stimuli thus could simply have been mispronounced as velar (or ambiguous) stop+/l/. Hallé et al. addressed this concern in two ways. They compared the acoustic properties of the stop release bursts of the /dl, tl/ versus /gl, kl/ clusters. Spectral cues to place of articulation were found to be more dental-like in the intended dentals than in the intended velars, in agreement with classic measures of the spectral properties of stop bursts (Halle, Hugues, and Radley, 1957; Kewley-Port, 1983; Stevens and Blumstein, 1978). These properties should provide sufficient information for reliable perception of place of articulation from bursts alone (Blumstein and Stevens, 1980). Hallé et al. then conducted a "phonetic gating" experiment and found that French listeners heard a dental rather than a velar stop in the shortest fragments, which corresponded to the stop bursts alone. It was thus argued that the initial stops of /dl, tl/ sufficiently met the requirements for dental place of articulation. It is possible,

however, that Hallé et al.'s (1998) French speaker produced ambiguous stops, intermediate between velar and dental, in the /dl, tl/ clusters. Indeed, assuming that speech perception and production are strongly linked (cf. Fowler, Brown, Sabadini, and Weihing, 2003, for an in-depth discussion), the dental-to-velar shift could conceivably occur in the French speakers' *productions* of /dl, tl/. These potential stimulus shortcomings are readily avoided by recording a native speaker of Hebrew, who must be able to produce differentiated velar- and dental-stop + /l/ clusters in word-initial position, because they are needed for contrasting minimal-pair words such as *tlulim–klulim* ('steep'–'included').

A second important concern is that of possible lexical influences. Although the nonwords in Hallé et al.'s (1998) study, such as *dlopta* or *tlabdo*, had been designed to bear little similarity with existing words, use of nonwords does not prevent lexical access processes from operating. The /dl, tl/ nonwords could conceivably activate words with phonetically similar onsets, such as /gl, kl/. Such lexical neighborhood activation was plausible given that the nonwords used were quite word-like in terms of syllabic pattern and phonotactic probabilities (except of course for the /dl, tl/ onsets), and were mixed with a wealth of very French-like nonword fillers. The perceptual repair of /dl, tl/ does not appear to be *strictly* lexically driven because, although French words beginning with /pl, bl/+/a, o, ɔ/ (the vowel context used in Hallé et al., 1998) are as frequent as words beginning with /gl, kl/+/ a, o, ɔ /, according to the "LEXIQUE" database (New, Pallier, Ferrand, and Matos, 2001) (/pl, bl/ words vs. /kl, gl/ words: 438 vs. 536 types; 1124 vs. 962 per million cumulative frequencies), there were exceedingly few dental-to-labial repairs. Thus, a strict lexical account of the /tl, dl/→/kl, gl/ perceptual repair does not work. However, the dental-to-velar shift could nevertheless be lexically *biased* to some extent. Specifically, the stronger dental-to-velar shift found for /tl/ than for /dl/ is compatible with the higher frequency of /k/- than /g/-initial words in the /l/+/a, o, ɔ/ context (/kl/ vs. /gl/ words: 369 vs. 167 types, 715 vs. 247

per million cumulated frequency). The cross-linguistic approach we employ here, using classic discrimination and categorization tasks with Hebrew stimuli, should help solve this puzzle. The Hebrew stimuli sound foreign rather than French-like, which should encourage French listeners to attend to their phonetic rather than lexical properties, thus minimizing the likelihood of lexical influences. To further minimize lexical effects, monosyllabic stimuli were used, because they are presumably less likely to suggest lexical items to French listeners than the multisyllabic stimuli of the earlier study (Hallé et al., 1998).

II. EXPERIMENT 1

We first examined the perception of Hebrew /dl, tl/ clusters by native listeners of French versus native speakers of Israeli Hebrew. If the dental-to-velar perceptual shift holds for French listeners presented with Hebrew stimuli, they should perceive Hebrew /dl, tl/ and /gl, kl/ as similar and thus have trouble discriminating the Hebrew /dl/-/gl/ and /tl/-/kl/ contrasts. The Hebrew /dr/-/gr/ and /tr/-/kr/ contrasts were used as control baseline contrasts to gauge French listeners' performance at perceiving Hebrew dental-velar contrasts in permissible, contrastive stop+liquid clusters, which French listeners should label and discriminate quite well. Hebrew listeners should encounter little difficulty with all the dental-velar stop+liquid cluster contrasts, whether with /r/ or with /l/, because they are all legal and contrastive in Hebrew. Yet, were the dental-to-velar perceptual shift for /dl, tl/ a universal tendency, Hebrew listeners could show difficulty discriminating /dl/-/gl/ and/or /tl/-/kl/.

A. Method

1. Participants

Twelve native French-speaking students at Paris V University (mean age 22, age range 19-25) and eleven native speakers of Hebrew recruited in Paris (mean age 23, age range 21-

29), participated in the experiment for a small participation payment. The data for one additional French participant were not retained because of failure to participate in the labeling part of the experiment. The data for one additional Hebrew participant were not retained because she had been residing in France for more than 20 years (the other participants were in Paris on a short stay) and was also well outside the age range of the other participants.

2. Stimuli

Twenty-four CCV monosyllabic items were constructed by crossing the vowels /a, i, u/ with the clusters /dl, tl, gl, kl, dr, tr, gr, kr/, which are all legal word-initially in Hebrew. The clusters were thus composed of a dental or a velar stop, voiced or voiceless, followed by /l/ or /r/. The three cardinal vowels /a, i, u/ were chosen so that the full extent of the Hebrew vowel space (/a, e, i, u, o/) would be represented. A randomized list, containing eight tokens of each item, was read aloud by a male native speaker of Hebrew who resides in Israel. The recording was made in an anechoic chamber, using a professional quality microphone and a DAT tape recorder. The speech materials were then digitized (16 kHz sampling rate, 16 bit resolution) and transferred to individual audio files. Two native speakers of Hebrew, students in phonetics, judged whether each stimulus had been correctly pronounced by checking them against a list of the intended pronunciations transcribed in Hebrew script. All the tokens were approved by both judges and therefore retained for selection as experimental stimuli.

For each of the 24 stimulus syllable types (4 stops x 2 liquids x 3 vowels), four tokens (out of eight repetitions) were selected for use in the perceptual experiments, for which the prosodic characteristics (syllable duration, F0 contour, and loudness) were balanced as well as possible within each set of items to be compared (e.g., the /gla/ versus /dla/ set). In order to enhance prosodic homogeneity, peak intensities in the vowel portion were equalized within and across these sets. The prosodic features of the selected materials are shown in Table I.

Insert Table I about here

The phonetic-acoustic characteristics of the clusters used, in particular of the initial stops, were critical to know for the velar-dental distinctions tested. Because few phonetic-acoustic data are available for Hebrew consonants and clusters, we acoustically analyzed the stimuli to be used in the planned perceptual tasks. In the stimuli produced by our speaker, the voiced stops had substantial pre-release voicing, whereas the voiceless stops had substantial post-release aspiration, thereby somewhat departing from the medium lag VOT that has been reported for Hebrew (e.g., Raphael et al., 1995). Thus, the voiced stops were pre-voiced while the voiceless stops were long-lag aspirated voiceless.

Three main characteristics distinguished velars and dentals. They are detailed in Table II. First, the spectral center of gravity (henceforth, SCG) in the release burst portion was lower overall for velars than for dentals by about 250 Hz, $F(1, 72) = 41.16$, $p < .00001$, but the difference was more marked for the /r/ than the /l/ context (333 Hz vs. 176 Hz), as suggested by a significant interaction between liquid context (/l/ vs. /r/) and place of articulation (dental vs. velar), $F(1, 72) = 3.91$, $p < .05$. There was some variability, especially with regard to vowel context and voicing, as can be seen in Table II: SCG for /dl, tl/ was much lower in the -/u/ than in the -/i/ or -/a/ contexts (3.2 kHz vs. 4 or 4.2 kHz) while SCG for /gl, kl/ was much less affected by the -/u/ context and remained in the 3.4-3.8 kHz range. As a result, SCG values for /dlu, tlu/ (3.2 kHz) are more typical of velar than dental place, whereas those for /dli, tli/ or /dla, tla/ (~4 kHz) are more typical of dental than velar place. This would predict more dental-velar shifts for /dlu, tlu/ than for /dli, tli/ or /dla, tla/, if SCG is an important cue in perception. For all other contexts, SCG was always lower for velars than dentals. Within the /l/ context, the difference was numerically, although not statistically larger for /dl/-/gl/ (250 Hz) than /tl/-/kl/ (100 Hz), $t(22) < 1$. Second, integrated energy in the

burst was greater for velars than for dentals (3.74 vs. 2.38 dB s), $F(1, 72) = 172.75$, $p < .00001$. This measure is interesting because perception of loudness depends on energy integrated over time, at least for rather short fragments of speech signal (Hughes, 1946; Scharf, 1978). It produced much smaller values for voiced than voiceless stops (1.53 vs. 4.59 dB s), $F(1, 72) = 875.50$, $p < .00001$, in keeping with the observation that voiced stops have quieter bursts than voiceless stops (Zue, 1976). Not surprisingly, then, the velar-dental differentials for this measure are smaller for voiced than voiceless stops (0.53 vs. 2.20 dB s), as suggested by the Place x Voice interaction, $F(1, 72) = 64.93$, $p < .00001$. Third, VOTs were significantly longer for velars than for dentals, consistent with the literature on VOT variation according to place of articulation (Fischer-Jørgensen, 1954; Lisker and Abramson, 1964; Nearey and Rochet, 1994; Peterson and Lehiste, 1960; Saerens, Serniclaes, and Beeckmans, 1989; see Cho and Ladefoged, 1999, for an overview). For voiceless stops, VOTs were longer for velars than dentals (95 vs. 55 ms on average) in both /l/ and /r/ contexts, $F(1, 36) = 205.5$, $p < .00001$. Voiced stops were all produced with voicing lead (i.e., negative VOTs). Voicing leads were significantly shorter for velars than dentals (93 vs. 112 ms on average), $F(1, 36) = 12.43$, $p < .005$.

Insert Table II about here

Dental clusters did not differ from velar clusters with respect to other measured acoustic cues. The duration of the liquid component was 115 vs. 113 ms for /tl/ vs. /kl/, 122 vs. 135 ms for /dl/ vs. /gl/, 114 ms for both /tr/ and /kr/, and 134 vs. 124 ms for /dr/ vs. /gr/ (all non-significant pairwise differences). The /r/s of the C+/r/ clusters were realized as uvular approximants or, more often, as uvular trills (cf. Devens, 1980), which occurred half of the time after voiced stops and virtually always after voiceless stops, regardless of place of articulation of the initial stop. The first three formants of the vowel following the liquid (/l/ or

/r/) were measured at 25%, 50%, and 75% of the vowel duration. The formant patterns were virtually identical in all dental-velar pairs (less than 50 Hz differences). They were conditioned only by the vowel itself and by the preceding liquid, not by the initial stop. F2 had a lower locus after /r/ than after /l/ (ca. 1 kHz vs. 1.9 kHz), consistent with the literature on French /l/ vs. /r/ (Chafcouloff, 1979). Analogously, the three first formants measured midway through the steady-state of the liquids /l/ and /r/ only depended on the liquid itself and the following vowel, not the place of articulation of the initial stop.

Lexical biases in French (Exp. 1) or American English (Exp. 2) listeners' perception are prima facie unlikely with stimuli that were produced by a native Hebrew speaker, and indeed sounded foreign to both French and American listeners. However, among the six /kl, gl/ syllables, three could induce a lexical bias in French: /klu/, /glu/, and /gla/ resemble a French word (*clou* 'nail', *glou* onomatopoetic, and *glas* '(bell) toll'); /klu/, /glu/, and /gli/ resemble an English word (*clue*, *glue*, and *glee*). These individual words could attract more dental-to-velar shifts in the matched /dl, tl/ syllables than in the others, and result, for example, in lower discrimination performance for /tlu/-/klu/, /dlu/-/glu/, and /dla/-/gla/ than other dental-velar contrasts by French listeners. Aside from individual words, lexical bias might be mediated by the cohorts of words beginning with /kl, gl/. Among these, in both French and English, words with /kla/ or /gla/ as initial syllable are by far the most frequent; those with /klu/ or /glu/ are the least frequent; those with /kli/ or /gli/ fall in between.[4] Thus, a cohort-mediated lexical bias would yield the poorest discrimination performance for the /a/ vowel context and the best one for /u/ in both French and English.

3. AXB discrimination task

The selected stimulus tokens were used to make up categorial AXB discrimination triads, in which A and B differed with respect to place (velar vs. dental) and X served as the

target item, to be judged as to its category match with A or B. It was always a categorial match, in that it was always a different token than the A or B item that it matched phonologically. For each of the four possible triad orders (AAB, ABB, BAA, and BBA), eight triplets were constructed (hence 32 combinations) in such a way that each token appeared equiprobably in each position and was not repeated within a given triplet. This yielded 384 triplets (32 combinations x 3 vowels x 2 voicings x 2 liquids), hence 384 test trials. The trials were presented in random order, blocked by 12 trials. The interstimulus interval was set to 1 s, the intertrial interval to 4 s, and the interblock interval to 8 s. Participants were allowed to pause midway during the test phase. The test phase was preceded by a training phase of ten AXB trials, which did not appear in the test phase. Three trials were 'easy' in that A and B differed in both initial consonant *and* liquid. No feedback was given to participants in either the training or the test phase.

