TEMPORAL PATTERNS IN SYRIAN ARABIC VOICING ASSIMILATION

Martin Barry† and Ryad Teifour*
†Department of Linguistics, The University of Manchester, UK,
**Department of English, Al-Baath University, Homs, Syria*

ABSTRACT

Acoustic analysis of CC sequences in colloquial Syrian Arabic shows that voicing assimilation is a gradient phonetic phenomenon correlated with speaking rate. Clusters in which the voicing values for the two consonants are opposed undergo regressive assimilation, to a greater extent at faster speaking rates. In [+voi][-voi] sequences devoicing extends progressively 'leftwards', i.e. as speaking rate increases, voicing in C_1 ceases at progressively earlier points in time; while in [-voi][+voi] sequences C_1 undergoes voicing which extends 'rightwards', i.e. as rate increases, carryover voicing from a preceding vowel increases in duration during C_1, leaving a progressively shorter voiceless portion of C_1 immediately prior to C_2. The complexity of the temporal patterns observed point towards a revision of the Window Model of coarticulation involving explicit representation of the time domain in the specification of window width. Both syllabic position and speaking rate may influence the window size for a given speech sound.

1. OBSTRUENT DEVOICING AND VOICING ASSIMILATION

The aim in this paper is to examine time-domain patterns in Syrian Arabic obstruent sequences which are said to undergo assimilation of voicing, and to consider the implications of these and other data for models of speech production. While the literature concerning voicing assimilation is relatively slender, a number of studies in the last fifteen years of the related phenomenon of final devoicing of obstruents have shown that some instances of so-called 'neutralisation' of the voicing contrast in final position in fact display phonetic characteristics which indicate that an underlying contrast may still be discerned in close scrutiny of one or more temporal parameters associated with the voicing contrast. It has been argued [1, 2] that parameters such as consonant duration, duration of glottal pulsing and duration of preceding vowel pertaining to a final 'devoiced' (underlyingly voiced) consonant may show values distinct from those observed in final consonants which are underlyingly voiceless. The issue is controversial, and a number of sources (cf. [3]) suggest that reported instances of so-called incomplete neutralisation may in fact be accounted for by a range of factors pertaining to the experimental method such as the elicitation procedure and the choice of lexical and textual materials, as well as by variability within and among speakers.

Assimilation of voicing occurs in a cluster of two consonants which differ underlyingly in terms of voicing, when one of the two consonants adopts the value for voicing for the other consonant. The process may be either anticipatory or perseverative in character, and typically involves a sequence of two obstruents. In some cases, however, a consonant which in other respects patterns as a sonorant may be susceptible to assimilation to the voicing value of an adjacent obstruent. This is the case for Russian /v/, which may undergo assimilation before a voiceless obstruent, but does not itself trigger voicing of a preceding voiceless obstruent [4]. Empirical investigation of the phenomenon in Russian [3] confirmed the standard account of the process as anticipatory (or regressive) in character, and identified some time- and frequency-domain characteristics of assimilated tokens as different from those observed in underlyingly homogeneous sequences, thus giving support to the claim that in processes such as this the voicing contrast is neutralised only incompletely. One of the parameters measured was the duration of voicing into the C_1 stop closure or fricative phase in CC clusters showing voicing assimilation, in which small (non-significant) differences were observed for underlyingly voiced and voiceless C_1.

Obstruent voicing assimilation has also been studied in other languages, e.g. Dutch [5] and Catalan [6]. In the latter study the assimilation process is held to be complete rather than incomplete.

