2012-04-25 J.-Y. Choi

**Standard generation of landmarks and non-oral events**

**from underlying phoneme contexts**

1. **Landmarks**
2. We will assume that landmarks are generated at the locations where information about underlying phonemes is most salient. These locations are at maximum and minimum articulatory points for vowels and glides, and at abrupt changes for consonants (when changes occur in the articulator-free features). The 6 articulator-free features in the distinctive feature representation of speech sounds are: [vowel], [glide], [consonantal] (or [cons]), [sonorant] (or [son]), [continuant] (or [cont]), and [strident] (or [strid]). These may be used to indicate the different manner classes of speech sounds as follows:

[+vowel] monophthongs (reduced, lax, tense), diphthongs,

(may also include syllabic liquids, syllabic nasals, nasalized or aspirated vowels)

[+glide] semivowels (y, w), liquids (r, l), aspirant (h)

[+cons], [+son], [-cont] nasal consonants m, n, ng

[+cons], [-son], [+cont] fricative consonants (unmarked for stridency) v, f

[+cons], [-son], [+cont], [-strid] fricative consonants (non-strident, no obstacle) dh, th

[+cons], [-son], [+cont], [+strid] fricative consonants (strident, obstacle) z, zh, s, sh

[+cons], [-son], [-cont] stop consonants b, d, g, p, t, k

[+cons], [-son], [+/-cont] affricate consonants jh, ch

Including silence, there are 9 classes to consider. A shorthand notation for these manner classes is given below:

# silence before and after an utterance

v vowel monophthongs (reduced, lax, tense), diphthongs

(may also include syllabic liquids, syllabic nasals, nasalized or aspirated vowels)

g glide semivowels (y, w), liquids (r, l), aspirant (h)

n nasal consonant m, n, ng

fu fricative consonant unmarked v, f

fn fricative consonant non-strident dh, th

fs fricative consonant strident z, zh, s, sh

s stop consonant b, d, g, p, t, k

a affricate consonant jh, ch

1. From these manner classes, several types of landmarks (LMs) are generated, according to underlying phoneme. The landmarks correspond to primary articulator descriptions, or oral cavity articulations (oral events). The various types are listed below.

V vowel LM energy peak

G glide LM energy dip

Nc nasal closure LM abrupt onset of nasal

Nr nasal release LM abrupt offset of nasal

Fc fricative closure LM abrupt onset of frication

Fr fricative release LM abrupt offset of frication

Sc stop closure LM cessation of vocal tract activity due to oral closure

Sr stop release LM abrupt release of burst energy due to oral release

Tn stridency onset LM onset of stridency

Tf stridency offset LM offset of stridency

(other possible types not dealt with here include abrupt onset and offset of liquids)

Lc liquid closure LM abrupt onset of liquid

Lr liquid release LM abrupt offset of liquid

1. Landmarks generated from underlying phoneme-pair contexts

Since landmarks occur at saliency points (maxima/minima for vowels and glides, articulator-free feature changes for consonants), we may select manner class-pair contexts as a general unit for schematically generating landmarks from phoneme sequences. To allow concatenative generation from these units, we will define a manner class-pair unit as the interval from the midpoint of the preceding manner class (including the maximum/minimum point) to just before the midpoint (or maximum/minimum) of the following manner class. A quasi-mathematical notation would be in the form of [ preceding segment midpoint, following segment midpoint ). With this convention, the 9 manner classes will lead to 81 manner-class pair contexts: 9 underlying phoneme manner classes x 9 underlying phoneme manner classes = 81 contexts, i.e. {#,v,g,n,fu,fn,fs,s,a} x {#,v,g,n,fu,fn,fs,s,a} = { # - #, # - v, # - g, … }. These are listed below, with the generated landmarks for each context. Commas denote sequential events, while slashes denote concurrent events.

<NB> The +n and –n notations indicate nasal onset and offset, which is described in Section B. Non-oral events.

