Retroflexes and Dentals in the FUL-Model

Aditi Lahiri and Henning Reetz

Department of Linguistics, University of Konstanz, Germany
E-mail: Aditi.Lahiri@uni-konstanz.de, Henning.Reetz@uni-konstanz.de

ABSTRACT

The FUL model (Featurally Underspecified Lexicon) of speech perception claims that the listener extracts features from the acoustic signal and matches them to a phonologically underspecified lexicon. No syllables or segments are computed from the signal nor are word-variants listed in the lexicon. FUL predicts an asymmetry in the mapping between the signal and the mental representation depending on whether the lexical representation under consideration is underspecified or not. Thus [LABIAL] or [DORSAL] from the signal would not mismatch with an underspecified [CORONAL] but the opposite is not true. This holds even for a language like Bengali with two coronal consonants, dental and retroflex, since the feature distinguishing them is not dominated by [CORONAL] and hence not specified. Experimental results support this hypothesis.

1. THEORY

Words are recognised by native listeners with amazing accuracy, in spite of the distortions and variations present across speakers and environments. The FUL model of speech perception ([10], [12], [13], [14]) asserts that the listener is able to make sense of the imperfect acoustic signal by extracting phonological features by broad heuristic means and mapping them onto a lexicon consisting of morphemes whose phonological representation are underspecified. The underspecification is based on universal principles and a language specific system of contrasts [7], [12]. Each word has a unique phonological representation such that no word-variants are listed. The entire speech recognition process, crucially depends on a system of phonological features and the mapping mechanism from signal to the representation. We discuss each in turn.

FUL assumes that the feature inventory is universal. Our feature tree (developed in [11] and elaborated in [7]) crucially assumes that the PLACE node is divided into articulator features and aperture or height features and the PLACE features are shared by vowels and consonants. The height features are not dominated by a place feature like [DORSAL] [8] or within an independent V-place [5]. Since we are primarily concerned with PLACE features in this paper, a partial tree with the crucial features are shown in Fig. 1.

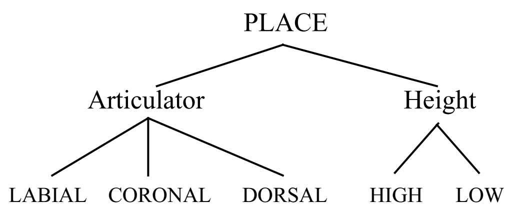

Figure 1: Partial feature tree

Within the FUL model we assume that phonological features are extracted from the speech signal by heuristic methods. For instance (depending on the language) [HIGH] is extracted on the basis of a low F1 (e.g. < 450Hz). The features are then directly compared with all entries in the lexicon without segmenting words into segments, syllables, or spectral templates. The mapping from the features in the signal to the lexicon is based on a ternary logic: match, nomismatch and mismatch. This ternary matching constrains the list of word candidates which are fed directly into the phonological and syntactic parser. All words in the lexicon are activated with matching and nomismatching features; as mismatching features are encountered, words are deactivated. Thus word candidates are expanded to possible word hypotheses, even without acoustic evidence and the degree of match is scored [12]. When a feature like [LABIAL] is extracted from the signal and does not find a relevant match (a nomismatch situation), but does not mismatch either, the word remains activated. With the extracted [LABIAL] from a word like [bæn], all words beginning with a /b/ as well as all words with a /t/, where the feature [CORONAL] is underspecified, are activated. The /b/ words get a higher score than the nomismatch /t/ cases.

Features from the signal mismatch with the features in the representation by universal principles. For instance, [CORONAL] mismatches with [DORSAL] and [LABIAL]. But [DORSAL] and [LABIAL] have a nomismatch mapping with an unspecified [CORONAL]. Mismatch is defined independently for vowels and consonants. Since [SONORANT] is specified, it makes the initial cut between obstruents and vowels. The match and mismatch conditions relevant for the present study are given below.

Signal		Representation
[CORONAL]	*MISMATCH*	[LABIAL], [DORSAL]
[DORSAL]	*MISMATCH*	[LABIAL]
[LABIAL]	*MISMATCH*	[DORSAL]
[LABIAL], [DORSAL]	*NOMISMATCH*	[CORONAL]

As a case study, we focus on the contrast between dental and retroflex consonants in Standard Colloquial Bengali (henceforth Bengali, spoken in Kolkata, India) which are both articulated with the front part of the tongue. The retroflex stop is articulated with the tongue tip curled slightly and raised to touch just behind the alveolar ridge. The retroflexion is not as much as in the Dravidian languages ([3] p. 268). The dental stop is produced with the tongue tip pressed flat against the teeth. We focus here on the phonological contrast between the dental and retroflex stops - their distribution, their representation in terms of phonological features, and the way in which they are processed.