2. VOICING ASSIMILATION IN SYRIAN ARABIC

A recent study [7] investigated the phonetic facts of voicing assimilation in obstruent clusters in Syrian Arabic. In informal accounts two parallel phenomena are said to occur, namely regressive voicing and regressive devoicing by assimilation across word boundaries. An example of voicing is given in (1) and an example of devoicing is given in (2) below:

(1) /ʃtareːt daːr/ → [ʃtareːt daːr] "I bought a house"

(2) /lwalad taːni/ → [lwalat taːni] "another child"

Table 1: *Voice assimilation in Syrian Arabic*: Mean percentages and standard deviations (within brackets) of voicing into closure pooled across all speakers, in both voicing and devoicing, in all three speaking rates. *Source: [7]*

Sp. rate	*Voicing* *underl vls*	*vd by ass* (N=5)	*underl vd*	*Devoicing* *underl vd*	*dev by ass* (N=5)	*underl vls*
SS	5.7 (0.8)	57.7 (11.9)	72.6 (10.4)	80.4 (7.2)	18.1 (4.5)	6.5 (0.8)
NS	10.8 (2.8)	92.0 (11.2)	95.4 (5.7)	92.7 (4.5)	17.1 (4.5)	8.4 (6.5)
FS	12.6 (2.4)	96.7 (6.0)	99.8 (0.5)	95.4 (10.9)	17.5 (8.0)	11.9 (2.3)
$\bar{x}$	9.7 (3.6)	82.1 (21.3)	89.3 (14.6)	89.5 (8.0)	17.6 (4.9)	8.9 (3.3)

Table 2: *Voice assimilation in Syrian Arabic*: Mean percentages of voicing into **friction** and standard deviations (within brackets) pooled across all speakers in both voicing and devoicing, in all three speaking rates. *Source: [7]*

Sp. rate	*Voicing* *underl vls*	*vd by ass* (N=5)	*underl vd*	*Devoicing* *underl vd*	*dev by ass* (N=5)	*underl vls*
SS	7.5 (1.9)	17.4 (2.8)	76.7 (4.7)	78.6 (3.7)	17.1 (1.7)	8.2 (1.5)
NS	12.0 (3.8)	77.2 (14.7)	91.6 (11.8)	95.9 (9.2)	15.9 (3.5)	10.7 (3.1)
FS	12.7 (2.0)	94.1 (6.2)	96.4 (5.1)	95.9 (3.9)	14.8 (3.9)	12.6 (2.1)
$\bar{\bar{x}}$	10.7 (2.8)	62.9 (40.3)	88.2 (10.3)	90.1 (10.0)	15.9 (1.2)	10.5 (2.2)

The study formed part of a larger investigation into the description and modelling of connected speech processes in Syrian Arabic, encompassing aspects of overlap and assimilation in supraglottal articulatory activity as well as the laryngeal features discussed further in the present paper. The larger study also aimed to explore aspects of assimilation sensitive to speaking rate and speaking style, since assimilatory phenomena of the type under discussion are typically evidenced in informal and rapid speech rather than in the more formal and monitored style of slow careful speech.

3. METHODS

Test materials were prepared containing instances of five voiceless stops, five voiced stops, eight voiceless fricatives and eight voiced fricatives, each in a potential 'assimilation site', that is, followed (across a word boundary) by an obstruent having a contrary value for the voicing feature. In addition, control materials were devised containing voiced-voiced and voiceless-voiceless sequences in similar environments. All the test sequences were embedded in carrier sentences, and were placed in similar syntactic and prosodic environments.

Five subjects, native speakers of urban varieties of Syrian Arabic, read the test materials at self-selected 'slow' and 'fast' rates of speech. In addition, a third series of data was obtained at normal colloquial speaking rate, by the use of a simple word-game technique in which the experimenter (the second author) read sentences to each subject and asked them to alter the meaning of a given word in the sentence (not one bearing a test sequence) and to repeat the sentence with this modification. This technique served a twofold purpose: to elicit speech at a rate intermediate between the more extreme values obtained in the reading task, and to ensure a speaking style less prone to the charge levelled against other similar studies that the phenomena observed are experimental artifacts brought about by a dependence solely on read materials.

