

Logo

- To **save** changes: Press the [Save Changes] button.
- To make **more changes**: Go back in your browser.
- To **cancel**: Go back twice.

Release edit lock ([help](#)) Minor changes, don't notify ([help](#)) Save Changes

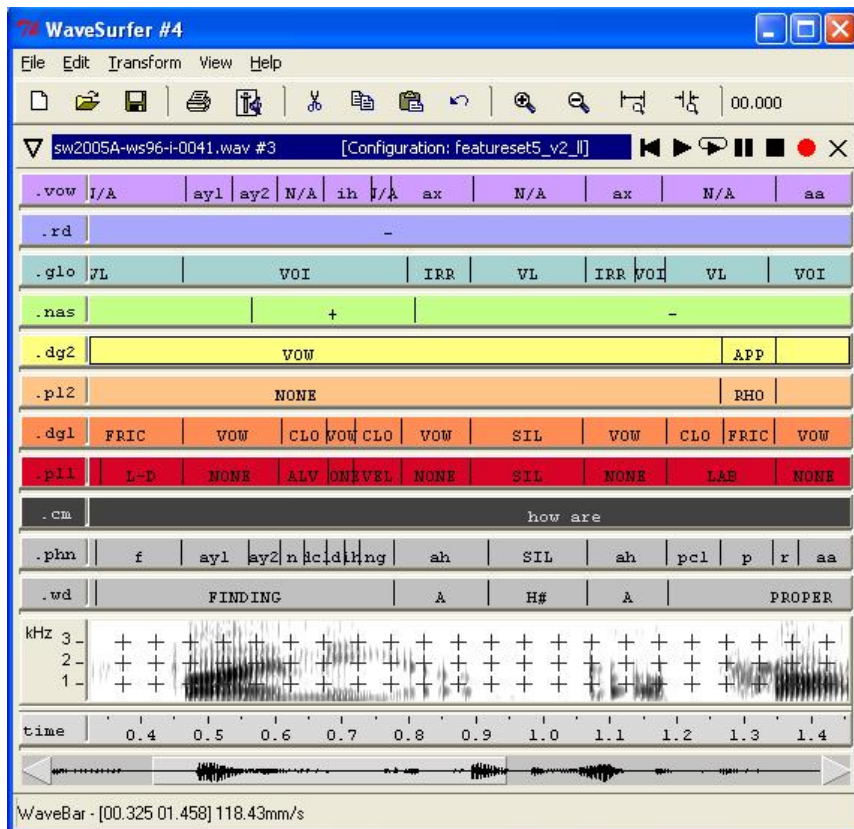
WS06.FeatureTranscription (preview)

Note:This is a preview. **Do not forget** to save your changes.

Feature transcriptions

This page is used to maintain information and materials for developing a set of "ground-truth" transcriptions at the feature level.

- [Background](#)
- [Plan](#)
- [Status](#)
- [Materials](#)
 - [Feature set](#)
 - [Phone-to-feature-set-4 mapping](#)
 - [Modified feature set \(Feature Set 5\)](#)
 - [Phone-to-feature-set-5 mapping](#)
 - [Transcription tools](#)
- [Analysis](#)
- [Discussion](#)



Background

There are several motivations for generating a set of feature-level transcriptions:

- To serve as reference for measuring feature classifier accuracy
- To train pronunciation models separately from acoustic models
- To study asynchrony and reduction effects

In the past, classifier accuracy has been measured by comparison against a reference phonetic transcription, assuming some mapping from phones to feature values. However, especially for conversational speech, we cannot assume that such a mapping would give us accurate reference feature values; there is too much coarticulation and reduction.

We are not aware of any data set that has been labeled at the feature level. There are, of course, some corpora of measured articulation, such as MOCHA or the Wisconsin X-ray microbeam database. These could also be used, but the mapping from measurements to feature values is non-trivial, and often the measurements do not include some important information, such as nasality. This motivates us to generate this new data set.

Plan

1. Manually transcribe a small set of utterances, say 50-100. These will serve as testing material for acoustic classifiers.
2. If the classifiers are accurate enough, use them, in combination with word transcriptions, to force-align a larger set of utterances. This larger set would serve as our "ground truth" set.

