STOP PLACE CONTRASTS BEFORE LIQUIDS

Edward Flemming

Department of Linguistics & Philosophy, MIT
flemming@mit.edu

ABSTRACT

All languages allow stop place contrasts in pre-vocalic position. Many languages allow stop place contrasts before liquids [l, r]. Indeed, stop-liquid clusters like [br, gl] are among the most common word-initial consonant clusters [2]. The preference for stop-liquid clusters is commonly attributed to Sonority Sequencing constraints – specifically a preference for large sonority rises in onset clusters. Here we pursue an analysis in terms of the availability of cues to place contrasts: contrasts are preferentially permitted in environments where they are most distinct. According to this analysis the high sonority of liquids is relevant only insofar as more sonorous sounds are generally better able to support the realization of cues to preceding consonants.

This line of analysis also offers an account of more specific restrictions on place contrasts that are not amenable to an analysis in terms of sonority sequencing. Specifically, we will present evidence that the cross-linguistic dispreference for coronal-velar contrasts before laterals is due to the acoustic similarity of these clusters.

1. INTRODUCTION

There is evidence that phonological restrictions on sound sequences are organized to ensure that adequate perceptual cues to all the segmental contrasts can be realized, e.g. [5, 9, 10, 13]. For example, all languages allow stop place contrasts before vowels since the availability of release cues make place contrasts less confusable in pre-vocalic position than in pre-consonantal or pre-pausal position. In this paper we develop this line of analysis towards an account of the cross-linguistic prevalence of stop place contrasts before liquids.

It is well established that the primary cues to stop place in word-initial pre-vocalic context reside in the quality of the stop burst and the second and third formant transitions (e.g. [1]). Given a cue-based approach to cluster markedness, we might expect that liquids should not be far behind vowels in their ability to support stop place contrasts because they involve relatively open constrictions, so even if a stop is released into a liquid, it should generate an audible burst [7]. Liquids also have well-defined formant structure, so it should be possible to realize formant transitions on a liquid. These expectations are broadly confirmed in a study of American English, although we will see that stop place is not differentiated by formant transitions before [l].

2. STOP-LIQUID CLUSTERS IN AMERICAN ENGLISH

Like many other languages, English contrasts three places of articulation before [r]: [br, dr, gr] and two before [l]: [bl, gl].

Five speakers produced two repetitions of a set of words consisting of 6 near-minimal triplets contrasting [br, dr, gr] (e.g. *brown, drown, ground*), 6 near-minimal pairs contrasting [bl, gl]. (e.g. *blow, glow*), and nine triplets contrasting [b, d, g] before nine different vowels.

2.1. Measurements

F2 and **F3** were measured at formant onset.

Burst duration was measured from the release of the stop to the onset of the first formant (which usually coincided with the onset of voicing), so the burst is taken to include the transient at the release of the stop and any frication generated at the place of the stop constriction.

Burst shape was measured from a smoothed spectrum [4] derived by averaging a series of seven squared-magnitude FFT spectra calculated from 3 ms Hamming windows, with a 1 ms shift between windows (so the centers of the first and last windows are 6 ms apart). The first window was centered on the stop release.

	bl		gl		br		dr		gr	
F2 onset (Hz)	1081	(137)	1133	(155)	1225	(150)	1641	(177)	1272	(215)
F3 onset (Hz)	3042	(257)	2961	(252)	1800	(214)	1998	(224)	1957	(205)
Burst durn (ms)	11	(3)	31	(10)	16	(6)	44	(15)	35	(9)
Burst peak (Hz)	109	(115)	1209	(285)	118	(186)	1924	(764)	1062	(333)
Amid-Ahi (dB)	8.73	(5.1)	19.8	(4.5)	12.01	(3.87)	17.00	(5.46)	18.01	(3.18)

Table 1: Mean measurements from stop-liquid clusters produced by 5 speakers of American English.

The shape of the stop burst was quantified in terms of two parameters, **Burst Peak** and a relative amplitude measure, **Amid-Ahi** (cf. [11]). Burst Peak is the frequency of the amplitude peak in the burst spectrum. Amid is the average amplitude of the burst over a mid-frequency range from 1.25 kHz–3 kHz. Ahi is the average amplitude in the high frequency range 3.5 kHz–8 kHz. The difference between these two quantities provides a measure of spectral shape.