For each trial, participants of both language groups were asked to indicate whether they perceived X as similar to A or to B by pressing one of two buttons labeled '1' and '3' (for A and B). They were instructed to respond on each trial, even if they had to guess, and to respond as fast as possible, as soon as they were confident of their response. There was no constraint that they wait until the third stimulus B. Hence, when listeners are quite confident, they may respond even before the final item of the triad, actually performing an AX-like speeded task in this case. We call this variant of the AXB discrimination procedure the "free RT" speeded AXB paradigm. The response times were measured from the release burst of the initial consonant of the third stimulus. Thus, early responses (i.e., occurring before the third item of the triad) have negative RTs. This procedure has proved useful to enhance RT differences between listener groups or between stimulus contrasts (Hallé, Chang, and Best, 2004). One potential drawback is that within-group RT variability may be large.

4. Categorization task

Only French participants had to categorize the Hebrew clusters in terms of French segmental categories. The task was not administered to Hebrew participants, given that all stimuli had been found to match intended pronunciation by two phoneticians, native speakers of Hebrew. In the test phase, each of the six /dl/ and /tl/ syllable types (2 clusters x 3 vowels) was used eight times (4 tokens x 2 repetitions), making 48 trials. The other 18 syllable types (6 clusters x 3 vowels) were used only four times (4 tokens), because they were presumably easier, making 72 trials. The test phase thus consisted of 120 trials. It was preceded by a training phase of 16 trials, which included two /dl/ and two /tl/ trials. No feedback was given to participants in either the training or the test phase.

For each trial, participants were first presented with a given syllable twice in a row. They were instructed to categorize its *initial consonant* by choosing one of ten consonants illustrated with the help of rhyming French keywords displayed on the screen (*paon*, *temps*, *Caen*, *banc*, *dent*, *gant*, *sang*, *Zan*, *rang*, and *lent*, consisting of an initial /p/, /t/, /k/, /b/, /d/, /g/, /s/, /z/, /r/, and /l/, respectively, followed by /ɑ̃/). Participants first made their choice by pressing an appropriate key. They then rated how well the syllable just presented (which they could hear again by pressing a key) matched the French keyword consonant they had indicated, using a 1-5 scale, in which 1 = "poor match" and 5 = "excellent match." The next trial was then administered until the test phase was completed.

B. Results

1. Discrimination

The performance of the French and Hebrew participants, pooled across vowels, is summarized in Figure 1A-B: correct discrimination and latencies for correct responses. Analyses of variance were conducted in two steps: first, only the structural factors of triad

Target (dental- vs. velar-initial X in AXB trials) and triad *Pattern* (XXY vs. XYY trials, usually referred to as primacy vs. recency trials), as well as the *Vowel* context factor, were examined separately for Hebrew and French participants. The factors of critical interest, cluster *Liquid* (/r/ vs. /l/) and cluster stop *Voicing* (voiced vs. voiceless), were examined in further analyses of variance which included listeners' *Language* (French vs. Hebrew) as a between-subject factor. Because the distributions of percent correct data often fail to meet the normality criterion, especially when they lie close to the upper boundary, we applied an arcsine transform to them in order to approximate more closely a normal distribution, and also ran analyses on the transformed data. The results of these analyses are reported in the following only in cases of discrepancy with the analyses on raw percentages.

Triad Pattern had no significant effect in any analysis and will not be discussed further. Triad Target generally had a significant though numerically small effect, with slightly better discrimination performance for velar targets.[5] The triad Target effect did not otherwise interact with the main patterns found in the data, so we will not discuss it further.

In both the Hebrew and French percent correct discrimination and RT data, the Vowel factor had no significant main effect or interaction effects, which is not surprising for Hebrew participants, given their near-ceiling performance. Although Vowel had no significant effect overall, the French data showed numerical trends for poorer and slower discrimination in the -/li/ context (/d/-/g/ and /t/-/k/: 75.6 and 60.4% correct; 1408 and 1873 ms) than the -/la, lu/ contexts (/d/-/g/ and /t/-/k/: 77.8 and 65.7% correct; 1034 and 1420 ms).

The language comparison analyses for the percent correct data show that the Language x Liquid interaction was highly significant, $F(1, 21) = 128.8$, $p < .00001$, reflecting that while French and Hebrew participants performed equally well for /dr/-/gr/ and /tr/-/kr/ contrasts (above 97%), Hebrew participants outperformed French participants for the critical /dl/-/gl/ and /tl/-/kl/ contrasts (96 vs. 71%), $F(1, 21) = 132.3$, $p < .00001$. French performance was

poorer for /tl/-/kl/ (64%) than for /dl/-/gl/ (77%), $F(1, 11) = 13.7$, $p < .005$. For Hebrew participants, the differences across contrasts were numerically tiny. However, performance was significantly lower for the /tl/-/kl/ contrast (94.7%) than for the other contrasts pooled (97.9% on average), $F(1, 10) = 15.2$, $p < .005$. This small difference remained significant in the arcsine-transformed data.

Insert Figure 1 about here

The RT data (Figure 2) paralleled the percent correct data: the higher the discrimination score, the shorter the RT in each language group. The Hebrew listeners responded much more quickly overall than the French, $F(1, 21) = 58.50$, $p < .00001$. Indeed, for the Hebrew listeners, the average response time, measured from the onset of the third stimulus, was negative (-916 ms), indicating that they actually responded without attending to the third stimulus. In contrast, the average RT for French participants was 977 ms, suggesting that, most of the time, they responded after they heard the third stimulus. In spite of this qualitative difference, a robust negative correlation between RT and percent correct discrimination was found in both groups, showing that the association of "free" RT and percent correct data provides a coherent picture of discrimination difficulty. The correlations were computed on 12 pairs of data points: the dependent variable values (percent correct and RT) averaged across participants for each segmental contrast (4 onset-cluster contrasts x 3 vowels). For French participants, the correlation between percent correct score and RT was $r(10) = -0.96$, $p < .00001$. For Hebrew participants, although the range of variation in percent correct scores was very narrow (93% to 99.7%), the correlation was nevertheless surprisingly robust, $r(10) = -0.79$, $p < .005$. The main trends in the percent correct data were thus found, in reverse, in the RT data: French participants were slower for the /l/ than for the /r/ cluster contrasts (1365 ms vs. 590 ms), $F(1,11) = 27.0$, $p < .0005$, and slower for the /tl/-/kl/ than for

the /dl/-/gl/ contrast (1571 ms vs. 1158 ms), $F(1,11) = 17.5$, $p < .005$. For Hebrew listeners, the /tl/-/kl/ contrast yielded the slowest responses (by about 200 ms), $F(1,10) = 22.4$, $p < .001$, confirming the numerically small but reliable difficulty they encountered with this contrast.

2. Categorization

For all the stimuli, including /dl/ and /tl/, the rate of voicing confusion was negligible (below 1%). For the clusters that are legal in French (/gl, kl, gr, kr, dr, tr/), place confusions were rare (3.8% and 1.7% velar-to-dental for /gl, kl/ and /gr, kr/, 1.7% dental-to-labial confusions for /dr, tr/). For /dl, tl/, French participants reported mainly velar or dental responses; they seldom indicated labial responses (none for /dl/ items and 6% for /tl/ items). An analysis of variance was run with *Liquid* context (/l, r/), *Vowel* context (/a, i, u/), *Place* of articulation (dental, velar) and *Voicing* (voiced, voiceless) of stimulus initial stop as within subject factors. Percentage of "correct responses" was the dependent variable. A response was scored as correct if it was a stop with the stimulus' intended place of articulation, regardless of whether reported voicing was correct. Note that correct dental responses to /dl, tl/ items may correspond to either "faithful" perception of /dl, tl/ as a whole, or to perceptual repairs such as /dr, tr/, /t, d/, or /dəl, təl/. We treat all these as cases of faithful perception of the initial consonant. The Liquid x Place interaction was highly significant, $F(1, 11) = 114.12$, $p < .00001$, reflecting nearly perfect performance for *all* the /r/ clusters (98.3%), regardless of Place but very different performance for the /kl, gl/ and /tl, dl/ clusters (96.2% vs. 42%), $F(1, 11) = 86.44$, $p < .00001$. In other words, perceptual shifts in place of articulation almost exclusively occurred for the /dl/ and /tl/ cluster stimuli. We therefore focused on the dental and velar place responses that were given to the /dl/ and /tl/ stimuli.

The raw percentages of dental, velar, and labial responses, according to cluster, are shown in Figure 2A. Following Guion et al. (2000), we also computed the "fit index" for

each subject and each response type as the proportion of responses of a given type multiplied by the corresponding mean rating on a 1-5 scale. This produces fit indexes in the 0-5 range. The fit index data (Figure 2B) yielded essentially the same patterns as the raw data. Analyses of variance were run on the raw data (F_{raw}) and the fit index data (F_{fit}), with Voicing (/dl/ or /tl/) and Vowel (/a, i, u/) as within-subject factors. Response (dental or velar) was treated as a repeated measures factor. The Voicing x Response interaction was significant, (ps < .0001) reflecting a strong asymmetry in the perception of the /dl/ and /tl/. In /dl/, the initial consonant was less frequently judged to be velar than in /tl/, ps < .005, as can be seen in Figure 2. This asymmetry is consistent with the better French performance in discriminating the /dl/-/gl/ contrast than the /tl/-/kl/ contrast. The discrimination performance correlated negatively with the rate of velar responses for the critical /dl, tl/ clusters, $r(22) = -0.79$, $p < .00001$ (Figure 3: one data point per subject for each contrast type): the poorer the discrimination, the higher the velar response rate to the /dl, tl/ cluster involved.

Insert Figures 2 and 3 about here

Variation across vowel context was significant within the /dl/ and /tl/ stimuli, $F_{raw}(2, 22) = 7.10$, $p < .005$; $F_{fit}(2, 22) = 4.80$, $p < .05$. There were fewer velar responses with /a/ than with /i, u/, $F_{raw}(1, 11) = 8.73$, $p < .05$; $F_{fit}(1, 11) = 6.28$, $p < .05$ (/dl/ stimuli: 21%, vs. 29% and 38% velar responses with /a/, vs. /i, u/; /tl/ stimuli: 67% vs. 94% and 83%).

C. Discussion

Experiment 1 confirmed the robustness of the French listeners' dental-to-velar perceptual repair of utterance-initial /dl, tl/. The repair extends to nonnative speech items beginning with /dl/ or /tl/ for French listeners, but does not extend to native listeners of the target language, Hebrew, in which initial /dl/ and /tl/ are permissible. Two important

concerns about Hallé et al.'s (1998) interpretation of the dental-to-velar shift as a case of "contextual perceptual assimilation" (ibid., p. 604) have been resolved by the current study.

First, the phenomenon is not simply attributable to the intrinsic acoustic properties of the speech stimuli of Hallé et al., which could have reflected the French speaker's pronunciation difficulty in producing /dl/ and /tl/ targets. Monosyllabic items such as /tla/ or /dli/ produced by a native speaker of Hebrew, which allows word-initial /dl/ and /tl/, gave rise in the present experiment to a very clear pattern of dental-to-velar perceptual shift by French but not Hebrew listeners.

Second, the use of monosyllables and the cross-linguistic design, involving stimuli that have a clearly non-French accent, appear to have successfully minimized the likelihood of lexical influences (i.e., feedback) on the French listeners' performance. For example, the observed effects of vowel context were indeed inconsistent with lexical bias accounts of French listeners' performance. French listeners tended to have more difficulty discriminating /dl/-/gl/ and /tl/-/kl/ in the /i/ than in the /a/ or /u/ contexts and, in the categorization test, produced significantly more velar responses to /dl, tl/ in the /i/ or /u/ than in the /a/ context. This pattern indicates that dentals are the most confusable with velars to a French ear when they occur in the -/li/ context, and the least confusable in the -/la/ context. In the French lexicon, words beginning with /gli/ or /kli/ are outnumbered by those with /gla/ or /kla/ in terms of both types and tokens (see footnote 4) and therefore constitute smaller lexical cohorts. Thus, a cohort-based lexical bias would predict more dental-velar confusability for /dl/-/gl/ and /tl/-/kl/ in the /a/ than in the /i/ context, a pattern which is exactly opposite to that observed. As for a lexical bias induced by individual words, only "clou," "glas," and "glou" could bias French listeners to hear a velar stop in /tlu/, /dla/, and /dlu/, predicting less dental-velar confusability in the /i/ than in the other vowel contexts, and more confusability in the /u/ context. Again, the variation in French performance according to vowel context, in both

discrimination and categorization, is not consistent with the predictions based on such lexical bias. Nor is it explainable by acoustic differences: SCG differences would predict -/lu/ to be the most difficult context (according to Table II, dentals are acoustically more similar to velars in this context than they are to dentals in other contexts). Nor should dental-velar shifts be the most frequent in the -/li/ context, based on acoustic differences. One possible account of the stronger velar shift with /tli, dli/, however, is compensation for coarticulation, given that high-front /i/ presumably induces more anticipatory fronting than the other vowels.

To sum up, the dental-to-velar perceptual repair effect for /dl, tl/ is more likely induced by phonological factors (i.e., language-specific phonotactics) than by either low-level acoustic properties, coarticulatory properties, or lexical influences. If this is correct, the effect should be found for other languages which, like French, disallow /dl, tl/ word-initially. Yet, the asymmetry between /dl/ and /tl/ found in the previous studies on French, as well as in the discrimination data and even more dramatically in the categorization data of the present study, does not fit well with an exclusively phonotactic motivation. It suggests that the basic effect of phonotactic repair is modulated by factors other than a cluster's illegal status.