Recordings were made using a Casio DA-7 DAT recorder with an AKG C1000S microphone placed around 10 cm from the mouth of the subject. The recorded utterances were digitised at 16 bits and 44 kHz, and analysed using GWI SoundScope running on an Apple Macintosh in the phonetics laboratory at the University of Manchester. A number of time-domain measures were extracted in the wider study. Of interest to this paper was the measurement of the duration of voicing into the C_1 closure or fricative phase, defined as the interval observable in the acoustic signal between the vowel offset and the cessation of periodicity. Statistical analysis was carried out (the techniques employed being ANOVA plus the Tukey Test for pairwise comparisons between means) using SPSS.

4. RESULTS

Table 1 shows the mean percentages of voicing in C_1 stops for the various conditions, pooled across all subjects. Table 2 shows the same information for fricatives.

The results here show that for all speaking rates the mean percentage duration of voicing in C_1 in an underlyingly heterogeneous CC cluster (the assimilation site) is intermediate between the expected value for C_1 in either a voiced-voiced or a voiceless-voiceless cluster. The difference between C_1 voicing in the assimilation site and in the voiced-voiced cluster or voiceless-voiceless cluster (whichever represents the cluster that would be created by complete assimilation taking effect at the assimilation site) was significant at the slow speaking rate alone for the stops and non-significant at the normal and fast rates; and significant at the slow and normal rates for the fricatives, and non-significant for fast speech. The results thus closely mirror the findings for Russian [3], suggesting at least some evidence that assimilation of voicing operates in a manner yielding only an incomplete neutralisation of the voicing contrast at slower speaking rates.

5. THE NON-CONTIGUOUS NATURE OF VOICING ASSIMILATION

Of significance to the discussion here is that the data showed evidence of a spread of values for voicing into the C_1 stop closure or fricative phase. If voiced-voiceless clusters are ordered according to the extent of voicing in C_1, a continuum of forms emerges in which voicing ceases progressively later; that is, devoicing may be argued to be spreading leftwards from the (underlyingly voiceless) C_2. For voiceless-voiced clusters, however, although a similar pattern emerges (shown schematically in figure 1), the interpretation is less straightforward, since in this continuum what we observe is the progressive spread rightwards (i.e. ostensibly from the preceding vowel) of voicing into C_1. Thus the process of voicing by assimilation appears to involve the spread of a phonetic property that might be termed *non-contiguous*, in that the process by which C_1 acquires voicing is triggered by a voiced C_2, but a voiceless portion of C_1 (of variable duration, and shorter in more completely assimilated tokens) always intervenes between the two voiced portions.

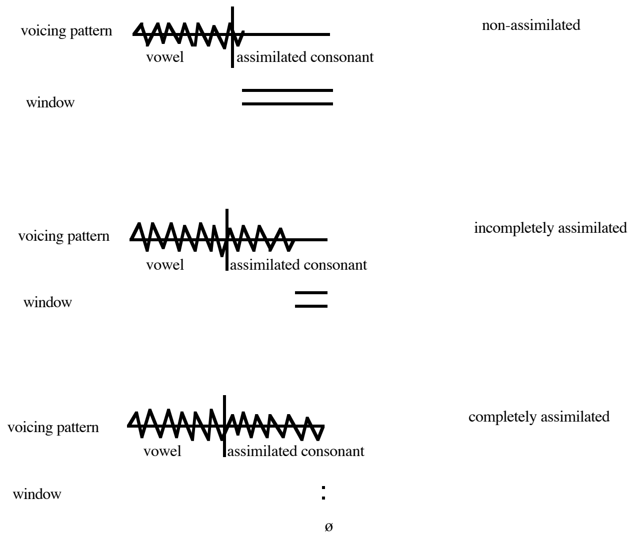

Figure 1. Schematic illustration of progressive rightward spread of voicing as speaking rate increases, in CC clusters with underlyingly voiceless C_1. (C_2 is not shown.)