2.2. Results

The measurements are summarized in table 1. [bl] and [gl] are distinguished by burst but not formant onsets. The differences in burst duration, Burst Peak and Amid-Ahi are all significant.

[br, dr, gr] clusters are distinguished by both formant transitions and burst properties. [dr] has a significantly higher F2 onset than [br, gr], while [br] has a lower F3 onset than [dr, gr]. Labial [br] has a shorter burst than [dr, gr] and all three places of articulation are distinguished by frequency of the Burst Peak.

Stop place before [r] is distinguished by burst quality and formant transitions in much the same way as in prevocalic contexts, hence stop place contrasts are common in this context. Note that a wide variety of rhotics can license a full range of preceding stop place contrasts, so additional studies are required to test the hypothesis that these rhotics also allow for good place cues. Taps, flaps and trills generally allow audible release of preceding stops and the realization of formant transitions during the open phases of these sounds, but the nature of place cues before uvular fricative rhotics, as in French, is much less clear.

Before [l], stop place is distinguished by burst properties only. We will see in the next section that this applies to [dl] clusters as well, and may account for the fact that only two stop places contrast before [l] in many languages, including English.

3. CORONAL-VELAR CONTRASTS BEFORE LATERALS

The liquids are not equal in their ability to license place contrasts. Many languages with Cr and Cl clusters allow more place contrasts before /r/ than before /l/, and few, if any, reverse this pattern. A particularly widespread restriction is the exclusion of coronal stop+l clusters /dl, tl/, as in English [5].

Kawasaki [5] argues that this is part of a broader pattern in which the contrast between coronal and velar stops is neutralized before laterals. While the result of neutralization is usually a velar, there are languages in which coronal-lateral clusters are permitted while velar-lateral clusters are illegal, (e.g. Haroi [8] and other Chamic languages, and Katu dialects [12]) or where there is free variation between the two (Mong Njua [6]). We provide evidence in support of Kawasaki's suggestion that coronal-velar contrasts are dispreferred before [l] because of the comparative poverty of place cues in this context.

Some support for this proposal comes from the findings that French listeners usually identify initial [dl] and [tl] as [gl] and [kl] respectively, and have difficulty in discriminating these pairs of clusters [3].

To investigate the acoustic characteristics of coronal stop-lateral clusters we analyzed medial [-dl-] and [-gl-] clusters from three of the speakers of American English from the previous study, following the same methods, except that Burst Peak here refers to the highest spectral peak above f0, since f0 peaks reflect voicing, not place.

Sample burst spectra are shown in fig. 1 and measurements are summarized in table 2. The clusters are differentiated primarily by burst quality. F2 onset is higher on average in [-dl-] than in [-gl-], but this difference is minimal unless the proceeding vowel has a high F2 ([i, ɪ] – words like *Fiedler*, *Ridley*), so it appears that [d] does not itself raise F2 at onset of [l], rather it permits more coarticulatory influence of a preceding front vowel on the [l] compared to [g].

	-dl-		-gl-	
F2 onset (Hz)	1311	(149)	1158	(131)
F3 onset (Hz)	3089	(403)	3044	(233)
Burst durn (ms)	17	(10)	25	(10)
Burst peak (Hz)	2141	(627)	1233	(256)
Amid-Ahi (dB)	12.37	(4.63)	20.42	(5.04)

Table 2: Mean measurements from medial [-dl-, -gl-], 3 speakers of American English.

	d- (*dole*)		g- (*goal*)	
F2 onset (Hz)	1846	(139)	1075	(106)
F3 onset (Hz)	2754	(160)	2758	(177)
Burst durn (ms)	11	(2)	30	(7)
Burst peak (Hz)	3225	(490)	1240	(98)
Amid-Ahi (dB)	3.6	(4.18)	16.11	(6.66)

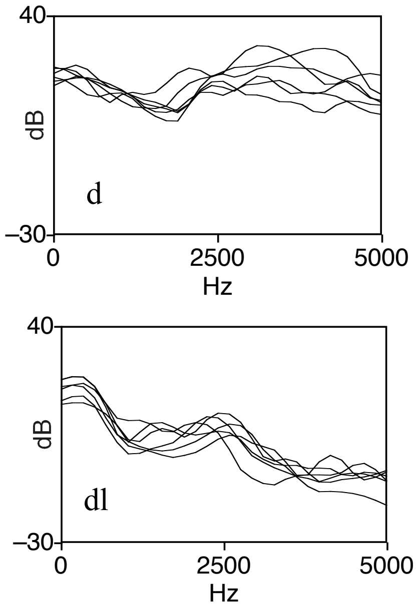
Table 3: Mean measurements from initial [d, g], 5 speakers of American English.