Those other factors could be phonetic. There is a better phonetic match between the French and Hebrew stops for /d/ than for /t/. The voiced stops of both Hebrew and French have clear pre-release voicing lead, but the Hebrew voiceless stops we used had a long lag aspirated VOT, unlike the typically short lag unaspirated voiceless stops in French. If the dental-to-velar perceptual repair effect is modulated by such language-specific phonetic factors in stop voicing settings, rather than, for example, structural factors at a more abstract level (cf. Moreton, 2002), then native listeners of English should 1) show the dental-to-velar shift (i.e., phonotactic repair at the phonological level) but 2) should exhibit a voiced-voiceless /dl, tl/ asymmetry opposite to that in French listeners (i.e., language-specific phonetic modulation of the effect). This is because there is a better phonetic match between

the English and Hebrew stops for /t/ than for /d/, contrary to the French case. The voicing distinction in English stops relies on the VOT difference between short lag and long lag, as there is not systematic pre-voicing for English voiced stops (cf. Lisker and Abramson, 1964). We need, however, to first establish whether English listeners also perceive utterance-initial /dl, tl/ as velar, that is, display phonotactic repair at all.

In Experiment 2, we tested native listeners of English on their perception of Hebrew /tl, dl/. The first purpose was to establish whether English listeners display the same kind of perceptual shift as French listeners. If yes, this would strengthen the language-specific phonotactic repair account of this shift. To this date, indeed, *direct* evidence for a perceptual repair of /dl/ or /tl/ in languages that disallow them has been reported only for French. For English listeners, there is indirect evidence for a perceptual repair, showing that the location of the categorical boundary in /r/-/l/ continua is biased by a preceding consonant context towards legal obstruant-liquid responses (Massaro and Cohen, 1983; Pitt, 1998). If, as this indirect evidence suggests, English listeners also display a perceptual shift for /dl, tl/, a second purpose was to examine whether /dl/ and /tl/ are affected asymmetrically by the shift and if yes, whether the asymmetry is congruent with the English phonetic settings of the voicing distinction. Differences between French and English listeners in the perceptual asymmetry between /dl/ and /tl/, related to language differences in stop voicing settings, would provide evidence of language-specific *phonetic* modulation of the dental-to-velar perceptual repair.

III. EXPERIMENT 2

In this experiment, native speakers of American English, who lacked contact with any language allowing /dl, tl/ word-initially, were run on the same tasks and materials as the French participants in Experiment 1.

A. Method

1. Participants

Fourteen students at Wesleyan University (mean age 19 years, age range 18-22) participated in the experiment for course credit. All of them lacked exposure to Hebrew, or to any language allowing /dl, tl/ word-initially, and lacked experience with French. None of them reported any speaking, hearing, or reading problems. Five additional Americans were tested; four failed to complete all tests and one completed the tasks but was an outlier with respect to miss rate (3.4% vs. an average 0.25% for the retained subjects).

2. Stimuli, design, and procedures

Stimuli were the same as in Experiment 1. The categorization test keywords were *pang*, *tang*, *kang*, *bang*, *dang*, *gang*, *sang*, *zang*, *rang*, and *lang*. All were words except *kang*, *lang*, and *zang*, which, however, all are possible and easy to pronounce surnames. They were modeled after the French keywords in Experiment 1: an initial /p, t, k, b, d, g, s, z, r, l/ followed by /æŋ/, an English rime that is phonologically similar to French /ɑ̃/.

B. Results

1. Discrimination

As can be seen in Figure 4, the discrimination performance of the American participants was very similar to that of French participants. They performed close to ceiling level for the /dr/-/gr/ and /tr/-/kr/ contrasts, and much more poorly for the contrasts involving the illegal clusters /dl/ or /tl/ ($p < .00001$). They thus differed from Hebrew listeners in much the same way as French did. Analyses of variance similar to those run on the discrimination data in Experiment 1 were conducted, examining first the possible effect of the structural factors triad Pattern (primacy vs. recency) and triad Target (dental vs. velar) and Vowel

context (/a, i, u/). As in Experiment 1, triad Pattern did not approach significance and will not be discussed further. Triad Target was significant – for the /l/ cluster contrasts only – with better performance for velar than dental targets (67% vs. 59.5% correct), $F(1, 13) = 7.85$, $p < .05$. Vowel was not significant overall. However, Vowel and Liquid context interacted significantly, $F(2, 26) = 18.25$, $p < .0001$, reflecting inconsistent vowel effects for the /r/ cluster 'baseline' contrasts and increasingly better discrimination performance from /i/ to /u/ to /a/ for the /l/ cluster contrasts (59%, 63%, and 68% correct, respectively): performance was the lowest for the /li/ context, $F(1, 13) = 7.10$, $p < .05$ (/li/ vs. /lu, la/), and the highest for the /la/ context, $F(1, 13) = 17.38$, $p < .005$ (/la/ vs. /lu, li/), just like in the French data. This pattern of performance is not consistent with the acoustic differences across vowels or with lexical influences for the same reasons as in the case of the French data. Again, as with the French listeners, the Americans' greater difficulty in the /i/ context could only be consistent with a compensation for coarticulation account.

Language comparison analyses were then run on the percent correct and RT data of all three listener groups, with Liquid context (/l/ vs. /r/) and cluster stop Voicing (voiced vs. voiceless) as within-subject factors and listener Language (American, French, and Hebrew) as a between-subject factor. American participants showed a pattern of results similar to that of French listeners but performed less well overall, $F(1, 24) = 8.25$, $p < .01$. Because Hebrew participants outperformed French participants for all contrasts, they obviously outperformed American participants as well.

We therefore focus on the French-American comparisons in the following. American participants performed less well than French participants on the /r/ clusters, $F(1, 24) = 12.19$, $p< .005$. They also performed less well on the /l/ clusters but the difference was significant only for /dl/-/gl/ (66% vs. 77%), $F(1, 24) = 6.41$, $p< .05$, not for /tl/-/kl/ (61% vs. 64%), $F(1, 24) = 1.01$, $p = .33$. Whereas the French data showed a clear voicing asymmetry, with better

performance on /dl/-/gl/ than /tl/-/kl/ (77% vs. 64%, $p < .005$), the asymmetry was not clear-cut in the American data (66% vs. 61%, $F(1,13) = 3.2$, $p = .092$), though also marginally in favor of the /dl/-/gl/ contrast. The weaker voicing asymmetry in the American than in the French data is reflected in a marginally significant Language x Voicing interaction for /l/ clusters alone, $F(1,24) = 3.10$, $p = .088$.

Insert Figure 4 about here

As in Experiment 1, the RT data paralleled the percent correct data. In the American data, RT negatively correlated with percent correct discrimination, $r(10) = -0.98$, $p < .00001$ (computed on 12 pairs of data points, as in Experiment 1). American participants were much slower for the /l/ than for the /r/ cluster contrasts (1167 ms vs. 576 ms), $F(1,13) = 38.73$, $p < .0001$. American participants, unlike French listeners, showed equivalent RTs for the /dl/-/gl/ and /tl/-/kl/ contrasts (1194 ms vs. 1139 ms), $F < 1$. Hence, both percent correct and RT data indicate that the clear voiced-voiceless asymmetry found in the French discrimination data for the /l/ clusters was virtually absent in the American data. RTs for the /tl/-/kl/ contrast were shorter overall for American than for French participants (1194 vs. 1571 ms) but the difference was statistically significant only after discarding two subjects (one French and one American) who had negative RTs, $t(22) = 2.41$, $p < .05$; this contrast thus might have been somewhat easier for American than for French listeners with respect to response times.

2. Categorization

American participants showed virtually no place confusions for the legal clusters (/gl, kl, gr, kr, dr, tr/), except for one /f/ response to a /gli/ trial. For all the stimuli, including the critical /dl/ and /tl/ stimuli, the rate of voicing confusion was negligible (below 0.2%). Like the French participants, the American participants seldom gave labial responses for /dl/ or /tl/

(0% for /tl/ items, 0.6% for /dl/ items). An analysis of variance similar to that run on the French categorization data (Experiment 1) was run on the American categorization data, with the percentage of correct responses (see Exp. 1) as the dependent variable. However, Liquid context (/l/ vs. /r/) could not be included as a factor because there was no variance for the subset of /r/ clusters (100% correct for each cluster and each participant). The mean percentage of correct responses was 99.7% for /gl, kl/ whereas it was only 36% on average for /dl, tl/, $F(1, 13) = 123.99$, $p < .00001$. We therefore focused on the dental and velar place responses that were given to the /dl/ and /tl/ stimuli, as in Experiment 1. The mean percentages of dental, velar, and labial responses according to cluster are shown in Figure 5A; the fit index data (Figure 5B) exhibited essentially the same patterns. The American and French data were entered into two analyses of variance, one for the percentage data and the other for the fit index data, in which Language (American vs. French) was a between-subject factor. As in Experiment 1, Voicing (/dl/ vs. /tl/) and Vowel (/a, i, u/) were within-subject factors, and Response (dental vs. velar) was treated as a repeated measures factor. The same voiced-voiceless asymmetry as in the French categorization data held for the American categorization results, reflected by a significant Voicing x Response interaction, ps $< .001$. The initial consonant was less consistently judged as velar in /dl/ than it was in /tl/, ps $< .01$. For the /dl/ stimuli, however, the advantage of dental over velar responses was less marked for American than French listeners, especially with respect to the fit index. The significant Language x Response interaction for the /dl/ stimuli in the fit index data, $F_{fit}(1, 24) = 5.34$, $p < .05$, reflected the smaller fit index for dental responses to /dl/ stimuli in the American than the French data (1.7 vs. 2.4), $F_{fit}(1, 24) = 10.09$, $p < .005$. (This interaction was not significant in the raw percentage data.) There was thus some evidence for a smaller voicing asymmetry in the American than in the French categorization data for dental stop+/l/ stimuli. In a correlation analysis of the American data, similar to that run in Experiment 1, the percent

correct score for /dl/-/gl/ and /tl/-/kl/ correlated negatively with the percentage of velar categorizations for /dl/ and /tl/: $r(26) = -0.62$, $p < .0005$ (Figure 6: one data point per subject for each contrast type, as in Figure 3). The correlation was numerically smaller than that found for French listeners ($r(22) = -0.79$) but not significantly so (*Fisher Z-transformed difference*: -1.12, $p = .26$).[6] That difference is nevertheless in line with the observation that discrimination data parallels identification data more closely for French than for American listeners: for American listeners, the voiced-voiceless asymmetry is significant in the identification data, not in the discrimination data; for French listeners, the asymmetry is found for both and is stronger than that of Americans for the fit index of the identification data.

Insert Figures 5 and 6 about here

The American data were also similar to the French data with respect to the Vowel effect, with a significant variation of velar judgments across vowel context, $F_{raw}(2, 26) = 21.08$, $p < .00001$; $F_{fit}(2, 26) = 5.44$, $p < .05$. Velar judgments were less frequent with /a/ than with /i/ or /u/ overall (/dl/: 23%, 36%, and 57%, respectively; /tl/: 73%, 92%, and 92%).[7] This runs counter to the lexical bias account that would predict more velar judgments in the /a/ context, for essentially the same reason as for French participants.

C. Discussion

The discrimination test showed that American participants, just like French participants, had difficulty distinguishing /tl/ from /kl/ or /dl/ from /gl/. Thus, the basic phonotactic repair effect extends to other languages for which initial /dl, tl/ is illegal. Nonetheless, there are differences from the French pattern. American participants performed less well than French participants on all contrasts, and significantly so, except for /tl/-/kl/. For instance, they performed slightly but significantly less well on /dr/-/gr/ and /tr/-/kr/ (95.5% vs.

97.5% correct). This could reflect a general perceptual difficulty induced by the Hebrew /r/ context, which is phonetically further away from English than French /r/. American English /r/ (bunched /r/ or retroflex /r/, cf. Boyce & Espy-Wilson, 1997; Zawadzky & Kuehn, 1980) is not uvular, as are both French and Hebrew /r/s (Delattre, 1969; Laufer, 1990). The same remark could apply to the lower performance overall of American participants for the /l/ contrasts: English /l/s tend to be "darker" ([ɫ]) than both French and Hebrew "light" /l/s.

American listeners performed more poorly than French participants on /dl/-/gl/ (66% vs. 77% correct discrimination) but not on /tl/-/kl/ (61% vs. 64%), on which they were actually faster than French listeners. In the categorization test, they showed the same voicing asymmetry as did the French listeners, categorizing dentals as velars less often for /dl/ than for /tl/, although this pattern was less marked than in the French data. This particular pattern of similarities and differences between the American and French asymmetries clearly runs counter to the hypothesis that the language-specific phonetic settings of stop voicing in French vs. English factors affect the /dl, tl/ → /gl, kl/ phonotactic repair effect. This hypothesis predicted that the voiced-voiceless asymmetry in the effect would go in opposite directions for English and French listeners. Although the lower discrimination performance of American than French listeners on the voiced contrast /dl/-/gl/ could reflect a real difference in the way French and American listeners perceive language-specific phonetic specifications for voiced stops, we found a similar asymmetry in the repair effect for both groups in the categorization data. The *direction* of asymmetry in phonotactic repair thus is not driven by the specific phonetic settings of stop voicing of the listeners' native language, although those details do appear to modulate its *magnitude*. Some other factor must be responsible for the stronger dental-to-velar shift for /tl/ than for /dl/, indeed even for Hebrew listeners to a very small but significant extent; we take up this issue in the General Discussion.