This is a tentative timeline:

1. By some time T
 - Finalize the set of transcribers (2-4)
 - Finalize the feature set and phone-to-target-features mapping
 - Finalize the transcription interface
 - Develop detailed transcription guidelines
2. By T + 2 weeks
 - Transcriber training: go over feature set in detail, practice transcribing together
3. By T + 4 weeks
 - Preliminary experiments: Measure transcriber agreement and speed for different variants of the interface (e.g. with/without target transcriptions provided, with optional collapsing of tiers into phones in careful speech regions)
4. By T + 12 weeks
 - Transcribe 50-100 utterances
 - Weekly meetings to discuss transcriber questions, ambiguities, disagreements
5. > T + 12 weeks
 - Use transcriptions! Analyze, use for workshop experiments, etc.

Status

- Transcriptions done! We have 78 SVitchboard and 9 STP transcriptions done. Of the 9 STP transcriptions, 5 have been done in an all-feature format and 4 in a phone-feature hybrid format (see below).
- We are currently finishing up remaining cleanup, agreement measures, etc.
- Thanks to our transcribers, Xuemin Chi (a graduate student) and Lisa Lavoie (a phonetician). Since both have a speech background, we did not do significant training.
- We also have benefited from help/input from Nancy Chen, Edward Flemming, Jim Glass, Daryush Mehta, Stefanie Shattuck-Hufnagel, and Janet Slifka (thanks!).

Excerpts of meeting notes/status reports:

- Mar. 8

- We went over one of Xuemin's transcriptions and discussed issues related to rhotics. Any comments on these are welcome! The issue is what the canonical feature values should be for /r/ vs. /er/ (bird) vs. /axr/ (retroflex schwa), and whether pre-vocalic /r/ should have different values from post-vocalic. As it stands, pre-/post-vocalic /r/s are identical, and /er/ and /axr/ are labeled as vowels with a RHO/APP constriction. This is a bit weird as all other vocalic regions are labeled NONE/VOW in both place/degree tiers.
- Mar. 15
 - Based on conversations with Lisa, Xuemin, and Edward, we are trying an experiment. We will compare two transcription setups:
 1. All-feature: As we've done until now.
 2. Hybrid phone-feature: Use phone labels for segments that look/sound like a standard phone, and use the feature tiers for segments that don't. Typical cases where the feature tiers are needed: Fricated/approximant realizations of stops, retroflexion/lateralization during a stop burst. This might both save time and, to some extent, keep us from making arbitrary judgments about feature values.
 - Lisa and Karen met today and Lisa transcribed 7 utterances, alternating between phone-feature and all-feature setups and timing herself on each utterance.
 - Xuemin will transcribe the same 7 utts with the same alternation before our next group meeting.
- Mar. 22
 - We went over 4 out of the 7 utts that Xuemin and Lisa transcribed this week. We found some typos and disagreements in the transcriptions, but overall, they are quite consistent (qualitatively--no quantitative measures yet).
 - Dividing transcription time by total speech time in each utt (excluding initial/final silence), we get real-time factors of around 500 for Lisa and 1500 for Xuemin. The difference is expected based on Lisa's previous transcription experience, and on the fact that Lisa used some shortcuts that Xuemin didn't.
 - The phone-feature hybrid takes 25-35% less time than the all-feature setup.
 - Both Lisa and Xuemin commented that it's not clear which setup is easier: The phone labeling is faster, but they still need to examine the speech carefully, and they might be tempted to label segments as canonical when they are not.
 - For the next meeting, Xuemin and Lisa will do some additional utterances, still alternating between the two setups. This time we will use utterances from STP, which are longer and more varied in phonetic content.
 - Laterals were problematic in some of this week's utts, and we decided to tweak the feature set: LAT is now a place, with no distinction between light and dark [l]s, other than that dark [l]s are more likely to be approximants. Feature Set 5 and the [TranscriptionNotes](#) are updated to reflect this.
 - We discussed the labeling/non-labeling of transitional regions between steady-state segments. We decided not to label them as separate segments if they are obligatory motions from one state to the next; if they are more than the minimum necessary, label them (e.g. in "feel" --> [f iy ax l], the [ax] region should be labeled separately from the [iy] and [l]).
- Mar. 29
 - Xuemin & Lisa transcribed 9 utts from STP, alternating between phone-feat and all-feat transcriptions. These were much longer than last week's, so we only went over two of them.
 - Average real-time factors over the 9 utts:
 - XC: phone-feat 465, feat 1003
 - LL: phone-feat 260, feat 586
 - In comparing transcriptions, we found a number of typos. We decided that instead of looking at transcriptions as a group as soon as they are done, Xuemin & Lisa will first do a 2nd pass through each one, comparing with each other's 1st pass transcriptions and fixing any typos (**NOT** disagreements; those stay!) At the next meeting, we will go over the remaining 10 transcriptions from the last two weeks, after Xuemin & Lisa do a 2nd pass.
 - Lisa & Xuemin commented that the baseform phonetic alignment that is given may be distracting/confusing when doing the phone-feat transcriptions. For the next set of utterances to be transcribed, we will have some with baseform phone alignment given and some without. We will look at this next set as a group two meetings from now, after both have had a chance to do a