The limited role of formant transitions in distinguishing stops before laterals is supported by a preliminary study of initial [bl-, dl-, gl-] clusters from three speakers of Hebrew, a language that permits coronal-lateral onsets. In all three clusters, F2 and F3 at onset are very close to their values at the mid-point of [l]. The only difference that approaches significance is a slight lowering of F3 after [b], compared to [d, g].

The low F2 onset observed in English [-dl-] clusters is quite striking because alveolar stops are generally characterized by relatively high F2 at release. Even preceding the extreme back allophone of /o/ found in *dole*, F2 onset averaged 1846 Hz. Possibly the low F2 onset in [dl] arises because the transition from [d] to [l] can be made by lateral release–lowering the sides of the tongue body by lateral contraction of the tongue–rather than lowering the tongue tip and retracting the tongue body as in a coronal-back vowel sequence.

Medial [-dl-] and [-gl-] are differentiated by the shape of their burst spectra. The Peak of the [dl] burst is higher than in the [gl] burst, and the difference between Amid and Ahi is smaller. However, the properties of the [dl] burst deviate substantially from prevocalic [d] bursts in ways that render them more similar to [g]/[gl] bursts. This can be observed qualitatively by comparing the representative [d], [dl] and [gl] spectra in Figs. 1 & 2. The [d] bursts are diffuse (lacking any prominent peaks) and relatively flat, i.e. intensity is similar at low and high frequencies, as indicated by the small difference between Amid and Ahi (3.19 dB). The [dl] bursts are more compact, with a peak around 2100 Hz, and a drop in intensity at higher frequencies, reflected in the higher difference between Amid and Ahi (12.37 dB). Velar bursts are typically compact and [gl] bursts follow this pattern, showing a very pronounced peak at around 1200 Hz, a little above F2 in the following lateral. Intensity drops off sharply at higher frequencies, resulting in an average Amid-Ahi of 20.42 dB (initial [gl] is very similar).

Figure 1: Smoothed burst spectra from one speaker: initial [d-] (top), medial [-dl-] (bottom).



So while [-dl-] and [-gl-] bursts differ, [-dl-] bursts deviate significantly from canonical prevocalic [d] bursts and are intermediate between [d] and [gl] on our measures of burst shape, suggesting that the coronal-velar difference should be less distinct in this context. Average measurements from [d] and [g] in *dole* and *goal* are presented in table 3 as the pre-vocalic context most comparable to [l] in F2.

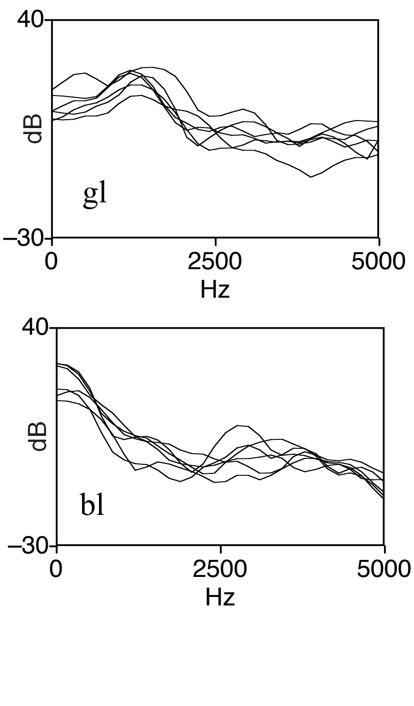
The properties of the [dl] bursts are expected consequences of laterally releasing [d]. The release burst of a centrally-released alveolar has significant high frequency energy because the release burst is filtered by the cavity in front of the closure, and in an alveolar stop this is small and thus has high resonant frequencies. But where the coronal is laterally released, the front cavity includes the side passages opened up by lowering the sides of the tongue body, and thus is significantly longer, resulting in the lower frequency peak and more compact burst shape.