IV. GENERAL DISCUSSION

In Experiment 1, French listeners had substantial difficulty discriminating the Hebrew dental-velar stop + /l/ contrasts, especially /tl/-/kl/. Consistent with their discrimination performance, they often categorized Hebrew /tl/ and /dl/ as velar-initial, especially /tl/. They thus showed a strong dental-to-velar shift with Hebrew /tl/ and, to a lesser extent, /dl/. Their performance was contrary to any account in terms of lexical bias, suggesting that the dental-to-velar shift operates at a sub-lexical level of perception. Surprisingly, Hebrew listeners had some difficulty discriminating /tl/-/kl/, although they performed near ceiling. Experiment 2 extended the findings to native speakers of American English, who differed from French listeners in their equally low discrimination performance on /dl/-/gl/ and /tl/-/kl/, and showed somewhat less /dl, tl/ asymmetry in their categorization responses.

The results provide unequivocal answers to several key issues raised in the Introduction. First, using /dl, tl/ stimuli produced by a native speaker of Hebrew, hence stimuli that must convey a dental quality because both /dl/-/gl/ and /tl/-/kl/ are contrastive in Hebrew, did not prevent French and American listeners from experiencing a dental-to-velar perceptual shift, as in Hallé et al.'s (1998) study, which used stimuli produced by a native speaker of French. Second, the dental-to-velar shift was not modulated by lexical feedback for either the French or the American listeners, and likely operated at a sublexical level of speech perception. The results also provide a clear answer to the main issue that was raised: the dental-to-velar shift effect is language-specific rather than universal. Indeed, the effect is substantial in the two languages examined that disallow /dl/ and /tl/ in initial position, but in Hebrew, which allows initial /dl/ and /tl/, the effect is either absent altogether (/dl/-/gl/) or numerically tiny and still near-ceiling (/tl/-/kl/). A claim that the dental-to-velar shift is due to a *universal* perceptual bias toward hearing a dorsal stop in coronal stop + coronal liquid

clusters is thus untenable. For example, the universal compensation for coarticulation mechanism suggested by Dupoux et al. (2001) as a possible explanation of the dental-to-velar shift obviously fails to account for the large difference between Hebrew and French or American data. Yet, as the slight difficulty encountered by Hebrew listeners with /tl/-/kl/ may suggest, we should ask whether there remains some residue of experience-independent perceptual difficulty with word-initial coronal-dorsal stop + /l/ contrasts. To probe this issue, we turn to the question of universal tendencies across languages.

Universal tendencies

Diachronic and synchronic data indicate that /tl/-/kl/ and /dl/-/kl/ are difficult contrasts. Diachronically, coronal-dorsal contrasts in stop + /l/ clusters have often become neutralized in the languages in which they have existed (cf. Hallé et al., 1998, for a brief survey), suggesting intrinsic difficulty with these contrasts. Importantly, though, in those languages where a coronal-dorsal distinction in /l/ clusters is lacking or has been lost, it is often the coronal, not the dorsal clusters that were maintained. This holds for languages of quite different linguistic families (Hmong-Mien dialects, Setswana or Sesotho in the Bantu family, Aztec languages such as Nahuatl, etc.), suggesting that the difficulty does no lie in the low perceptibility of the coronal stop + /l/ clusters per se, but, rather, in the auditory similarity of the coronal and dorsal stop + /l/ clusters. It is also true that coronal stop + /l/ clusters are unstable diachronically and have a much lower incidence than dorsal stop + /l/ clusters in the languages of the world, a fact which would usually be interpreted as reflecting a structural constraint against adjacent coronal consonants: avoidance of homorganic CC sequences, especially when the two Cs are close in sonority (Padgett, 1991; Selkirk, 1988). Avoidance of coronal stop + /l/ clusters may alternatively be understood as resulting from the avoidance of clusters that are confusable with the readily acceptable dorsal stop + /l/ clusters.

Bradley (2006) proposed that coronal stop + /l/ clusters are intrinsically confusable with their dorsal counterparts, due to the articulatory adjustments necessary to produce them. Bradley reasoned that "In oral stops that are released into a following lateral, coronal and velar constrictions are produced at or behind the lateral constriction. As a result, these stops are not well differentiated by their release bursts…" Flemming (2002) proposes that clusters such as [tl] and [kl] share the same critical "auditory features" (F2 transition, burst diffuseness and frequency) and are thus poorly discernible, whereas they both differ from [pl] on all three features. Our acoustic measurements of the Hebrew dorsal and coronal stop + liquid clusters (Table II) also suggest that the bursts of /tl/ and /kl/ (or /dl/ and /gl/) are less well differentiated than those of /tr/ and /kr/ (or /dr/ and /gr/).

It is difficult, however, to conclude that the physical (or psychophysical) similarity between coronal and dorsal stops released into /l/ is sufficient to universally entail perceptual confusion. The much better discrimination performance of Hebrew listeners compared to French or American listeners instead suggests that a phonological grammar allowing /tl/-/kl/ (and possibly /dl/-/gl/) greatly enhances this otherwise poorly perceptible contrast: Hebrew listeners are tuned to attend to the subtle acoustic-articulatory differences between /tl/ and /kl/, however small these differences are. Because French and English both disallow /tl, dl/ while allowing /kl, gl/, the discrimination data of American and French listeners could provide measures of objective (i.e., unbiased by native phonological grammar) similarity between /tl/ and /kl/, and between /dl/ and /gl/. French and American listeners indeed exhibited roughly the same dental-to-velar perceptual shift for /dl, tl/. Yet, systematic differences in magnitude for the /dl, tl/ asymmetry, in both their discrimination and identification data, suggest their performance is influenced by language-specific aspects other than purely phonological ones.

Differences between French and English

The difference between the French and American data is not dramatic but, still, is robust. While the discrimination performance of French and American listeners was virtually the same on /tl/-/kl/, French listeners outperformed American listeners on /dl/-/gl/, and showed more marked /dl, tl/ asymmetry in their categorization data. Because French and English equally disallow /dl, tl/, these differences must be explained by some language-specific factors other than the phonological grammar alone.

One such factor is the language-specific phonetic settings for stop voicing. The better match of Hebrew /d/ ([d̪]) with French /d/ ([d̪]) than English /d/ ([t/d]) and, conversely, the better match of Hebrew /t/ ([t̪^{h}]) with English /t/ ([t^{h}]) than French /t/ ([t̪]), could have led to opposite patterns of /dl, tl/ asymmetry for French and American listeners. But this did not occur. The two groups of listeners differed in the magnitude of the /dl, tl/ asymmetry, not in its direction. Both groups exhibited more dental-to-velar shifts for /tl/ than for /dl/ in their identification data and, congruent with identification, lower discrimination performance for /tl/-/kl/ than /dl/-/gl/ (at least, numerically). The French data differed from the American data in that the advantage for /dl/-/gl/ over /tl/-/kl/ was substantial only in the French data, not in the American data. The pattern of results is perhaps best captured in the following way: both French and American listeners performed poorly on the /tl/-/kl/ contrast – slightly above 60% correct discrimination – but only French and not American listeners exhibited improved performance for /dl/-/gl/.

The reason why American listeners did not perform better on /dl/-/gl/ than /tl/-/kl/ could be their lack of consistent experience with voicing leads in phonetically voiced stops. They might be deaf to these voicing leads. French listeners, in contrast, are likely to perceive the voicing murmur of pre-voiced [d]s, which is the main cue to the voicing contrast in French stops. They thus could parse inputs such as murmur+[d]+[l]+[a] into acceptable disyllabic

sequences close to /əd.la/.[8] American listeners would not be able to produce that parse because they just do not hear the murmur. Another possibility is that French but not American listeners perceive and use the place information that could be conveyed by voicing leads. Although we are not aware of acoustic or perceptual studies demonstrating that cues to place of articulation are present in voicing leads and are exploited in perception, this possibility deserves consideration. While pre-release voicing murmurs are low in amplitude, they are somewhat audible and could contain place information insofar as the articulators arrive at their intended place of articulation during stop closure. We did measure spectral differences between the voicing leads of the Hebrew dental and velar stops stimuli used, and found some differences: "velar" murmurs tend to have a lower F1 and a higher F2 than "dental" ones. Yet, further investigation is needed to determine whether place information in voicing leads is perceptible, at least to listeners whose native language employs voicing lead.

To sum up, while the main determinants of the dental-to-velar perceptual repair effect must lie at an abstract phonological level (the native phonological grammar), we nevertheless found some modulation of the effect by the language-specific phonetic settings of voicing in stops. French listeners perceived Hebrew /dl/ more "faithfully" – as a coronal-initial cluster – than American listeners did. Yet, common to both groups of listeners, Hebrew /dl/ induced less dental-to-velar shift than /tl/, and this is most clearly evidenced in the categorization data. We dubbed this common pattern "voicing asymmetry." We turn to this pervading aspect of the data, found throughout the present study as well as in our previous work.

The /dl, tl/ asymmetry

Voicing asymmetry, which suggests higher perceptibility of dental place in Hebrew /dl/s than /tl/s, was found for all three groups but to different extents. It was especially clear in the categorization data of French and American listeners, though somewhat less so in the

latter group. As for discrimination, the asymmetry was substantial in the French data, very tiny and near ceiling in the Hebrew data, and nonsignificant in the American data despite numerical trends also suggesting easier discrimination for /dl/-/gl/ than /tl/-/kl/.

Intriguingly, the voicing asymmetry we consistently observed is at odds with the distribution of the /dl/-/gl/ and /tl/-/kl/ contrasts in the languages of the world. While both contrasts are infrequent across languages, /dl/-/gl/ is even less frequent than /tl/-/kl/, as shown by cross-linguistic evidence: if a language allows /dl/-/gl/, it also allows /tl/-/kl/ but not vice-versa (cf. Tobin's [2002] survey of forty languages with complex onset clusters).[9] Assuming that languages avoid contrasts doomed by perceptual similarity, the cross-linguistic pattern would suggest that /dl/-/gl/ is less perceptible than /tl/-/kl/, rather than the other way around.

In the account offered by Flemming (2002; also see Bradley, 2006), the lesser stability across languages of /dl/-/gl/ than /tl/-/kl/ is attributable to the poorer acoustic information appearing in the release burst of voiced than voiceless stops. There is generally less information in the bursts of /dl, gl/ than /tl, kl/ to distinguish velars from dentals because voiced stops have quieter bursts than voiceless stops (Zue, 1976). Our measurements of burst integrated energy in the Hebrew clusters (Table II) indeed confirm Zue's observation, with the qualification that the velar-dental difference, in terms of burst energy, is smaller for voiced than voiceless stops, thus predicting better discrimination for /tl/-/kl/ than for /dl/-/gl/.

The "quiet burst" argument, as well as the cross-linguistic data, are thus at odds with the asymmetry we observed. A possible explanation is that stop voicing contrasts generally involve a short vs. long lag VOT voicing distinction, not a pre-voiced vs. short or medium lag as in French or Hebrew. There may be perceivable information on place of articulation in the voicing lead of /dl/ and /gl/, as we discussed earlier, but this remains quite speculative at the moment. The /dl, tl/ asymmetry is also in line with the larger velar-dental SCG differential for /dl/-/gl/ than for /tl/-/kl/ (250 vs. 100 Hz), but this SCG differential varied across vowel

context and did not reach significance. Finally, no discernible acoustic difference was found between velar- and dental-initial /l/ clusters in either the /l/ or the following vowel formant patterns, whether the initial consonant was voiced or not. The formant patterns thus cannot explain the observed advantage of /dl/-/gl/ over /tl/-/kl/ in discrimination. Overall, then, the acoustic evidence does not explain well why /dl/-/gl/ is discriminated more easily than /tl/-/kl/: /tl/ does not clearly appear to be physically (or psychophysically) closer to /kl/ than /dl/ is to /gl/ in our Hebrew stimuli.

What about, as an alternative, "structural" linguistic arguments? Is the greater acceptability of /dl/ explainable by structural constraints? An example of differential structural acceptability has been recently offered by Moreton (2002), who showed there is a stronger perceptual bias against utterance-initial /dl/ than /bw/ for English-speaking listeners, although both clusters are illegal in English. Moreton interpreted this finding in terms of universal structural constraints, reasoning that "the *[dl] sequence is closer in sonority than *[bw] …, and hence a worse structural violation" (ibid., p. 57). Indeed, there is a smaller differential in sonority in /dl/ than in /bw/ because /l/ is less sonorous than /w/ in proposed sonority scales (Clements, 1990; Kahn, 1980). The recent work of Berent, Steriade, Lennertz, and Vaknin (in press) also demonstrates the role of sonority profile in cluster acceptability. But the problem in our case is that the differential sonority account predicts that /tl/ is just as unacceptable as /dl/ because the stop and the liquid are not further apart in sonority in /dl/ than in /tl/. Indeed, they might even be closer in /dl/ than /tl/, as voiced /d/ might possibly be considered more sonorous than voiceless /t/. Yet, the lesser acceptability of /tl/ is resistant to language-specific phonetic differences. It holds for speakers of French, English, and even Hebrew, suggesting that the asymmetry is driven by some structural, universal factor. Further research will be needed to pinpoint what kind of structural factor could adequately account for the lesser acceptability, hence the greater dental-to-velar perceptual shift, for /tl/ than /dl/.