2nd pass on them.

- Apr. 10
 - Xuemin and Lisa did a 2nd pass through most of their STP transcriptions, each comparing against the other's transcriptions.
 - In our meeting, we looked at two of the STP transcriptions that had been checked. We found that, qualitatively at least, there are fewer consistent differences between the two transcribers.
 - We still found a number of typos, and decided that in order to do a better job of catching them, we need a better way of doing the 2nd pass. We decided that Karen will make a wavesurfer config file that allows Xuemin and Lisa to look at both of their transcriptions on top of each other.
- Apr. 19
 - Karen has made a new wavesurfer config file for comparison of pairs of transcriptions.
 - Karen met separately with Lisa and Xuemin to go over the new config and additional comments/questions. The new config seems to be extremely helpful for doing comparisons and error-checking. Woo-hoo! [Here](#) is an example of a comparison of two transcriptions of the same utt; for each feature, the top tier in each pair is Xuemin; the bottom is Lisa.
 - For the next week, Xuemin and Lisa will transcribe the new set of utts, comparing (1) feat vs. phone-feat hybrid setups and (2) being given vs. not being given an initial phonetic alignment. Then they'll pass their transcriptions to each other for the 2nd pass, following which we'll reconvene for a meeting. So next meeting should be in ~2 weeks.
- Apr. 23 (first planning meeting for WS06)
 - A suggestion from the planning meeting: We may want to transcribe only those utterances in SVitchboard that are also in STP (all of which are in the "E" set of SVitchboard), for comparison.
- May 8
 - Xuemin and Lisa have both transcribed another 16 utts, using the 4 transcription variants (i) phone-feat hybrid, canonical phone alignment given; (ii) phone-feat hybrid, no initial phone alignment; (iii) all-feat transcription, initial phone alignment given; (iv) all-feat transcription, no initial phone alignment.
 - Next meeting Wed. May 10, to go over a few examples and decide which of the 4 variants above we'll use for the remainder of the transcriptions.
- May 10
 - We decided to go with phone-feature hybrid transcriptions, with initial .phn transcriptions provided.
 - Karen will also provide boundaries for the initial/final silence, plus one extra boundary in the middle of the utt, for all feature tiers.
 - Xuemin and Lisa took an oath to not be tempted to stick to canonical phones.
 - From next week till the end of June, we'll do about 15 utterances per week, "due" on Sunday of each week. We'll meet about every other week to go over examples/issues. We have 6 sets of 15 utts to get through, call them Set 1-6. Call last week's set Set 0. Here's the projected schedule:
 - Sun. 5/21: Set 0 2nd pass, Set 1 1st pass done
 - Wed. 5/24: meeting
 - Sun. 5/28: Set 1 2nd pass, Set 2 1st pass done
 - Sun. 6/4: Set 2 2nd pass, Set 3 1st pass done
 - Mon. 6/5: meeting
 - Sun. 6/11: Set 3 2nd pass, Set 4 1st pass done
 - Sun. 6/18: Set 4 2nd pass, Set 5 1st pass done
 - Sun. 6/25: Set 5 2nd pass, Set 6 1st pass done
 - Mon. 6/26: meeting
 - Sun. 7/2: Set 6 2nd pass done!
 - Mon. 7/3: last meeting!
 - Through Set 1, the utterances were simply picked by Karen in such a way as to try to maximize variability in the data that is transcribed. Starting from Set 2, however, the utterances are picked randomly from among those SVitchboard utterances in the 500-vocab set containing 5 words or more (excluding initial/final silence).