1. # - # no landmarks or events generated
2. # - v
3. # - g
4. # - n +n/Nc
5. # - fu Fc
6. # - fn Fc
7. # - fs Fc/Tn
8. # - s Sc
9. # - a Sc, Sr, Fc
10. v - # V
11. v - v V
12. v - g V
13. v - n V, +n/Nc
14. v - fu V, Fc
15. v – fn V, Fc
16. v – fs V, Fc/Tn
17. v - s V, Sc
18. v - a
19. g - # G
20. g - v G
21. g - g G
22. g - n G, +n/Nc
23. g – fu G, Fc
24. g – fn G, Fc
25. g – fs G, Fc/Tn
26. g - s G, Sc
27. g - a G, Sc, Sr, Fc
28. n - # Nr/-n
29. n - v Nr/-n
30. n - g Nr/-n
31. n - n but (Nr, Nc) may be possible for different places
32. n – fu Nr/-n, Fc
33. n – fn Nr/-n, Fc
34. n – fs Nr/-n, Fc/Tn
35. n - s Nr/-n, Sc
36. n - a Nr/-n, Sc, Sr, Fc
37. fu - # Fr
38. fu – v Fr
39. fu – g Fr
40. fu – n Fr,+n/Nc
41. fu – fu but (Fr/(+/-g)/Fc) may be possible for different places
42. fu - fn Fr/Fc a change from unmarked for stridency to [-strid]
43. fu – fs Fr/Fc/Tn
44. fu – s Fr,+n/Nc
45. fu – a Fr/Sc, Sr, Fc
46. fn - # Fr
47. fn – v Fr
48. fn – g Fr
49. fn – n Fr,+n/Nc
50. fn – fu Fr/Fc a change from [-strid] to unmarked for stridency
51. fn – fn but (Fr/(+/-g)/Fc) may be possible for different places
52. fn – fs Fr/Fc/Tn
53. fn – s Fr,+n/Nc
54. fn – a Fr/Sc, Sr, Fc
55. fs - # Fr/Tf
56. fs – v Fr/Tf
57. fs – g Fr/Tf
58. fs – n Fr/Tf,+n/Nc
59. fs – fu Fr/Tf/Fc a change from [+strid] to unmarked for stridency
60. fs – fn Fr/Tf/Fc a change from [+strid] to [-strid]
61. fs – fs but (Fr/(+/g)/Fc) may be possible for differenct places
62. fs – s Fr/Tf,,+n/Nc
63. fs – a Fr/Tf/Sc, Sr, Fc
64. s - # Sr
65. s - v Sr Sr and (+g) time difference = VOT
66. s - g Sr
67. s - n Sr, +n/Nc
68. s – fu Sr/Fc
69. s – fn Sr/Fc
70. s – fs Sr/Fc/Tn
71. s - s Sr/Sc hot dog
72. s - a Sr/Sc, Sr, Fc cold juice
73. a - # Fr
74. a - v Fr
75. a - g Fr
76. a - n Fr, +n/Nc
77. a – fu Fr/Fc
78. a – fn Fr/Fc
79. a – fs Fr/Fc/Tn
80. a - s Fr/Sc church bell
81. a - a Fr/Sc, Sr, Fc

<NB> An alternative scheme describing standard generation of landmarks/events from articulatory states is also possible. In this case, the underlying states are {#, v, g, n, f, sc, sr}, for silence, vowel, glide, nasal, fricative, stop closure, stop release. An affricate will be considered a sequence of states comprising {sc, sr, f}. A similar list with generated landmarks/events is found, but a more compact representation is possible using a state transition diagram to show possible landmarks/events that can be generated.

<NB> Also, fricatives can be grouped into 2 categories: strident vs. unmarked/non-strident. This will cut down on the number of contexts to be considered, and the final number of contexts will be 8 x 8 = 64 contexts. The reduced set of contexts can be represented with f (= {fu, fn}) vs. fs, i.e., unmarked and non-strident fricatives, vs. strident fricatives.

1. **Non-oral events**
2. In addition to landmarks, which represent saliency points for articulator-free features, other events corresponding to secondary articulators also occur. These are also called non-oral events. Some non-oral events may be predictable from the underlying manner classes, but others are not predictable from the lexical specification, and may be related to other speech information, such as prosody, or with voice quality effects. The non-oral events that are (mostly) predictable from underlying manner classes are related to glottal voicing and nasality. These are indicated as follows:

+g glottal (voicing) onset

-g glottal (voicing) offset

+n velopharyngeal port opening (nasal) onset

-n velopharyngeal port opening (nasal) offset

Other non-oral events that are usually not predictable from underlying manner classes include the following:

(+q) glottalization onset (or irregular pitch period: ipp)

(-q) glottalization offset

(+h) aspiration onset

(-h) aspiration offset

(+p) pharyngealization onset

(-p) pharyngealization offset

1. It is straightforward to predict the glottal voicing and velopharyngeal events from the manner class of underlying phonemes. For +g and –g events, the classes are divided into the voiced and unvoiced classes.

voiced all vowels, semivowels, liquids, nasals, and voiced obstruents

unvoiced all other unvoiced phonemes and silence = {#, h, f, th, s, sh, p, t, k}

Then, the glottal voicing class-pairs (total of 4 contexts) and the corresponding events that are generated are:

1. voiced – voiced
2. voiced – unvoiced -g
3. unvoiced – voiced +g
4. unvoiced – unvoiced

For the +n and –n events, we may similarly divide all classes into the nasal and non-nasal classes and generate a 4-context description identical to that for glottal voicing.

nasal nasal consonants, nasalized vowels

non-nasal all other phonemes and silence

1. nasal - nasal
2. nasal – non-nasal -n
3. non-nasal - nasal +n
4. non-nasal – non-nasal

However, since nasality is not distinctive for vowels in English, we may conflate the nasal events with landmark generation of Nc and Nr, as +n/Nc and Nr/-n. This has been done in Section 3 above.

<NB>If syllabic nasals, such as em, en, eng, are introduced into the phoneme list, then the conflation scheme will not be as effective, and a separate +n/-n generation scheme will be simpler.