Models of cross-language speech perception

Theoretically speaking, it is also important to ask whether extant models of nonnative speech perception can contribute to our understanding of the dental-to-velar perceptual shift. For this discussion, we limit ourselves to three widely known models: Speech Learning Model (SLM: Flege, 1995), Native Language Magnet (NLM: Kuhl & Iverson, 1995), and Perceptual Assimilation Model (PAM: Best, 1995). Could any of these models predict a dental-to-velar perceptual shift repairing /tl, dl/ into /kl, gl/?[10] How would the models explain the stronger repair effect for /tl/ than /dl/, for both English and French native listeners?

A general point must be made first, concerning the issue of whether stop + liquid clusters should be viewed analytically as a sequence of two phonemes, or rather, as unanalyzed syllable onsets. In the former case, all three models would agree that the perception of /dl, tl/ is reducible to that of /d, t/ plus /l/, and none would easily address the possible influence of the following /l/. In the second case only, if the clusters are perceived holistically as syllable onsets, might the three models differ in their predictions.

Although none of the three models explicitly addresses the issue of complex onsets, it is reasonable to assume that all should base their predictions about complex onset perception on *perceptual* distances. Such distances are conceivably determined both by objective stimulus properties and by how listeners represent speech sounds, depending on their native language. Models may differ on the latter issue of internal representations: NLM regards native categories as emerging from statistical clustering of auditory cues, whereas PAM instead regards them as the natural product of perceiving articulatory gestures that correspond to natively-produced gestures and as organized in a phonologically principled way. SLM is neutral with regards to this issue. As for the "objective" component in perceptual distances,

both NLM and SLM posit they are determined by the auditory properties of the stimuli, whereas PAM posits they are instead determined by their articulatory properties.

As discussed earlier, coronal stops released into /l/ may be intrinsically confusable with dorsal but not with labial stops for acoustic-articulatory reasons: dorsal and coronal but not labial constrictions are masked by the lateral constriction gesture (Bradley, 2006; Kawasaki, 1982). This distinction between non-labial and labial stop + /l/ would be captured in Articulatory Phonology by the major within- vs. between-organ tier distinction (Goldstein, Byrd, and Saltzman, 2006; Goldstein and Fowler, 2003): /tl/-/kl/ is a within-tier distinction (tongue tip vs. tongue dorsum, both on the tongue tier) and /tl/-/pl/ a between-tier distinction (tongue vs. lips), which is presumably easier (Goldstein and Fowler, 2003). PAM would follow Articulatory Phonology to account for the perception of /dl, tl/, whereas NLM and SLM would reason in terms of psychophysical distances. But all three models would agree that /gl, kl/ is a closer match with /dl, tl/ than /bl, pl/. Would they also agree that dental-to-velar shift is more likely with /dl/ than /tl/? NLM and SLM, as they reason in terms of psychophysical distances, would have difficulty explaining the asymmetry because the evidence for larger psychophysical distances between /dl/ and /gl/ than between /tl/ and /kl/ remains, at the moment, quite weak (SCG differentials) or speculative (perception of the pre-voicing murmur), and would have to address other acoustic measures that could favor the voiceless stops over the voiced ones (e.g., burst intensity). PAM, as it reasons in terms of underlying gestural dynamics, might offer a more principled account of the asymmetry.

The particular formalization used by the models does not help, either, to set them apart on the issue of voicing asymmetry. For SLM, nonnative speech sounds that are "similar" to some native category cause production and perception difficulties, due to interference from the native language; "new" speech sounds are easier because they escape such interference. Based on the observed data, that is, on a post hoc analysis, SLM could formalize the voicing

asymmetry in treating Hebrew /tl/, which is massively perceived as /kl/ by French and American listeners, as "similar" to a native /kl/, and /dl/, which receives mitigated judgments, as a "new" speech sound. For NLM, discrimination is more difficult around prototypes than non-prototypes. The better discrimination of /dl/-/gl/ than /tl/-/kl/ could be taken, then, to indicate lower typicality of Hebrew /dl/ as an exemplar of French or English /gl/ than that of /tl/ as an exemplar of /kl/. Again, this is a post hoc interpretation. PAM could also offer a post hoc interpretation of the data based on the majority of velar judgments for /tl/ and of dental judgments for /dl/, considering /tl/-/kl/ as a single-category assimilation, and /dl/-/gl/ as a category-goodness assimilation type, an easier type of contrast in PAM's classification. This interpretation would be just as problematic as that of SLM and NLM if PAM also based its predictions on psychophysical distances and auditory representations. But PAM explicitly posits that speech perception is driven by both the native phonological grammar and by the detection of underlying articulatory gesture properties. A plausible account from the PAM perspective might thus be framed in terms of gestural organization.

A speculative articulatory account

So far, there is no definitively satisfying answer to the question of why /dl/ is more discernible from /gl/ than /tl/ is from /kl/, and more easily categorized as dental than /tl/. The answer to the puzzle might lie in the different gestural organizations that could apply to the production of /tl/ and /dl/. For all stop + /l/ onsets, the production of the lateral segment /l/ requires retraction of the side(s) of the tongue, thinning it laterally by anterior extension of the tongue blade or tongue tip (TT), along with posterior movement of the tongue dorsum (TD). This maneuver, then, involves constriction at both the TT and TD. These constrictions are relatively innocuous to the production of a velar stop + /l/ onset, in which the stop involves a TD constriction anyway, but are somewhat detrimental to the dental quality of intended dental

stop + /l/ onsets. We speculate that the difference between /dl/ and /tl/ may reflect different inter-gestural phasing relationships between the stop and the /l/ for the voiced vs. voiceless stops, due to the fact that /tl/ requires a glottal abduction gesture during /t/, whereas this glottal gesture is absent for both /l/ and /d/. In other words, the phasing of /d/+/l/ is gesturally simpler and therefore less constrained, because there are fewer gestures to phase to one another. The perceptual side of this tentative gestural dynamics account is that 1) the intended dental gesture in /dl, tl/ may be misperceived as dorsal because it is contaminated by the tongue dorsum movement and 2) the confusion is all the more likely for the tighter gestural phasing relationship involved in [tl] than in [dl]. Thus, factoring out the intended place of articulation in the occlusion is conceivably more difficult in [tl] than in [dl]. This gestural dynamics account, however speculative at the moment, has the merit of 1) providing a low-level articulatory motivation of the dental-to-velar shift in general, and 2) accounting for the asymmetry in the perception of /dl/ versus /tl/. Unfortunately, there are no published studies on the gestural organization of English, French, or Hebrew stop + liquid onset clusters. The speculations offered here thus clearly call for further research.

Regardless of whether speech perception is determined by auditory analysis of acoustic events or recovery of articulatory gestures, it is clear from our findings that nonnative speech perception is systematically biased not only by individual native segments and contrasts, but also by phonotactic constraints. Perception of segmental contrasts that do exist in the native language is disturbed when the segments involved are part of a phonotactic pattern that is not permitted by that language. All extant models of nonnative speech perception fall short of anticipating these multi-segmental perceptual difficulties, because they have thus far focused only on perception of singleton segments. If we attempt to extend them to account for onset clusters, all would presumably agree that /kl/ is the best repair for */tl/

based on either acoustic or articulatory grounds. Yet, one intriguing pattern we consistently found, the /dl, tl/ asymmetry, is perhaps more easily explained in terms of articulatory gesture organization, in line with the current views of Articulatory Phonology. We provisionally favor PAM because it implicitly adopts the Articulatory Phonology framework. However, further articulatory studies of gestural organization in stop + liquid clusters are needed for a better understanding of their perception.

To summarize, much work is needed to develop predictions regarding multi-segmental influences (such as complex syllable onsets). More research is needed, as well, on understanding the relationships between the articulatory organization of speech at the relevant phonetic and phonological levels, and the perception of native and nonnative speech patterns, especially cross-linguistically.

ACKNOWLEDGMENTS

This work benefited from a "Cognitique" grant (LACO 1, French Ministère de la Recherche) to the first author and from a U.S. NIH (DC 00403) grant to the second. We are thankful to all the Israeli, American, and French subjects who participated. We also are indebted to Alice Faber for her help on Hebrew phonetics, to Christophe Pallier for his invaluable help with the R language for statistics, to Asaf Bachrach for recruiting and testing the Israeli subjects in Paris, to Nick Clements, Barbara Kunhert, Janet PierreHumbert, and Juan Segui for stimulating discussions on this work, and to three anonymous reviewers who helped improving an earlier version of this paper.

FOOTNOTES

[1] In the context of loanword phonology, words borrowed from a source language are introduced in the borrowing language by speakers who are aware of the necessary adjustments for the loanword to comply with the target language phonology (see, for example, Paradis and LaCharité, 1997). Loanword phonology primarily describes how "input" forms from a source language are adapted into "output" forms in a target language. In the context of speech perception, "phonological repair" refers to repairs made at a perceptual level by naive listeners who may not be aware of a phonological violation. Hence, the "perceptual repairs" we refer to do not involve conscious computations or cognitive strategies but instead occur automatically. Yet, the difference in meaning between loanword adaptation and perceptual repairs is rhetorical if one adheres to the view that adaptations largely reflect the layperson's perceptual assimilations (Vendelin and Peperkamp, 2004).

[2] In the languages which use vowel epenthesis in loanword adaptations, the inserted vowel is typically the least salient one (closest to Ø), that is, the phonetically most unmarked vowel in the vowel system (Kenstowicz, 2003). In Japanese, the default epenthetic vowel is /u/ realized [ɯ], the shortest vowel in Japanese (Han, 1962); only /u/ and /i/ undergo devoicing or reduction, and /i/ is the more central of the two (Keating and Huffman, 1984). The case of *drama* adapted into *dorama* or *zurama* (/dzurama/) is somewhat more complex. Because Japanese lacks /du/ in its /d/-initial syllables and has /dzu/ instead, a second repairing step is necessary with two possible solutions: maintain /d/ and use the non-default /o/ vowel, or use the default /u/ vowel and substitute /dz/ for /d/.

[3] Reports differ as to whether trilled /r/s are often rather than occasionally produced by Hebrew speakers. Trilling would be more frequent in the Oriental variety of Hebrew, but tends to spread to the non-Oriental variety as well (Laufer, 1990). Trilled /r/s occur in some varieties of continental French. They are rare in current Parisian pronunciation, but do occur in restricted circumstances (e.g., in songs), and should not be unfamiliar to French listeners.

[4] In French, the number of types is 46, 28, and 7 for /kla/, kli/, and /klu/ words, respectively; the number of tokens is 1634, 940, and 353. For /gl/ words, the number of types is 23, 13, and 10 for /gla/, gli/, and /glu/, respectively; the number of tokens is 1108, 584, and 271. These counts are drawn from the Brulex database (Content, Mousty, and Radeau, 1990) based on a 23.5 million word corpus. In English, the number of types is 46, 12, and 2 for /kla/, kli/, and /klu/ words, respectively; the number of tokens is 174, 90, and 15. For /gl/ words, the number of types is 29, 9, and 11 for /gla/, gli/, and /glu/, respectively; the number of tokens is 205, 12, and 40 (from the one million words Brown corpus: Francis and Kucera, 1982).

[5] In the Hebrew participants' data, the advantage of velar over dental targets was significant, although numerically tiny, for the percent correct but not for the RT discrimination performance In the French data, the advantage was found for both percent correct and RT performance but was confined to the /l/ cluster contrasts (75.1% vs. 69 % correct; 866 ms vs. 1281 ms) and virtually absent in the /r/ cluster contrasts (97.7% vs. 97.2 %; 284 vs. 337 ms).

[6] The following function (`compcorr`) can be used in the R environment (Becker, Chambers, and Wilks, 1988) to test for difference in correlation strengths (independent samples):

```
compcorr <- function(n1, r1, n2, r2)
{
# compares two correlation coefficients; returns difference and p-value
# Fisher Z-transform
      zf1 <- 0.5*log((1 + r1)/(1 - r1))
      zf2 <- 0.5*log((1 + r2)/(1 - r2))
# difference and p-value returned as list(diff, pval)
      dz <- (zf1 - zf2)/sqrt(1/(n1 - 3) + (1/(n2 - 3)))
      pv <- 2*(1 - pnorm(abs(dz)))
      return(list(diff=dz, pval=pv))
}
```

[7] The lesser incidence of velar judgments for /dl, tl/ in the /a/ context is consistent with the higher discriminability of /dl/-/gl/ or /tl/-/kl/ in that context for both American and French listeners. No such correlation is observed in the case of /dlu, tlu/, which should be more confusable with /glu, klu/ on acoustic grounds but actually induced less dental-to-velar shifts than /dli, tli/, or in the case of /dli, tli/, which led to the largest amount of velar shifts but should not have done so, given their acoustic characteristics, as suggested by Table II.

[8] Hallé et al. (1998) proposed that the voicing murmur of French [d], a "weakly resonant vowel-like sound" (ibid., p. 605) could be perceived as a vowel and enhance the perceptibility of /dl/ due to the fact that, in French, dental stop + /l/ clusters are acceptable after a vowel (word-medially) because they allow for a phonotactically legal parse in this position, with a syllable boundary before /l/ (Dell, 1995; for a review, see Goslin and Frauenfelder, 2000).

[9] This cross-linguistic pattern also holds for languages in which voiced stops are phonetically prevoiced, such as Mexican Spanish. Mexican Spanish has few /tl/-words, most of them borrowed from Nahuatl (e.g., *tlaco* 'coin', *tlapalería* 'hardware store'), but no /dl/-word. One exception is Creole French (both the Indian Ocean and Caribbean varieties), in which word-initial /dl/ but *not* /tl/ is attested. Word-initial /dl/ appears in adaptations from French words in which agglutination is involved, such as [dle] for "lait" (from *du lait*, 'some milk'), or quite commonly, [dlo] for "eau" (from *de l'eau*, 'some water') as in *mon dlo lé au fé* ('my water is on the stove'). These forms alternate with [dəle] or [dəlo] (Chaudenson, 1974: 654).