- May 29? (Lost track of the exact date)
 - We had some concerns about doing the 2nd pass in phone-feature format: It can be hard to see errors when one transcriber used a phone label and the other used features for a given segment; and in addition, the transcribers don't get to see the final feature values that will be used as their transcriptions. We concluded that instead, the 2nd pass will be done in an **all-feature** format, after an automatic post-processing to convert the 1st pass phone labels to features. So the new procedure is:
 - 1st pass: Phone-feature hybrid
 - Post-processing script convert phones to features. The post-processing script also produces a list of warnings about illegal/missing feature values that can help in detecting errors (but no effort has been made to detect **all** errors).
 - 2nd pass: Transcribers compare the post-processed, all-feature versions of their transcriptions, and each edits her own transcriptions directly in the feature tiers. The phone tier is no longer used in the 2nd pass.
 - Since this change is taking place after the 2nd pass for Set 1, Xuemin & Lisa will do a 3rd pass of just that set, starting from their 2nd pass transcriptions converted to all-feature format.
- June 7
 - Since there are still some errors after the 2nd pass of Set 2, we decided that Xuemin and Lisa will do a 3rd pass of this set.
- June 19
 - Xuemin & Lisa decided to do 3rd passes for **all** utterances to get rid of remaining errors. This 3rd pass is done by discussing any differences face-to-face and each transcriber altering her own transcriptions as appropriate. We will later check how much this affects their inter-transcriber agreement. At any rate, all 3 passes will be kept for posterity.
- June 20
 - Since we are now doing 3 passes, each utterance is taking a bit longer, so we will put off/get rid of the last set (Set 6) in favor of spending more time getting the rest of the transcriptions polished.
- June 28
 - We decided to not do Set 6, but to go back to the 9 STP utterances that Xuemin and Lisa had transcribed and do 2nd & 3rd passes of those. This will help us to compare this transcription effort to STP. Since the STP transcriptions alternated between hybrid phone-feature and all-feature transcriptions, only the phone-feature ones will be converted to all-feature. This means that this set won't use **exactly** the same procedure as the others, but it should still give us useful information.
- July 5
 - All final transcriptions done!

Materials

This section is used to maintain evolving materials for ongoing transcription work.

Feature set

The feature set we started out with (same as Feature Set 4 in [FeatureSets](#)) and some concerns we had about it:

| <u>Feature name</u> | <u>Values</u> | <u>Comments</u> |
|---------------------|--|--|
| place | LAB, LAB-DEN, DEN, ALV, POST-ALV, VEL, GLO, RHO, FRT, CEN, | GLOttal place is used for [hh]; POST-ALVeolar includes palato-alveolars (sh, ch, etc.) and palatals (y). |

| | | |
|-----------------------|-------------------------------------|--|
| | BK, SIL | |
| manner | VOW, GLI, LAT, FLAP, FRIC, CLO, SIL | GLI might be better called "approximant" and is intended to refer to any articulation in which there is a narrow closure (includes the usual glides, but also stops realized as approximants); LAT is for "l"; CLO refers to any <i>complete</i> closure (including nasal closures); FRIC refers to both fricatives and stop bursts. |
| nasality | +, - | + means "velum is open" and therefore is used for both nasal consonants and other nasalized sounds. |
| voicing | +, - | |
| lip-rounding | +, - | |
| vocalic tongue height | HI, MID, LO, NA | NA ("not applicable") includes all consonantal articulations and silence. |

Issues with the feature set:

- Should LAT be a place? What do we label an [l] realized as an approximant?
- How to handle aspiration? (Voiceless stop aspiration, [hh], aspirated vowels)
- Should there be separate places for dental and inter-dental? Reasoning for merging them into DEN above: We do not expect to see any inter-dental stops or dental fricatives/approximants.
- What to do about glottal stops/glottalization? (Currently we are ignoring them)
- Need more values for vowel front/back and high/low; currently, we cannot distinguish all vowels, even without reduction or coarticulation. Alternatively, have a separate tier for vowel phonetic labels? Doesn't seem like this would lose any information relative to separate front/back and high/low tiers.
- No way to represent multiple constrictions. Add additional place values such as labio-velar?

