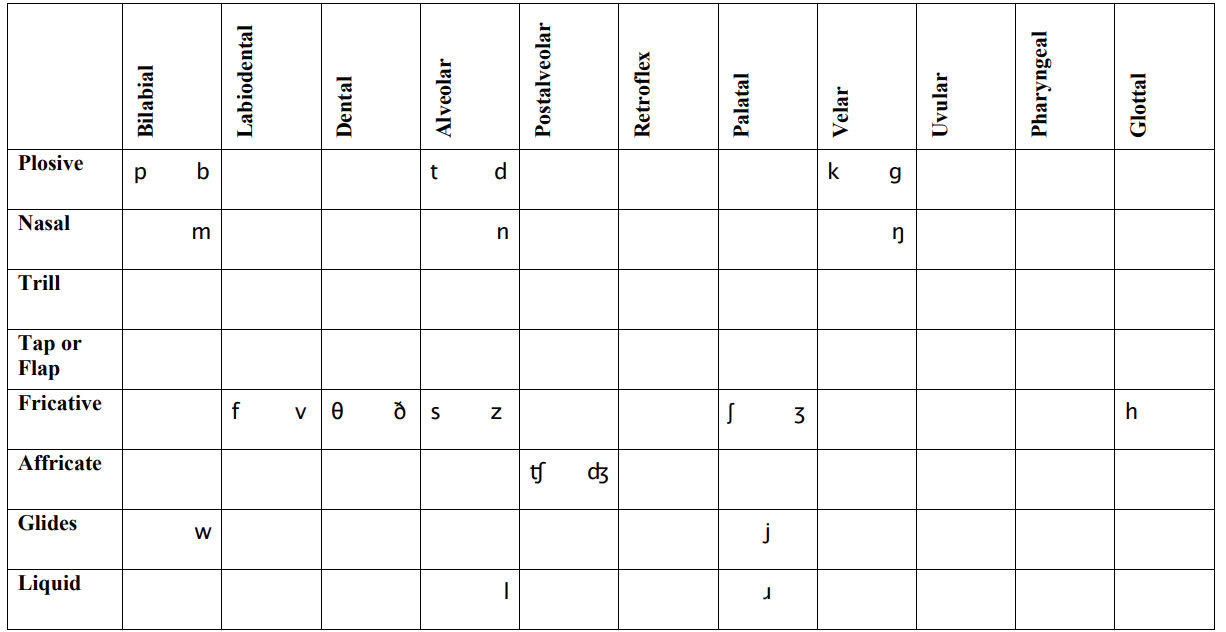
Consider this phoneme inventory chart for English from the American Speech-Language-Hearing Association:



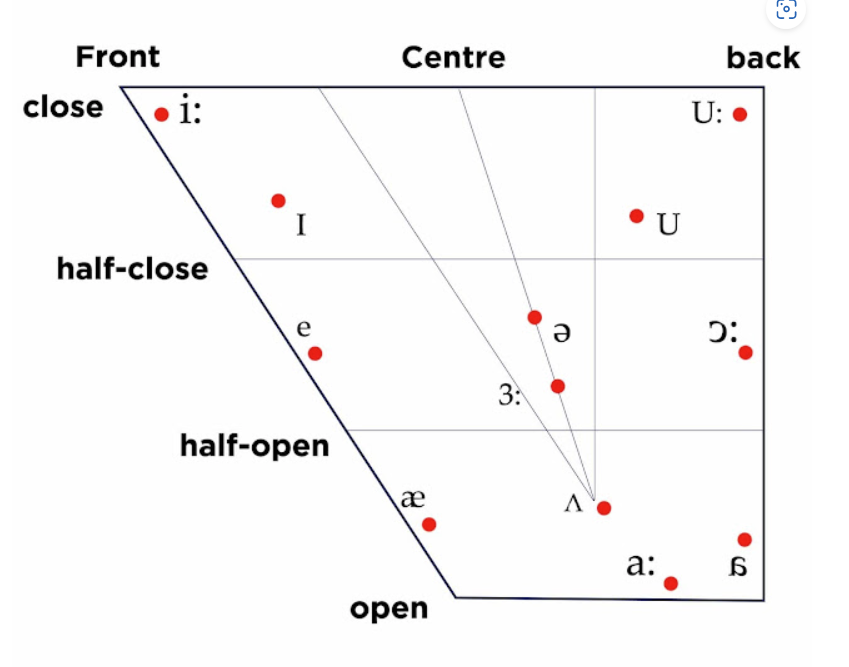
We simplify it with respect to NUPhone as this:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Bilabial | Labiodental | Dental | Alveolar | Palatal | Velar | Glotal |  |  |
| Plosive | p b |  |  | t d |  | k d |  | 2 | 2 |
| Nasal | m |  |  | n |  | ŋ |  | 1 | 1 |
| Fricative |  | f v | Θ ð | s z |  | ʃ ʒ | h | 0 | 0 |
| Glides | w |  |  |  |  |  |  | -1 | 5 |
| Liquid |  |  |  | l | ɹ |  |  | -2 | 6 |
|  | -3 | -2 | -1 | 0 | 1 | 2 | 3 |  |  |
|  | 7 | 6 | 5 | 0 | 1 | 2 | 3 |  |  |