ISBN 1-876346-48-5 © 2003 UAB

1.1 Distribution of dentals and retroflexes in Bengali

Bengali has four stops in each of the dental and retroflex places of articulation: [t̪ t̪ʰ d̪ d̪ʰ ʈ ʈʰ ɖ ɖʰ] (cf. [3]). All of the consonants occur word initially: [t̪ɑk, t̪ʰɑn, d̪in, d̪ʰɑn, ʈɑk, ʈʰik, ɖɑl, ɖʰɑk] 'shelf, white sari, day, grain, bald head, correct, lentils, type of drum' and as geminates word medially except for [ɖʰ]: [ʃot̪t̪i, pot̪t̪ʰo rod̪d̪ur, bɑd̪d̪ʰo, ɑʈʈɑ, kɔʈʈʰor, bɔɖɖo] 'true, medicine, sunlight, obedient, eight, difficult, very'. However as singletons, neither of the voiced retroflex consonants occur medially or finally. Singleton medial [ɖ ɖʰ] are rhotacised: [bɔɽo]. All the other consonants occur both medially ([pɑt̪ɑ, ʃɑt̪ʰi, god̪i, bɑd̪ʰɑ, ɑʈɑ, lɑʈʰi] 'leaf, companion, mattress, block, wheat, stick') and finally ([hɑt̪, pɔt̪ʰ, rod̪, bɑd̪ʰ, gʰɑʈ, mɑʈʰ] 'hand, path, sunlight, prevention, bank, field').

Both voiceless unaspirated stops [t̪] and [ʈ] are used as grammatical morphemes, but [ʈ] is also used as the initial prefix consonant of an extremely productive reduplication process. The retroflex [ʈ] is the pre-specified initial consonant for all reduplications above a word [6]: noun [pɑt̪ɑ] 'leaf', [pɑt̪ɑ ʈɑt̪ɑ] 'leaves and such'; verb [dækʰɑ] 'see-VERBAL NOUN', [dækʰɑ ʈækʰɑ] 'see and such'. Reduplications of this type are very common as in most Indo-Aryan languages and hence [ʈ] occurs rather frequently. The voiceless [ʈ] also appears in suffix morphemes. Bengali does not have determiners but it has a specifier particle [ʈɑ]: [boi, boi-ʈɑ] 'book, the specific book'. The voiceless dental stop [t̪] is used only in suffixes like the habitual past marker ([dækʰɑ, dekʰ-t̪-ɑm] 'see VERBAL NOUN, see-HABITUAL PAST-1P') and the conjunctive participle ([dækʰɑ, dekʰ-t̪-e] 'see VERBAL NOUN, see-CONJUNCTIVE'). Thus, although the dentals occur initially in many more words than the retroflexes (e.g. percentage count from Chatterji based on 6 texts: [t̪] 3.8, [ʈ] 0.7), due to the productivity of the reduplication process and the specifier morpheme, the sound [ʈ] is rather predominant in the language. Another phenomena which increases the total number of retroflexes are loans, where Bengali always uses the retroflex stops for English alveolars: [ʈɑim, ɔʈo, bæʈ, ɖiʃ, meɖel, pæɖ] 'time, 3-wheeler taxi, bat, dish, medal, pad'. In fact, contrary to the regular restrictions, the voiced unaspirated retroflex can occur word finally in borrowed English words. English [θ ð] are adapted as [t̪ʰ d̪]: [t̪ʰiŋk d̪en] 'think, then'. Consequently, although the retroflex stops are more marked than the dentals in terms of their distribution across words and the number of words in which they occur, the actual frequency of occurrence of the two places of articulation may well be equal.

1.2 Acoustics of dental and retroflex consonants

A detailed account of the acoustic characteristics of dental and retroflex stops is provided in [15] for Hindi speakers. Dental stops have a falling F2, F3, and F4 while the retroflex stops have a falling F2 but a sharply rising F3 and F4. Additionally, the burst energy is higher for the dentals (around 3.3 kHz) than for the retroflex (between 2.7 and 3 kHz). In the synthetic stimuli they used, a burst around 4 kHz was required for the dentals to be clearly identified.