Phone-to-feature-set-4 mapping

The phone set (mostly stable but might have slight changes) will be based on the ARPAbet. Mappings between IPA symbols and ARPAbet can be found [here](#). A mapping from phones to feature values is below. This is a slightly different phone set from the basic ARPAbet. The main differences are:

- Diphthongs are broken up into two "phones", corresponding to the initial and final configurations.

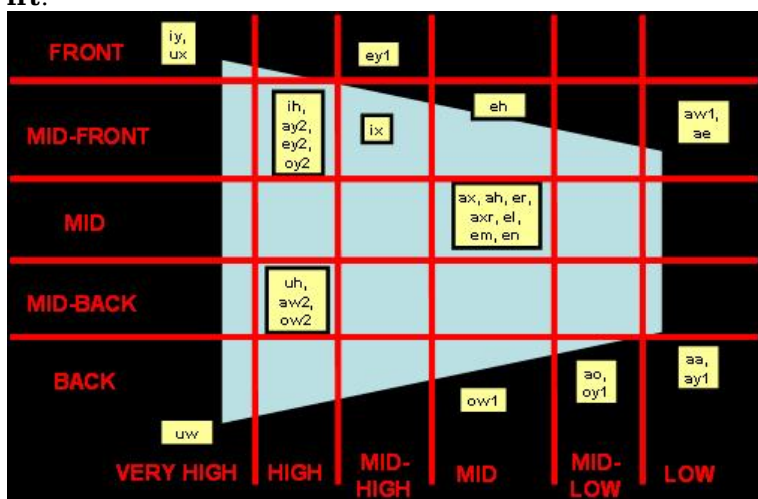
| phn | place | manner | nasal | voicing | lip-rd | voc-ht | phn | place | manner | nasal | voicing | lip-rd | voc-ht |
|-----|----------|--------|-------|---------|--------|--------|-----|----------|--------|-------|---------|--------|--------|
| aa | BK | VOW | - | + | - | LO | jh | POST-ALV | FRIC | - | + | - | NA |
| ae | FRT | VOW | - | + | - | LO | k | VEL | FRIC | - | - | - | NA |
| ah | CEN | VOW | - | + | - | MID | kl | VEL | CLO | - | - | - | NA |
| ao | BK | VOW | - | + | + | LO | l | ALV | CLO | - | + | - | NA |
| aw1 | FRT | VOW | - | + | - | LO | m | LAB | CLO | + | + | - | NA |
| aw2 | BK | VOW | - | + | + | HI | n | ALV | CLO | + | + | - | NA |
| ax | CEN | VOW | - | + | - | MID | ng | VEL | CLO | + | + | - | NA |
| axr | RHO | GLI | - | + | - | NA | ow1 | BK | VOW | - | + | + | HI |
| ay1 | BK | VOW | - | + | - | LO | ow2 | FRT | GLI | - | + | + | NA |
| ay2 | FRT | VOW | - | + | - | HI | oy1 | CEN | VOW | - | + | + | HI |
| b | LAB | FRIC | - | + | - | NA | oy2 | FRT | VOW | - | + | - | HI |
| bcl | LAB | CLO | - | + | - | NA | p | LAB | FRIC | - | - | - | NA |
| ch | POST-ALV | FRIC | - | - | - | NA | pcl | LAB | CLO | - | - | - | NA |
| d | ALV | FRIC | - | + | - | NA | q | GLO | CLO | - | - | - | NA |
| dcl | ALV | CLO | - | + | - | NA | r | RHO | GLI | - | + | - | NA |
| dh | DEN | FRIC | - | + | - | NA | s | ALV | FRIC | - | - | - | NA |
| dx | ALV | FLAP | - | + | - | NA | sh | POST-ALV | FRIC | - | - | - | NA |
| eh | FRT | VOW | - | + | - | MID | t | ALV | FRIC | - | - | - | NA |
| el | ALV | CLO | - | + | - | NA | tcl | ALV | CLO | - | - | - | NA |