6. CONSTRAINTS ON VARIABILITY IN THE COARTICULATORY WINDOW FOR VOICING

This non-contiguous character of the spread of voicing in a CC cluster presents a problem for any theory of speech production or of sound patterning which seeks either to provide an account rich in low-level detail of the phonetic character of the spreading, or to provide an explanation for the phenomenon grounded in what has been termed ‘the inertial properties of the speech apparatus’ [8: 297]. One such model is the Window Model of coarticulation [9], which proposes that articulator paths (or equivalently, though less concretely, acoustic paths such as formant trajectories) are specified not by monotonic interpolation between fixed target points specified for each ‘segment’ in a sequence of sounds, but by specifying a ‘window’, that is a range governing all the observed variation for a particular articulatory parameter for a given speech sound. The articulator path is then generated not by interpolation but by the tracing of a path within limits specified by the window for successive segments. The mechanisms governing how the precise articulator path might be determined are unclear, but the implication in the model is that the speaker has a certain amount of freedom within a given window, mirroring the fact that the window is itself determined empirically from observations of variability in speaker behaviour. The author of the model leaves open the question whether values within a given window have a flat distribution, or whether the window is subject to some function indicating the probability of the articulator path taking any particular value.

For the voicing assimilation phenomena discussed in this paper a version of the model is proposed which is illustrated in figure 2. On the vertical axis voicing is specified in a notional space with two states (vibrating and non-vibrating vocal folds), following Keating’s suggestion that the model may specify speaker behaviour at a more abstract level than the movements of specific articulators. The horizontal axis incorporates the time domain, again in a manner consistent with the original model, in which it is left open whether the horizontal axis reflects real time (in some sense) or merely sequential ordering among segments. In the present proposal the observed details of timing on which the discussion hinges require the assumption that the horizontal axis incorporates real time, at least in that windows of different horizontal extent must be taken to have different temporal extent; and in that windows situated closer together horizontally in the representation must relate to events that are closer together in time in the utterance being modelled.

Figure 2 also illustrates the range of tokens observed for the voiceless-voiced sequences. ‘Voicelessness’ in underlying C_1 is specified by a window of narrow ‘width’ to use Keating’s term for *vertical* extent within the representation, stipulating the non-vibrating glottal state. As tokens become successively more subject to assimilatory voicing, the duration, that is, *horizontal* extent within the representation progressively decreases, leaving successively shorter voiceless portions within C_1, but (as the diagram indicates) temporally adjacent to C_2. The limiting case is the reduction in the duration of the window for voicelessness to zero duration (indicated by the Ø symbol in figures 1 and 2), whereupon the 'path' for voicing follows only the windows for the surrounding vowels and the cluster C_2, all of which (in this simplified representation) show vocal-fold vibration.

Thus crucially this account proposes two modifications to the Window Model as originally set out [9]. Firstly, it is proposed to build time explicitly into the representation by allowing windows to differ one from another by virtue of their temporal extent in the representation. This allows us to model the progressive encroachment of voicing (and also of the simpler case of devoicing) in C_1 as tokens show more evidence of assimilation. Secondly, we propose a significant departure from the nature of the 'window' as originally defined in the model. In Keating's formulation, the range of values specified by a window of given width serves to encompass the full range of variability encountered in the behaviour of a given articulator path (or other parameter trajectory), regardless of the source of this variability. It is then left to mechanisms underdetermined by the model to generate a particular path in a given utterance. This is evidently potentially a weakness in the model, since phonetic variability, which the model is devised precisely to account for, becomes entirely extrinsic to the representation. Our proposal is instead to disaggregate the body of factors governing variability in articulatory trajectories, by expressly identifying particular contextual factors as having a predictable effect on window widths. Two such factors proposed in this study are speaking rate and syllable-position. Both a faster speaking rate and the coda syllable position are phonetic environments associated with weakening of articulatory gestures, and it seems to us appropriate to adapt the window model so as to show global reduction effects on window sizes for segments in these environments. In the data under consideration in the present study, this allows us to predict a reduction in magnitude of the temporal extent of the window specifying the non-vibrating glottal state in the underlyingly voiceless C_1: voicing assimilation takes place because the 'voicelessness' of C_1 is reduced as a consequence of its weak syllabic position and its utterance in a speaking style more conducive to reduction of this kind.