1.3 Patterns of alternations and features

Dental and retroflex stops trigger an alternation in the sibilant. Bengali has only one sibilant in the language with the default surface form [ʃ]. Medially in the context of a dental stop, the sibilant is a dental: [ɑʃ-ɑ, ɑʃ-b-e, ɑʃ-t̪-ɑm > ɑs̪-t̪-ɑm] 'come-VERBAL NOUN, come-FUTURE-3P, come-PAST-1P'. In the context of a retroflex and all other stops the sibilant surfaces as [ʃ]: [kuʃʈi] 'horoscope' vs. [kus̪t̪i] 'wrestling'. In an gating experiment [9], we argued that the posteriority feature extracted from the signal of a sibilant in the sequence [kuʃ-], encouraged listeners to choose [kuʃʈi] as a more likely continuation than [kus̪t̪i]. We concluded that the retroflex was specified as [-anterior] dominated by [CORONAL] (similar to SPE, [4]), the same features that were present in the surface form of [ʃ]. This feature configuration, however, does not fit Bhat's ([1], [2]) description of retroflexion which suggest that dentals become retroflexes quite often in the context of the high back vowel [u]. Furthermore, it also fails to explain Panini's well-known *ruki*-rule of retroflexion states that the front dental [s] becomes a retroflex [ʂ] in the context of the segments [r u k i]. It is somewhat daunting to group these sounds under one feature. Whatever it is, it cannot be a front/back distinction given that both [u] and [i] are part of the context. One possibility that fits both Bhat's generalisation and the *ruki*-rule is that the context is HIGH. If so, then possibly the retroflex is also [HIGH], which would fit the previous gating results.

Lahiri & Evers suggested that [HIGH] and [LOW] were used contrastively usually to distinguish secondary articulations in consonants, where palatalized consonants [kʲ, pʲ] could be captured by the primary place feature plus the feature [HIGH]. The dorsal [k] can have default [HIGH] in a language which does not contrast [kʲ] with [k]. The stops and affricates in Bengali other than the dentals and retroflexes include the labials [p pʰ b bʰ], palatoalveolars [tʃ tʃʰ dʒ dʒʰ] and dorsals [k kʰ g gʰ]. The palatoalveolar affricates are [CORONAL] like the dentals but are differentiated by [STRIDENT]. The retroflexes are, however, different. Although the point of contact is close to the palatoalveolar affricate, if we consider the acoustic characteristics as well as the phonological alternations like the *ruki*-rule and assume that the retroflex is [HIGH], the representation would look as in (2) with the unspecified features in gray.

Here, both dental and retroflex stops are [CORONAL], the contrastive feature being [HIGH]. For both places of articulation [CORONAL] remains unspecified in the lexical representation but can be extracted from the signal.

	ARITCULATOR			HEIGHT	
	LAB	COR	DOR	HIGH	LOW
/p/	√				
/t̪/		√			
/ʈ/		√		√	
/k/			√	√	

Figure 2: Featureal representation of some consonants

2. SIGNAL TO REPRESENTATION

2.1 Hypotheses

Given these representations, our hypotheses are that there is an asymmetric relationship between dentals and retroflexes versus labials and dorsals. Recall that in FUL, any feature extracted from the signal will activate all word candidates with which this feature does not mismatch. A word like [tʃʰɑt̪] 'roof' or [tʃʰɑʈ] 'spray of water' will not activate [tʃʰɑp] 'print, stamp' since [CORONAL] from the signal will mismatch with the [LABIAL] in the representation. But potentially, [tʃʰɑp] will permit the activation of both [tʃʰɑt̪] and [tʃʰɑʈ] since [LABIAL] from the signal does not mismatch with the unspecified [CORONAL]. To test this hypothesis and to ensure that we tapped the lexicon, we chose an indirect semantic priming task to avoid any hint whatsoever of a phonetic priming. We based the task on previous experiments which have shown systematic priming of semantically related words [12]. The logic behind a lexical decision task with semantic priming is explained below with English examples:

HEAR/SEE ↓				LEXICAL DECISION ↓	
Prime	Phon. Repr.	Meaning		Target RT	
PATH *Test*	/pɑθ/	place to walk on	*facilitates*	ROAD	(a)
CAKE *Control*	/keɪk/	baked dessert	*does not facilitate*	ROAD	(b)

Reaction time difference RT (b) - RT (a) gives the amount of facilitation, if the difference is significant. This design anticipates that when PATH is encountered, ROAD is activated via the meaning of PATH right away, such that on then hearing or seeing ROAD the listener is very fast to decide that it is a word. CAKE, on the other hand, does not activate ROAD in any way, and hence the reaction time to ROAD is slower than when PATH was encountered.

We used the same logic in our experiment except that for each set there were a pair of Test items and one Control (see Fig. 3). The hypothesis was that Test item A was directly related semantically to the Target word. The Test item B (which was played to a different subject) had phonologically a *nomismatch* or *mismatch* relationship with Test item A. In turn, the Test item B would in the nomismatch condition (set 1) activate the Test item A, but in the mismatch case (set 2) the Test item A would not be activated. Consequently, the Test item A would prime (nomismatch condition) or not prime (mismatch condition) the Target word semantically related to the Test word A.