[10] Other possible repairs of, for example, /tla/ are simplification (/ta/), epenthetic vowel insertion (/təla/), liquid substitution (/tra/), and labial substitution (/pla/). Labial responses suggesting labial substitution were negligible in our identification data. Dental responses to, for example, /tla/ might have occasionally corresponded to /ta/, /təla/, or /tra/ repairs. These repairs, however, very rarely appeared in open response transcription tasks (Hallé et al., 1998) and are not compatible with the difficulties encountered by non-Hebrew listeners with /dl/-/gl/ or /tl/-/kl/. We therefore conservatively restrain the discussion to velar substitution repairs.

REFERENCES

Becker, R., Chambers, J., and Wilks, A. (**1988**). *The New S Language* (Chapman & Hall, London).

Berent, I., Steriade, D., Lennertz, T., & Vaknin, V. (**in press**). "What we know about what we never heard: Evidence from perceptual illusions," Cognition.

Best, C. (**1995**). "A direct realist perspective on cross-language speech perception," in *Speech perception and linguistic experience: Issues in cross-language research*, edited by W. Strange and J. Jenkins (York Press, Timonium, MD), pp. 171-204.

Blumstein, S., and Stevens, K. (**1980**). "Perceptual invariance and onset spectra for stop consonants in different vowel environments," J. Acoust. Soc. Am. **67**, 648-662.

Boyce, S. and Espy-Wilson, C. (**1997**). "Coarticulatory stability in American English /r/," J. Acoust. Soc. Am. **101**, 3741-3753.

Bradley, T. G. (**2006**). "Contrast and markedness in complex onset phonotactics," Southwest J. Linguistics **25**, 29-58.

Chafcouloff, M. (**1979**). "Les propriétés acoustiques de /j, ɥ, w, r, l/ en français," ("Acoustic characteristics of French /j, ɥ, w, r, l/"), Travaux de l'Institut de Phonétique d'Aix, report **6**, 11-24.

Chaudenson, R. (**1974**). *Le lexique du parler Créole* (*Lexicon of Creole*, Honoré Champion, Paris).

Chayen, M. J. (**1973**). *The phonetics of modern Hebrew*. (Mouton, Paris).

Cho, T.-H., and Ladefoged, P. (**1999**). "Variation and universals in VOT: evidence from 18 languages," J. Phonetics **27**, 207-229.

Clements, G. N. (**1990**). "The role of the sonority cycle in core syllabification," in *Between the grammar and the physics of speech*, edited by J. Kingston and M. Beckman (Cambridge University Press, New York), pp. 283-333.

Content, A., Mousty, P., and Radeau, M. (**1990**). "Une base de données lexicales informatisée pour le français écrit et parlé," ("A computerized lexical database for written and spoken French"), Année Psychol. **90**, 551-566.

Dehaene-Lambertz, G., Dupoux, E., and Gout, A. (**2000**). "Electrophysiological correlates of phonological processing: A cross-linguistic study," J. Cog. Neurosscience **12**, 635-647.

Delattre, P. (**1969**). "L'R parisien et autres sons du pharynx," ("Parisian R and other pharyngeal speech sounds"), The French Review **43**, 5-22.

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

Devens, M. (**1978**). *The phonetics of Israeli Hebrew: 'Oriental' versus 'General' Israeli Hebrew*, PhD Dissertation (UCLA University, Los Angeles).

Devens, M. (**1980**). "Oriental Israeli Hebrew: A study in phonetics," Afroasiatic Linguistics **7**, 127-142.

Dupoux, E., Kakehi, Y., Hirose, C., Pallier, C., and Mehler, J. (**1999**). "Epenthetic vowels in Japanese: a perceptual illusion?" J. Exp. Psychol. Hum. Percept. Perform. **25**, 1568-1578.

Dupoux, E., Pallier, C., Kakehi, Y., and Mehler, J. (**2001**). "New evidence for prelexical phonological processing in word recognition," Lang. Cog. Processes **16**, 491-505.

Espy-Wilson, C. (**1992**). "Acoustic measures for linguistic features distinguishing the semivowels /wjrl/ inAmerican English," J. Acoust. Soc. Am. **92**, 736-757.

Fischer-Jørgensen, E. (**1954**). "Acoustic analysis of stop consonants," Miscellanea Phonetica **2**, 42-59.

Flege, J. (**1995**). "Second language speech learning: Theory, findings, and problems," in *Speech perception and linguistic experience: Issues in cross-language research*, edited by W. Strange and J. Jenkins (York Press, Timonium, MD), pp. 233-277.

Flemming, E. (2002). *Auditory representations in phonology* (Routledge, New York).

Fowler, C., Brown, J., Sabadini, L., and Weihing, J. (**2003**). "Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks," J. Mem. Language **49**, 396-413.

Francis, W., and Kucera, H. (**1982**). *Frequency analysis of English usage: Lexicon and grammar* (Houghton-Mifflin, Boston).

Goldstein, L., Byrd, D., and Saltzman, E. (**2006**). "The role of vocal tract gestural action units in understanding the evolution of phonology," in *From action to language: The mirror neuron system*, edited by M. Arbib (Cambridge Universty Press, MA), pp. 215-249.

Goldstein, L., and Fowler, C. (**2003**). "Articulatory phonology: a phonology for public language use," in *Phonetics and phonology in language comprehension and production: Differences and similarities*, edited by A. Meyer and N. Schiller (Mouton de Gruyter, Berlin), pp. 159-208.

Goslin, J., and Frauenfelder, U. (**2000**). "A comparison of theoretical and human syllabification," Language and Speech **44**, 409-436.

Guion, S., Flege, J., Akahane-Yamada, R., and Pruitt, J. (**2000**). "An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants," J. Acoust. Soc. Am. **107**, 2711-2724.

Halle, M., Hugues, G., and Radley, J.-P. (**1957**). "Acoustic properties of stop consonants," J. Acoust. Soc. Am. **29**, 107-116.

Hallé, P., Best, C., and Levitt, A. (**1999**). "Phonetic vs. phonological influences on French listeners' perception of American English approximants," J. Phonetics **27**, 281-306.

Hallé, P., Chang, Y.-C., and Best, C. (**2004**). "Categorical perception of Taiwan Mandarin Chinese tones by Chinese versus French native speakers," J. Phonetics **32**, 395-421.

Hallé, P., de Boysson-Bardies, B., and Vihman, M. (**1991**). "Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infants," Language and Speech **34**, 299-318.

Hallé, P., Segui, J., Frauenfelder, U., and Meunier, C. (**1998**). "Processing of illegal consonant clusters: A case of perceptual assimilation?" J. Exp. Psychol. Hum. Percept. Perform. **24**, 592-608.

Hughes, J. W. (**1946**). "The threshold of audition for short periods of stimulation," Proc. Royal Society B **133**, 486-490.

Han, M. (**1962**). "Unvoicing of vowels in Japanese," 音声の研究 (Phonetic Studies) **10**, 210-231.

Kahn, D. (**1980**). *Syllable-based generalizations in English phonology* (Garland, New York).

Kawasaki, H. (**1982**). *An acoustical basis for universal constraints on sound sequences*, PhD Dissertation (University of California, Berkeley).

Keating, P., and Huffman, M. (**1984**). "Vowel variation in Japanese," Phonetica **41**, 191-207.

Kenstowicz, M. (**2003**). "The role of perception in loanword phonology. A review of *Les emprunts linguistiques d'origine européenne en Fon*," Studies in African Linguistics **32**, 95-112.

Kewley-Port, D. (**1983**). "Time-varying features as correlates of place of articulation in stop consonants," J. Acoust. Soc. Am. **73**, 606-608.

Kuhl, P., and Iverson, P. (**1995**). "Linguistic experience and the "perceptual magnet effect"," in *Speech perception and linguistic experience: Issues in cross-language research*, edited by W. Strange and J. Jenkins (York Press, Timonium, MD), pp. 121-154.

Laufer, A. (**1990**). "Hebrew (Illustrations of the IPA)," J. Int. Phonetic Assoc. **20**, 40-43.

Laufer, A. (**1998**). "Voicing in contemporary Hebrew in comparison with other languages," Hebrew Studies **39**, 143-179.

Lisker, L., and Abramson, A. (**1964**). "(1964) Cross-language study of voicing in initial stops: acoustical measurements," Word **20**, 384-422.

Mann, V. A. (**1980**). “Influence of preceding liquid on stop consonant perception,” Percept. Psychophys. **28***,* 407-412.

Mann, V., and Repp, B. (**1981**). “Influence of preceding fricative on stop consonant perception,” J. Acoust. Soc. Am. **69**, 548-558.

Massaro, D., and Cohen, M. (**1983**). "Phonological context in speech perception," Percept. Psychophys. **34**, 338-348.

Moreton, E. (**2002**). "Structural constraints in the perception of English stop-sonorant clusters," Cognition **84**, 55-71.

Nearey, T., and Rochet, B. (**1994**). "Effects of place of articulation and vowel context on VOT production and perception for French and English stops," J. Int. Phonetic Assoc. **24**, 1-18.

New, B., Pallier, C., Ferrand, L., and Matos, R. (**2001**). "Une base de données lexicales du français contemporain sur internet: LEXIQUE^{TM}," ("An internet on-line lexical database for modern French: LEXIQUE^{TM}"), L'Année Psychologique **101**, 447-462.

Obler, L. (**1982**). "The parsimonious bilingual," in *Exceptional language and linguistics*, edited by L. Obler and L. Menn (Academic Press, New York), pp. 339-346.

Padgett, J. (**1991**). *Structure in feature geometry*, PhD Dissertation (University of Massachusetts, Amherst).

Paradis, C., and LaCharité, D. (**1997**). “Preservation and minimality in loanword adaptation,” J. Linguistics **33**, 379-430.

Paradis, C., and LaCharité, D. (**2001**). “Guttural deletion in loanwords,” Phonology **18**, 255-300.

Peterson, G., and Lehiste, I. (**1960**). "Duration of syllable nuclei in English," J. Acoust. Soc. Am. **32**, 693-703.

Pitt, M. (**1998**). "Phonological processes and the perception of phonotactically illegal consonant clusters," Percept. Psychopys. **60**, 941-951.

Polivanov, E. (**1931**). "La perception des sons d'une langue étrangère," ("The perception of non-native language sounds"), Travaux du Cercle Linguistique de Prague **4**, 79-96.

Raphael, L., Tobin, Y., Faber, A., Most, T., Kollia, H., and Milstein, D. (**1995**). "Intermediate values of voice onset time," in *Producing speech: Contemporary issues for Katherine Harris*, edited by F. Bell-Berti and L. Raphael (AIP Press, xx), pp. 117-127.

Raphael, L., Tobin, Y., and Most, T. (**1983**). "Atypical VOT categories in Hebrew and Spanish," J. Acoust. Soc. Am. **74**, S89A.

Rosen, H. (**1962**). *A textbook of Israeli Hebrew* (The University of Chicago Press, Chicago).

Saerens, M., Serniclaes, W., and Beeckmans, R. (**1989**). "Acoustic versus contextual factors in stop voicing perception in spontaneous French," Language and Speech **32**, 291-314.

Scharf, B. (**1978**). "Loudness," in *Handbook of perception, vol. 4: Hearing*, edited by E. Carterette and P. Morton (Academic Press, New York), pp. 187-242.

Segui, J., Frauenfelder, U., and Hallé, P. (**2001**). "Phonotactic constraints shape speech perception: Implications for sublexical and lexical processing," in *Language, Brain, and Cognitive Development: Essays in Honor of Jacques Mehler*, edited by E. Dupoux (MIT Press, Cambridge, MA), pp. 195-208.

Selkirk, E. (**1988**). *Dependency, place, and the notion "tier"*, unpublished manuscript (University of Massachusetts, Amherst).

Shinohara, S. (**2006**). "Perceptual effects in final cluster reduction patterns," Lingua **116**, 1046-1078.

Silverman, D. (**1992**). “Multiple scansions in loanword phonology: Evidence from Cantonese,” Phonology **9**, 289-328.

Simon, M. (**1967**). *Les consonnes françaises – Mouvements et positions articulatoires : La lumière de la cinématographie* (*French consonants – articulatory movements and positions: the contribution of cinematography*, Klincksieck, Paris).

Stevens, K., and Blumstein, S. (**1978**). "Invariant cues for place of articulation in stop consonants," J. Acoust. Soc. Am. **64**, 1358-1368.

Trubetzkoy, N. (**1939**). "Grundzüge der Phonologie," ("Principles of Phonology"), Travaux du Cercle Linguistique de Prague **7**, 1-271.

Téné, D. (**1972**). "The measured durations of Hebrew vowels – the schwa," in *Readings in Phonology*, edited by U. Ornan (Jerusalem Hebrew University Press, Jerusalem), pp. 239-251 (in Hebrew).

Tobin, Y. (**2002**). “Phonology as human behavior: initial consonant clusters across languages,” in *Signal, meaning, and message*, edited by W. Reid, R. Ortheguy, and N. Stern (John Benjamin, Amsterdam), pp. 191-255.

Vendelin, I., & Peperkamp, S. (**2004**). “Evidence for phonetic adaptation of loanwords: and experimental study,” in *Actes des Journées d’Etudes Linguistiques 2004*, pp. 129-131.