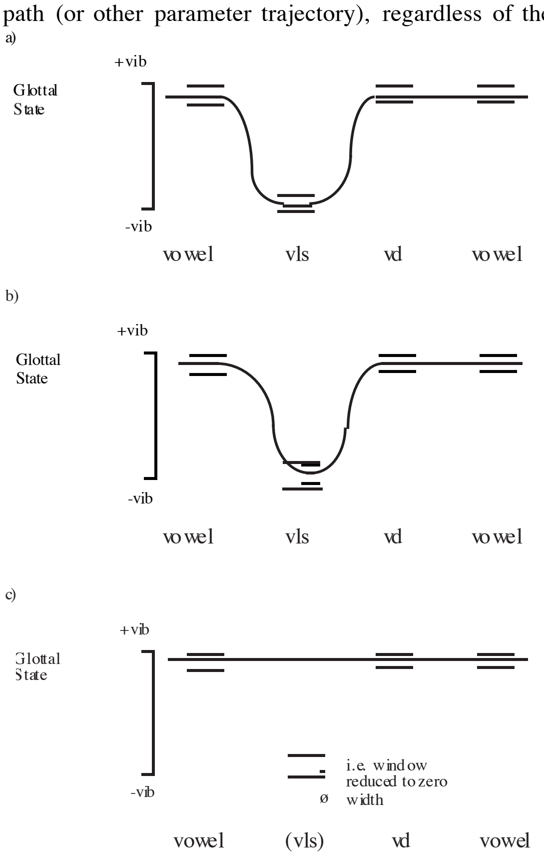

Figure 2. Sequence of schematic amplitude of voicing windows for vls-vd clusters showing (a) no assimilation together with the surrounding vowels, (b) incompletely assimilated and (c) completely assimilated (with reduction of window width to zero).

The modification we propose to the model may allow identification of additional factors having some global predictable effect on window width. The consequence, it is proposed, is a theory with greater explanatory force, in that it points to a move away from a version of the Window Model in which all the substantive information as to window width, formally extrinsic to the representation, is established inductively and simply plugged into the model, towards a theory in which window width is at least in part determined by factors which may be represented explicitly within the formalism of the model. Further work, inevitably, is needed in order to explore the potential for developments of this nature.

REFERENCES

[1] Port, R.F. and O'Dell, M. 1985. Neutralization of syllable-final voicing in German. *Journal of Phonetics*, 13, 455–471.

[2] Port, R.F. and Crawford, P. 1989. Incomplete neutralisation and pragmatics in German. *Journal of Phonetics*, 17, 257–282.

[3] Burton, M.W. and Robblea, K.E. 1997. A phonetic analysis of voicing assimilation in Russian. *Journal of Phonetics*, 25, 97–114.

[4] Jones, D. and Ward, D. 1969. *The Phonetics of Russian.* Cambridge, UK: Cambridge University Press.

[5] Slis, I. 1986. Assimilation of voice in Dutch as a function of stress, word boundaries and sex of speaker and listener. *Journal of Phonetics*, 14, 311–326.

[6] Charles-Luce, J.C. 1993. The effects of semantic context on voicing neutralization. *Phonetica*, 50, 28–43.

[7] Teifour, R. 1997. Some phonetic and phonological aspects of connected speech processes in Syrian Arabic. PhD thesis, The University of Manchester, UK.

[8] Lass, R. 1984. *Phonology: An introduction to basic concepts*. Cambridge, UK: Cambridge University Press.

[9] Keating, P. 1990. The window model of coarticulation. In Kingston, J. and Beckman, M. (eds.) *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge, UK: Cambridge University Press.