There is also a significant difference in burst duration between medial [gl] and [dl] (25 ms vs. 17 ms). So there are acoustic differences between [gl] and [dl], as we would expect, given that this contrast is maintained in some languages. However, maintaining a three way contrast between [bl, dl, gl] involves distinguishing the three places of articulation in terms of burst quality, which would be likely to result in less distinct place contrasts than can be realized pre-vocalically, or before [r], where significant formant transitions are available.

If one stop place is to be eliminated, the most distinct contrasts should be retained. It is plausible that the [dl]-[gl] contrast is less distinct than the differences between [gl] and [bl]. The [gl]-[bl] contrast seems to involve a larger difference in burst duration ([gl] burst is 2.8 times longer than [bl] in initial position, compared to a ratio of 1.5 for medial [gl] vs. [dl]) and a difference between a compact burst in [gl] vs. a diffuse falling burst in [bl] (fig. 2) whereas both [gl] and [dl] have relatively compact bursts. In addition there are visual cues to the labial closure in [bl].

It is not clear why the [dl]-[gl] contrast is usually neutralized in favor of the velar stop. Possibly velar-lateral clusters are favored because they allow for canonical velar bursts, whereas lateral release of a coronal results in a burst that is atypical for a prevocalic coronal.

Figure 2: Smoothed burst spectra from one speaker. [gl] (top), [bl] (bottom).



4. CONCLUSIONS

A preference for well-cued contrasts provides a basis for understanding cross-linguistic generalizations about the relative markedness of consonant clusters. Before [r], stop place contrasts can be realized by burst and formant cues comparable to those found in pre-vocalic position, so these are cross-linguistically preferred clusters. Formant transition cues are more limited before lateral [l], and stop place contrasts tend to be more limited in this context. In particular, lateral release has a substantial effect on the acoustics of coronal stops, shifting them acoustically closer to velars. This pattern of realization may account for the dispreference for contrasts between coronal and velar stops before laterals.

5. REFERENCES

[1] Dorman, M.F., Studdert-Kennedy, M., Raphael, L.J. 1977. Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent context-dependent cues. *Perc. and Psychophysics* 22, 109–122.

[2] Greenberg, J.H. 1978. Some generalizations concerning initial and final consonant clusters. In: Greenberg, J.H. (ed.) *Universals of Human Language vol. 2: Phonology*. Stanford: Stanford University Press, 243-279.

[3] Halle, P., Best, C.T., Bachrach, A. 2003. Perception of /dl/ and /tl/ clusters. *Proceedings of the 15th ICPhS*, 2893-2896.

[4] Hanson, H.M., Stevens, K.N. 2003. Models of aspirated stops in English. *Proceedings of the 15th ICPhS*, 783-786.

[5] Kawasaki, H. 1982. *An acoustical basis for universal constraints on sound sequences*. Ph.D. dissertation, University of California, Berkeley.

[6] Lyman, T.A. 1974. *Dictionary of Mong Njua.* The Hague.: Mouton.

[7] Mattingly, I.G. 1981. Phonetic representation and speech synthesis by rule. In: Myers, T, Laver, J., Anderson, J. *The Cognitive Representation of Speech*. Amsterdam: North Holland, 415-420.

[8] Mundhenk, A.T., Goschnick, A. 1977. Haroi phonemes. *Papers in S.E. Asian Linguistics* 4, 1-15.

[9] Ohala, J.J. 1992. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. *Papers from the Parasession on the Syllable: Chicago Linguistics Society*, 319-338.

[10] Steriade, D. 1999. Alternatives to syllable-based accounts of consonantal phonotactics. In: Fujimura, O., Joseph, B.D., Palek, B. (eds) *Proceedings of Linguistics and Phonetics 1998*. Prague: Karolinum Press, 205-245.

[11] Suchato, A., Proadpran, P. 2005. Factors in classification of stop consonant place of articulation. *Interspeech 2005*, 2969-2972.

[12] Wallace, J. 1969. Katu phonemes. *Mon Khmer Studies* 3, 64–73.

[13] Wright, R.A. 2004. A review of perceptual cues and robustness. In: Hayes, B., Kirchner, R., Steriade, D. (eds), *Phonetically Based Phonology*. Cambridge: CUP.