The claim is rather strong. Under FUL, the assumption is that on hearing [tʃʰɑp] 'print', all words which phonologically have a *nomismatch* relationship remain activated. This includes the phonologically underspecified [tʃʰɑʈ] 'spray', since the [LABIAL] extracted from the signal does not mismatch with the unspecified [CORONAL]. And along with the phonological activation, the semantic representation would be activated as well. Thus, [dʒɔl] 'water' would be facilitated by [tʃʰɑp] via [tʃʰɑʈ]. There is however a built

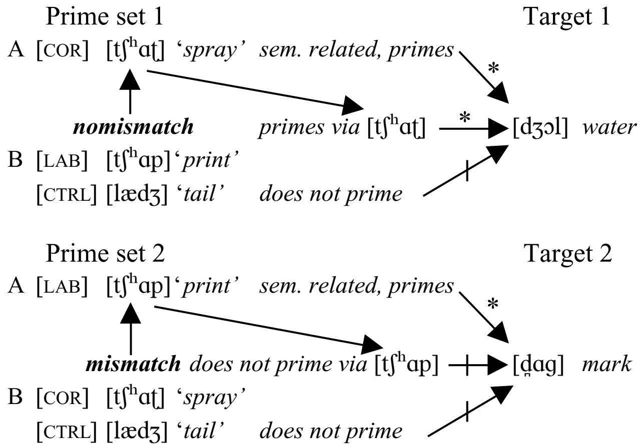

Figure 3: Direct and indirect activation (or non-activation) of target words with pairs of priming sets.

in asymmetry. On hearing [tʃʰɑʈ], [CORONAL] extracted from the signal prevents the activation of any word with a [LABIAL] in the same position, such that [tʃʰɑp] 'print' is not considered and its semantically related word [d̪ɑg] 'mark' is not facilitated. Thus the asymmetry in perception predicted by FUL is due to a combination of the underlying lexical representation and the ternary matching logic from the signal to the representation.

In the experiment, we matched pairs of dentals vs. labials/dorsals as well as retroflex vs. labials/dorsals. That is, in addition to the 2 sets of Prime/Target with retroflex stops vs. labial/dorsal above, we had a corresponding dental set.

2.2 Material, method and subjects

We chose a cross-modal priming task with presentation of the targets as written words to realise the experimental paradigm of Fig. 3. 48 prime-sets were distributed across three subject groups so that each subject saw each target only once. Controls, direct and indirect related primes as well as the relevant places of articulation were distributed equally across the groups. 32 filler and 128 non-word pairs were added to the randomised sequences in each group.

The primes were presented via headphones and at their offset the targets were displayed on a white wall in Bengali script with a video-beamer. The subjects were instructed to press a 'word' or 'non-word' button on a box placed in front of them as soon as they saw the target words. 40 male and 40 female students were paid for their participation. They were aurally instructed and performed a short practice session prior to the actual experiment.

2.3 Results

Data of 3 subjects were removed from the analysis because they made more than 20% errors in the word - nonword decision. From the remaining 77 subjects only the correct word-reactions faster than 1500 ms for the test and control prime and target pairs were extracted. From the 96 target words those which did not show a priming in the direct semantic condition were removed from further analyses, which left 66 prime-target pairs with a total of 4451 reactions.

ISBN 1-876346-48-5 © 2003 UAB

An analyses of variance (ANOVA) of the reaction times as dependent variable and the subjects as random factor, dental/retroflex set, prime-target relation, prime set, and target words nested under dental/retroflex and prime set with maximal factorisation was performed. Fig. 4 shows the differences of reaction times against the controls for the different conditions, and an asterisk indicates significant priming on the 5% level. The first pair of columns show that dentals do not prime labials or dorsals, whereas the second pair of columns show that labials and dorsals prime dentals, as predicted. The same is true for the retroflexes. Here, the priming is significant but the difference to the controls for the labial and dorsal primes is less, which will require further investigation.

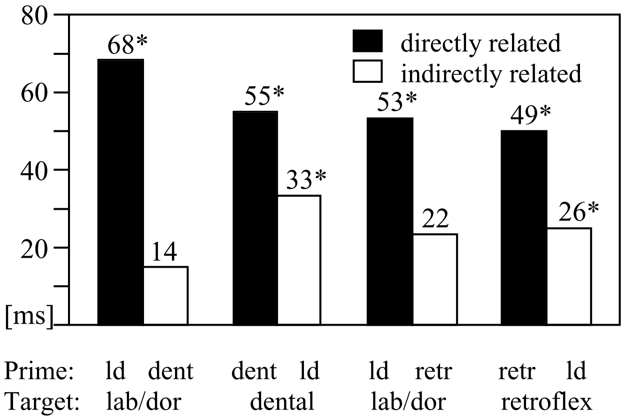

Figure 4: Milliseconds facilitation against the controls. Asterisks (*) mark significant facilitations.