Note that we delete all blank rows and columns in the inventory. The affricate is also deleted, because in English NUPhone, we represent t͡ʃ as t|ʃ and d͡ʒ as d|ʒ. Using uncertainty here aids is fuzziness for comparison purposes.

Also note that while two phonemes can exist in a single cell, we treat it as a single numeric. This is because we track voicing with its own distinct bit.

We do the same thing for the vowel chart, assigning numerics to vowell positions. The middle of the chart is (0,0). We refer to Centerness and Middleness.



|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Front | Center | Back |  |  |  |  |  |  | |
| close | i |  | u | -2 | 6 | (0x4 is the negative bit) | | | |
| ½ close | ɪ |  | ʊ | -1 | 5 | (0x4 is the negative bit) | | | |
| middle | e | ə |  | 0 | 0 |  |  |  |  |
| ½ open | ɛ | ɜ | ɔ | 1 | 1 |  |  |  |  |
| open | æ | ʌ ɑ | ɐ | 2 | 2 |  |  |  |  |
|  | 1 | 0 | -1 |  |  |  |  |  |  | |
|  | 1 | 0 | 3 | (0x2 is the negative bit) | | | |  |  | |

**Bit mappings**

|  |  |  |
| --- | --- | --- |
| Description | Mask | Values |
| Phoneme-type: 2 bits | 0xC\_ | 0x0\_: vowel  0x8\_: consonant-unvoiced  0xC\_: consonant-voiced  0x4\_: undefined |
| Vowel Middle-ness: 2 bits | 0x3\_ | see middle-ness in vowel table |
| Vowel Center-ness: last 3 bits | 0x07 | see center-ness in vowel table |
| Consonant: Place or articulation: 3 bits | 0x38 | see consonant table |
| Consonant: Manor or articulation: last 3 bits | 0x07 | see consonant table |

This means that we can store all phoneme embeddings for English NUPhone in an 8-bit integer (1 byte)

Each three-bit signed integer has a distance measurement. Distances of the two dimensions are additive. If we use percentage-like scores, we can establish these consonant scoring heuristics.

Not a consonant: 0%

Distance == 0 🡪 Similarity: 100%

Distance == 1 🡪 Similarity: 98% [-2%]

Distance == 2 🡪 Similarity: 94% [-4%]

Distance == 3 🡪 Similarity: 86% [-8%]

Distance == 4 🡪 Similarity: 70% [-16%]

Distance >= 5 🡪 Similarity: 38% [-32%]

Consider these two phonemes: k & ð

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Bilabial | Labiodental | Dental | Alveolar | Palatal | Velar | Glotal |
| Plosive | p b |  |  | t d |  | k d |  |
| Nasal | m |  |  | n |  | ŋ |  |
| Fricative |  | f v | Θ ð | s z |  | ʃ ʒ | h |
| Glides | w |  |  |  |  |  |  |
| Liquid |  |  |  | l | ɹ |  |  |

Manner delta = 2

Place delta = 3

Distance (2 + 3) = 5 🡪 Similarity: 38%

Voicing penalty = -1% 🡪 Similarity: 37%

**And a similar heuristic for vowels:**

Not a vowel: 0%

Distance == 0 🡪 Similarity: 100%

Distance == 1 🡪 Similarity: 99%

Distance == 2 🡪 Similarity: 97%

Distance == 3 🡪 Similarity: 93%

Distance >= 4 🡪 Similarity: 85%

Consider these two phonemes: ʊ & ə

|  |  |  |  |
| --- | --- | --- | --- |
|  | Front | Center | Back |
| close | i |  | u |
| ½ close | ɪ |  | ʊ |
| middle | e | ə |  |
| ½ open | ɛ | ɜ | ɔ |
| open | æ | ʌ ɑ | ɐ |

Middle-ness delta = 1

Central-ness delta = 1

Distance (1 + 1) = 2 🡪 Similarity: 97%