Yip, M. (**1993**). “Cantonese loanword phonology and Optimality Theory,” J. East Asian Ling. **2**, 261-291.

Zawadzki, P. A. and Kuehn, D. P. (**1980**). “A cineradiographic study of static and dynamic aspects of American English /r/,” Phonetica **37**, 253-266.

Zue, V. (**1976**). *Acoustic characteristics of stop consonants: a controlled study*. PhD Dissertation (MIT University, MA).

Table I. Characteristics of F0 contours in the speech materials retained: F0 at contour onset, F0 range, mean F0 (all in Hz), and mean $|d^2F0/d^2t|$ (an index of F0 fluctuations)[a], and contour duration (in ms), according to cluster.

Cluster	**Onset F0**	**F0 range**	**mean F0**	**$\|d^2F0/d^2t\|$**	**duration**
/tl/	147	15	152	5.6	258
/kl/	150	14	151	6.7	249
/tr/	146	19	151	8.7	280
/kr/	147	22	151	10.9	272
/dl/	119	43	139	17.4	385
/gl/	121	42	138	13.6	381
/dr/	124	37	138	28.2	397
/gr/	124	42	138	21.5	386

[a] see Hallé, de Boysson-Bardies, and Vihman (1991: 303-304) for further details on this fluctuation index.

Table II. Acoustic characteristics of the dental- vs. velar-initial cluster stimuli according to liquid and vowel context. Voice onset time (VOT) is expressed in ms, spectral centre of gravity (SCG) in Hz, and burst integrated energy (BIE) in dB sec. Δs stand for (velar - dental) differences.

	vowel	/l/ context						/r/ context					
		Voiced			Voiceless			Voiced			Voiceless		
		/dl/	**/gl/**	**Δ**	**/tl/**	**/kl/**	**Δ**	**/dr/**	**/gr/**	**Δ**	**/tr/**	**/kr/**	**Δ**
	/a/	-147	-97	***50***	49	78	***29***	-97	-85	***12***	40	99	***59***
VOT	/i/	-126	-105	***21***	60	101	***41***	-94	-76	***18***	51	101	***50***
	/u/	-106	-94	***12***	67	93	***26***	-102	-97	***5***	60	95	***35***
	mean	-126	-99	***27***	59	91	***32***	-98	-86	***12***	50	98	***48***
	/a/	4212	3552	***-660***	4066	3699	***-367***	3944	3631	***-313***	3958	3724	***-234***
SCG	/i/	4017	3719	***-298***	3963	3858	***-105***	3939	3340	***-599***	3915	3566	***-349***
	/u/	3206	3413	***+207***	3571	3742	***+171***	3833	3699	***-134***	3879	3518	***-361***
	mean	3812	3561	***-251***	3867	3767	***-100***	3906	3557	***-349***	3918	3602	***-316***
	/a/	1.25	1.49	***0.24***	3.17	4.72	***1.55***	1.13	1.81	***0.69***	2.57	5.57	***3.00***
BIE	/i/	1.30	1.55	***0.25***	3.86	6.22	***2.36***	1.22	1.93	***0.72***	3.47	5.82	***2.36***
	/u/	1.32	1.80	***0.47***	4.22	5.91	***1.70***	1.35	2.15	***0.80***	3.67	5.90	***2.22***
	mean	1.29	1.61	***0.32***	3.75	5.62	***1.87***	1.23	1.97	***0.73***	3.24	5.76	***2.53***

FIGURE CAPTIONS

Figure 1. Discrimination performance of Hebrew vs. French subjects: (A) percent correct discrimination, and (B) RT for correct responses in ms. Error bars indicate standard error.

Figure 2. Categorization of the initial stop of Hebrew /dl/ and /tl/ by French subjects: (A) percentages of dental and velar responses, and (B) corresponding "fit indexes." Error bars indicate standard error.

Figure 3. French subjects: Correct discrimination of /dl/-/gl/ or /tl/-/kl/ as a function of the percentage of velar responses to /dl/ (empty circles 'o') or to /tl/ (black diamonds '◆'), respectively.

Figure 4. Discrimination performance of American vs. French subjects: (A) percent correct discrimination, and (B) RT for correct responses in ms. Error bars indicate standard error.

Figure 5. Categorization of the initial stop of Hebrew /dl/ and /tl/ by American subjects: (A) percentages of dental and velar responses, and (B) corresponding "fit indexes." Error bars indicate standard error.

Figure 6. American subjects: Correct discrimination of /dl/-/gl/ or /tl/-/kl/ as a function of the percentage of velar responses to /dl/ (empty circles 'o') or to /tl/ (black diamonds '◆'), respectively.

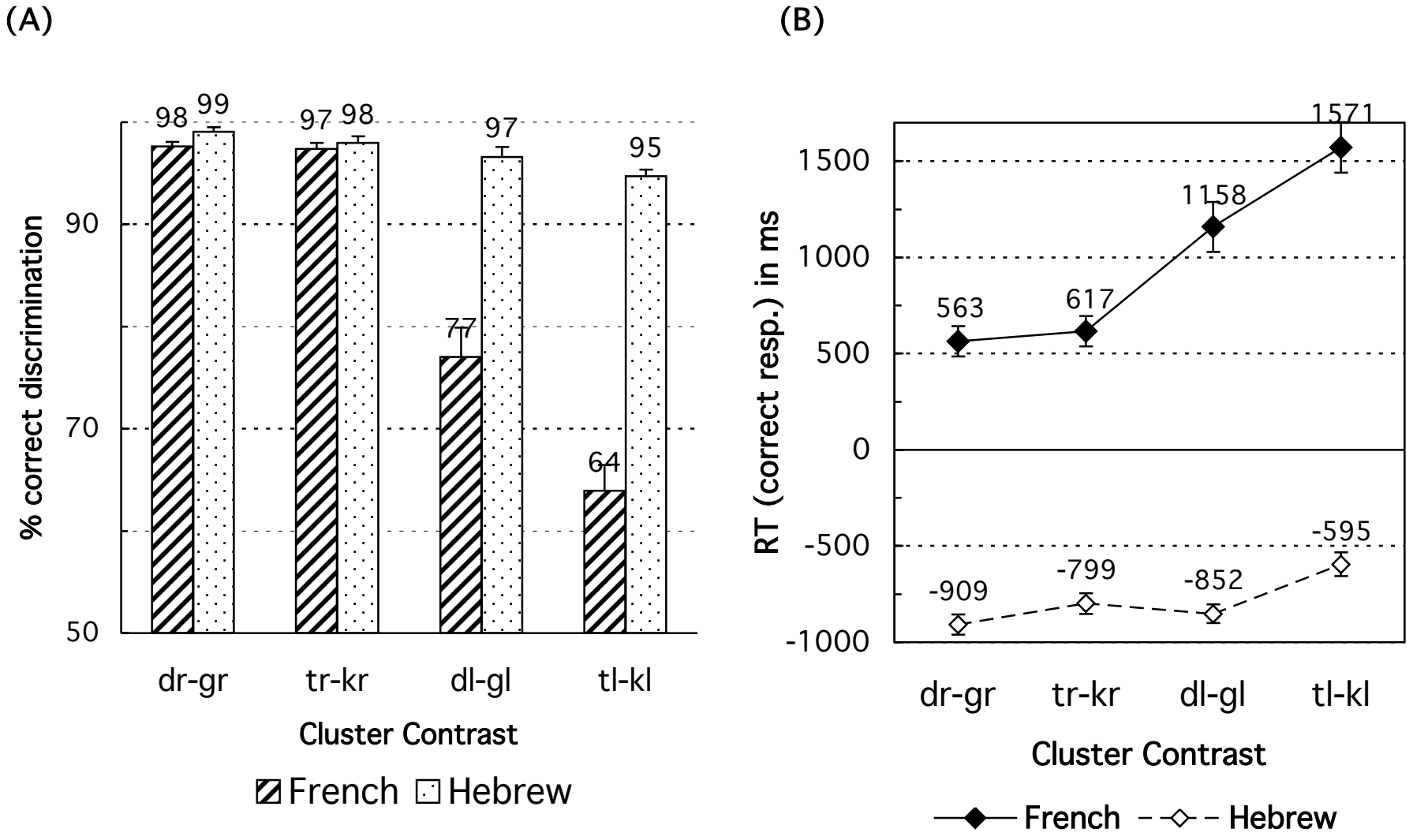
(A)
% correct discrimination
90
70
50
98
99
97
98
77
97
64
95
dr-gr
tr-kr
dl-gl
tl-kl
Cluster Contrast
French
Hebrew
(B)
RT (correct resp.) in ms
1500
1000
500
0
-500
-1000
563
617
1158
1571
-909
-799
-852
-595
dr-gr
tr-kr
dl-gl
tl-kl
Cluster Contrast
French
Hebrew

Figure 1

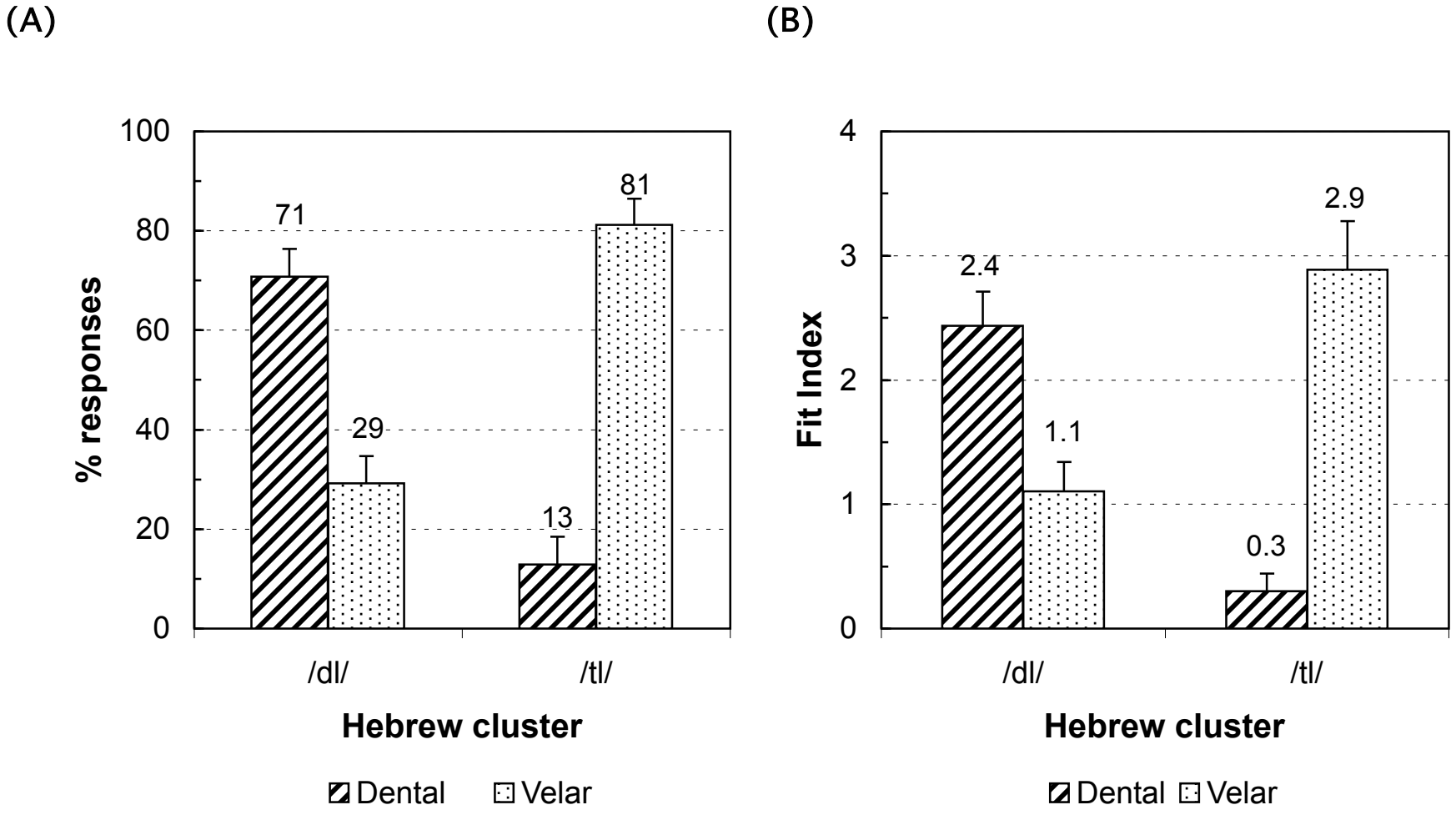
(A)
100
80
60
40
20
0
% responses
71
29
13
81
/dl/
/tl/
Hebrew cluster
Dental
Velar
(B)
4
3
2
1
0
Fit Index
2.4
1.1
0.3
2.9
/dl/
/tl/
Hebrew cluster
Dental
Velar