3. CONCLUSION

FUL assumes that all PLACE features mismatch each other. However, it also predicts that if a PLACE is not specified in the lexicon, other PLACE features from the signal cannot mismatch and therefore word candidates with the unspecified feature will also be activated. This is predicted even when there are two consonants underspecified for [CORONAL], dentals and retroflexes. Our experimental results with Bengali data confirmed the hypothesis. Labial and dorsal stops primed lexical dentals and retroflexes but not vice versa. The asymmetry in perception predicted by FUL-model assumes that the even if the features are not present in the representation, they can be extracted from the signal. This of course, can only happen if the feature is indeed present in the signal. We have examined here dentals and retroflexes in the final position, where the release of the consonant is not as clear as in the initial position. Word initially, where the voiceless unaspirated [ṭ] has a predominant role in the language, the two places of articulation may be more distinct. The feature [HIGH] for the retroflexes may play a bigger role in the nomismatch/ mismatch conditions in the initial position. Although the articulators mismatch in the retroflex and dorsals, [HIGH] matches. We need to examine whether there is a distinction in the between the labials vs. dorsals compared to the retroflex stops.

ACKNOWLEDGEMENTS

This research was partly supported by the DFG which funds the SFB 471 and the Leibniz Prize. We are extremely grateful to the Principal and Faculty of Lal Baba College, Belur, and particularly to Dr. Kalpana Sen Barat for all the organisation and planning, and for allowing us to run the experiments in the college premises.

REFERENCES

[1] D.N.S. Bhat, "Retroflexion: an areal feature," in *Working Papers on Language Universals,* vol. 13, pp. 27-67, 1973

[2] D.N.S. Bhat, "Retroflexion and retraction," in *Journal of Phonetics,* vol. 2, pp. 233-237, 1974.

[3] S.K. Chatterji, *The Origin and Development of the Bengali Language*, Calcutta: Rupa and Co, 1975. First Edition 1926 by Calcutta University Press, Calcutta.

[4] N. Chomsky & M. Halle, *The Sound Pattern of English*, New York: Harper and Row, 1968.

[5] G.N. Clements & E. Hume, "The internal organization of speech sounds," in *The Handbook of Phonological Theory*, J. Goldsmith, Ed., pp. 245-306, London: Blackwell, 1995.

[6] J. Fitzpatrick-Cole, "Reduplication meets the phonological phrase in Bengali," in *The Linguistic Review*, vol. 13, pp. 305-356, 1996.

[7] M. Ghini, *Asymmetries in the Phonology of Miogliola*, Berlin: Mouton, 2001.

[8] M. Halle, "Feature geometry and feature spreading," in *Linguistic Inquiry*, vol. 26, pp. 1-46, 1995. Reprinted in *From Speech to Memory and Back*, M. Halle, Ed., pp. 196-252, Berlin: Mouton, 2003.

[9] A. Lahiri, "Anteriority in sibilants," in *Proc. of the 12th International Congress of Phonetic Sciences*, vol. 1, pp. 384-388, Aix-en-Provence, 1991.

[10] A. Lahiri, "Speech recognition with phonological features." *Proc. of the 14th International Congress of Phonetic Sciences* 1, 715-718, San Francisco, 1999.

[11] A. Lahiri & V. Evers, "Palatalization and coronality," in *The Special Status of Coronals*, C. Paradis & J.F Prunet, Eds., pp. 79-100. San Diego: Academic Press, 1991.

[12] A. Lahiri & H. Reetz "Underspecified Recognition," in *Laboratory Phonology 7*, C. Gussenhoven and N. Warner, Eds., pp. 637-675, Berlin: Mouton, 2002.

[13] H. Reetz, *Automatic Speech Recognition with Features*, Habilitationschrift, Saarbrücken. 1999.

[14] H. Reetz, "Converting speech signals to phonological features," in *Proc.of the 14th International Congress of Phonetic Sciences* 3, 1733-1736, San Francisco, 1999.

[15] K.N. Stevens & S.E. Blumstein "Quantal aspects of consonant production and perception: a study of retroflex stop consonants," in *Journal of Phonetics*, vol. 3, pp. 215-233, 1975.