Figure 2

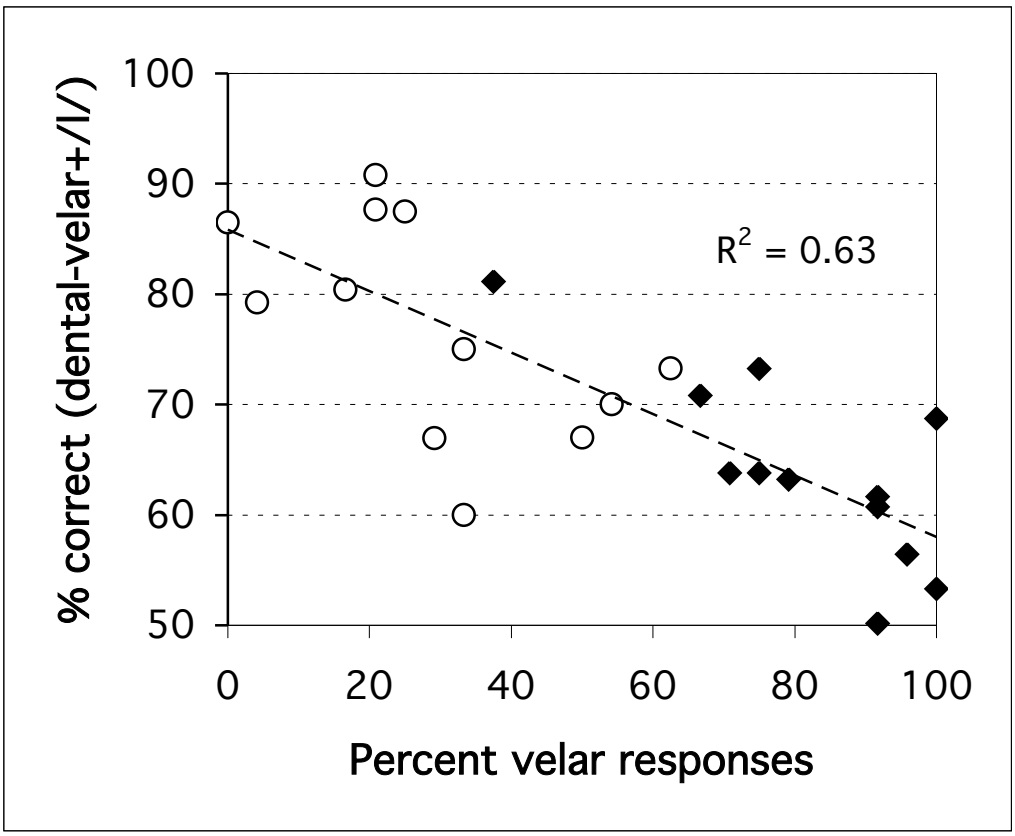
100
90
80
70
60
50
% correct (dental-velar+/l/)
R^2 = 0.63
0
20
40
60
80
100
Percent velar responses

Figure 3

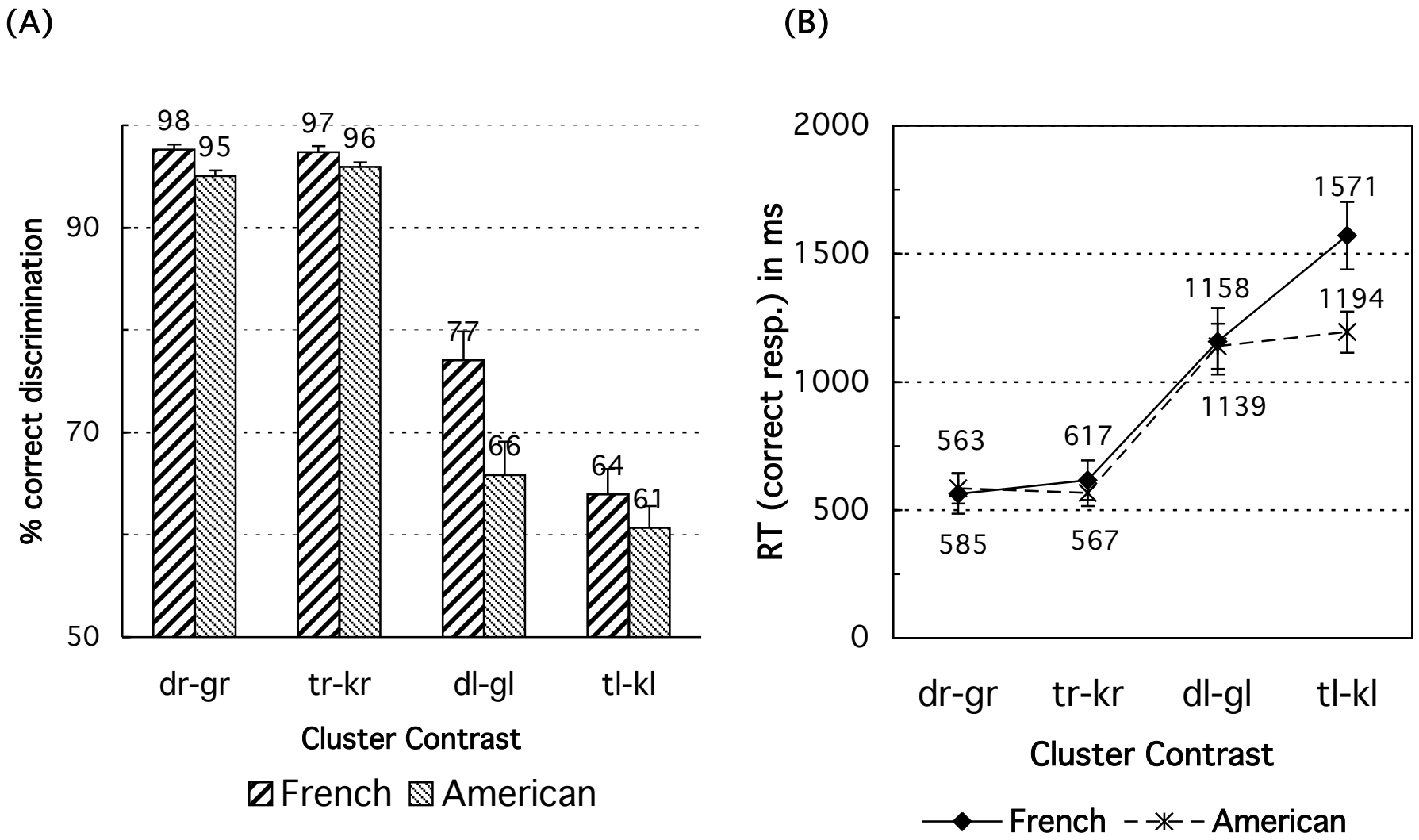
(A)
% correct discrimination
90
70
50
98
95
97
96
77
66
64
61
dr-gr
tr-kr
dl-gl
tl-kl
Cluster Contrast
French
American
(B)
RT (correct resp.) in ms
2000
1500
1000
500
0
1571
1158
1194
1139
563
617
585
567
dr-gr
tr-kr
dl-gl
tl-kl
Cluster Contrast
French
American

Figure 4

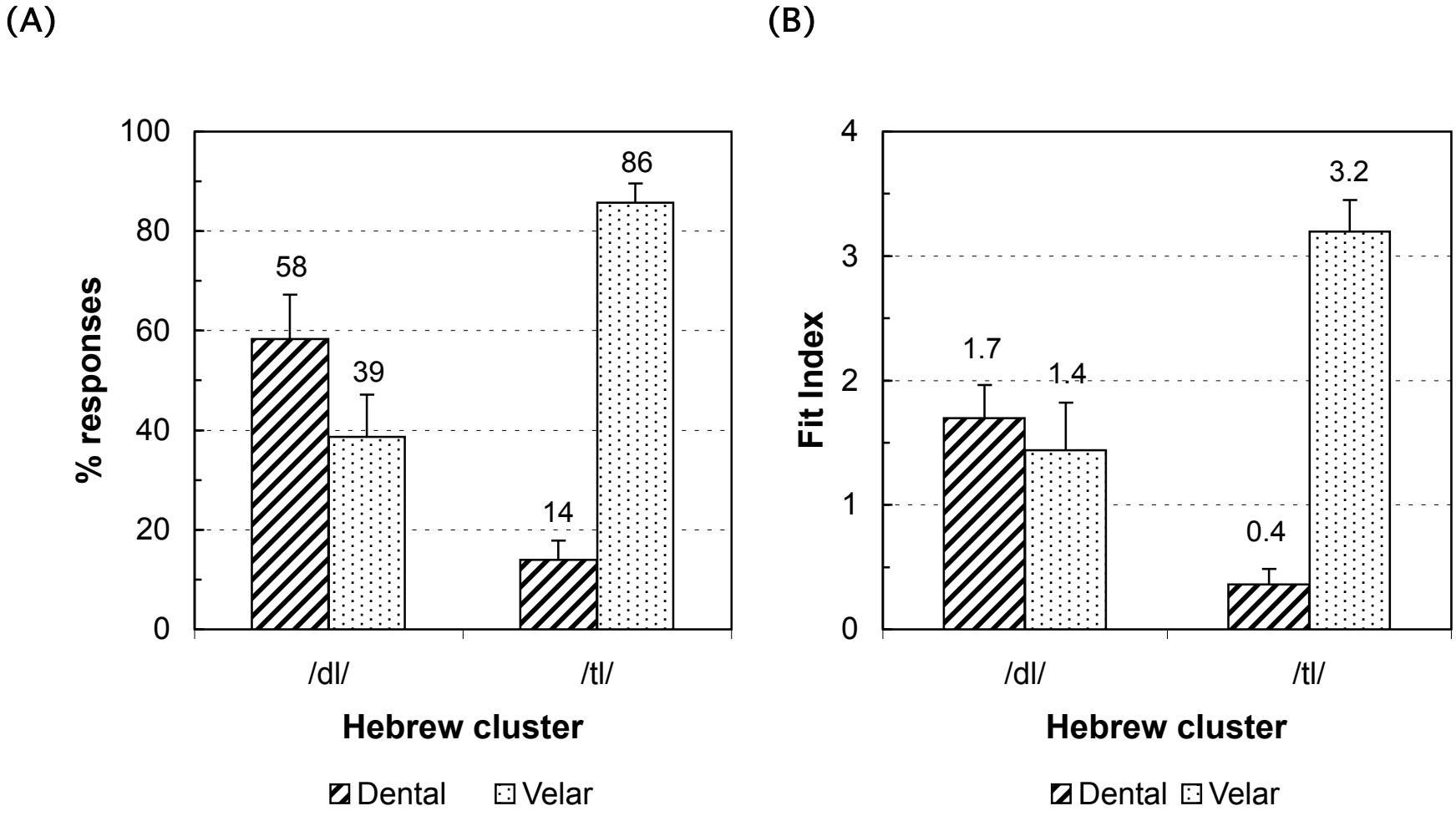
(A)
100
80
60
40
20
0
% responses
58
39
14
86
/dl/
/tl/
Hebrew cluster
Dental
Velar
(B)
4
3
2
1
0
Fit Index
1.7
1.4
0.4
3.2
/dl/
/tl/
Hebrew cluster
Dental
Velar

Figure 5

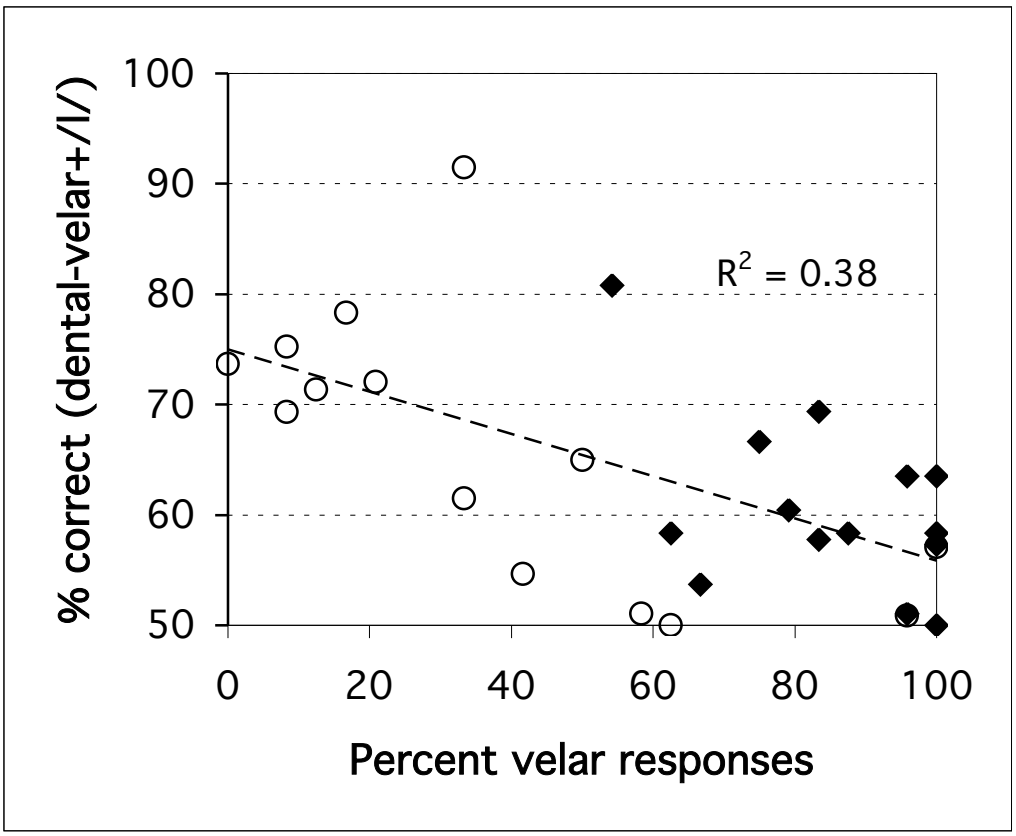
% correct (dental-velar+/l/)
100
90
80
70
60
50
0
20
40
60
80
100
Percent velar responses
R2 = 0.38

Figure 6