| | | | | | | | | | | | | | |
|-----|---------|------|---|---|---|-----|-----|---------|------|---|---|---|----|
| em | LAB | CLO | + | + | - | NA | th | DEN | FRIC | - | - | - | NA |
| en | ALV | CLO | + | + | - | NA | uh | BK | VOW | - | + | + | HI |
| er | RHO | CLO | - | + | - | NA | uw | BK | VOW | - | + | + | HI |
| ey1 | FRT | VOW | - | + | - | MID | v | LAB-DEN | FRIC | - | + | - | NA |
| ey2 | FRT | VOW | - | + | - | HI | w | BK | GLI | - | + | + | NA |
| f | LAB-DEN | FRIC | - | - | - | NA | y | FRT | GLI | - | + | - | NA |
| g | VEL | FRIC | - | + | - | NA | z | ALV | FRIC | - | + | - | NA |
| gcl | VEL | CLO | - | + | - | NA | sil | SIL | SIL | - | - | - | NA |
| hh | GLO | FRIC | - | - | - | MID | | | | | | | |
| ih | FRT | VOW | - | + | - | HI | | | | | | | |
| iy | FRT | GLI | - | + | - | NA | | | | | | | |

Modified feature set (Feature Set 5)

Another proposed feature set, which has come out of discussions at feature transcription meetings. This is mainly intended to make the set more expressive, including the ability to distinguish a larger number of both canonical and non-canonical articulations.

A main difference from Feature Set 4 is that height/front-back have been replaced with a single vowel tier; height and front-back fail to distinguish certain vowel sets and were found to be hard to use in practice when transcribing. However, in case we want to use them for recognition experiments, height & front-back features are included below. Here is a diagram showing the relationship between **vow**, **ht** and **frt**:



| <u>Feature name</u> | <u>Values</u> | <u>Comments</u> |
|---------------------|--|--|
| p11 (place 1) | LAB, L-D, DEN, ALV, P-A, VEL, GLO, RHO, LAT, NONE, SIL | Place of forward constriction. |
| dg1 (degree 1) | VOW, APP, FLAP, FRIC, CLO, SIL | Degree of forward constriction. This is not exactly a degree of constriction feature, though; e.g., the same physical degree of constriction could result in a fricative or not, depending on the pressure behind it. We label a constriction as a fricative only if there is turbulence noise. FLAP is also not really a degree of constriction; it's really a closure which is short in duration. |
| p12 (place 2) | L-D, DEN, ALV, P-A, VEL, GLO, RHO, LAT, NONE, SIL | Place of rear constriction. One value fewer than place I, because can't have two labial constrictions. |

| | | |
|------------------------|--|--|
| dg2 (degree 2) | VOW, APP, FLAP, FRIC, CLO, SIL | Degree of rear constriction |
| nas (nasality) | +, - | + means "velum is open" and therefore is used for both nasal consonants and other nasalized sounds. |
| glo (glottal state) | stop (STOP), irregular pitch periods (IRR), regular pitch periods (VOI), voiceless (VL), aspiration (ASP), aspiration + voicing (A+VO) | Replaces voicing feature to deal with more states. "Voiceless" refers to both silence and non-silence voiceless. "Aspiration" refers to voiceless with aspiration (e.g. aspirated part of voiceless stop burst). "Aspiration + voicing" is used for voiced [h] and aspirated vowels/liquids/glides. When we label something as "aspirated", we are including aspiration noise that may originate elsewhere other than the glottis (so it is not really a "glottal state", but we are lumping it into this feature anyway). |
| rd (lip rounding) | +, - | |
| vow (vowel) | aa, ae, ah, ao, aw1, aw2, ax, axr, ay1, ay2, eh, el, em, en, er, ey1, ey2, ih, ix, iy, ow1, ow2, oy1, oy2, uh, uw, ux, N/A | Replaces front-back and high-low features. Doesn't seem like there's any information loss in doing this. |
| ht (vowel height) | LOW, MID-L, MID, MID-H, HIGH, V-HI (very high), N/A | |
| frt (vowel front-back) | BK, MID-B, MID, MID-F, FRT, N/A | |

Phone-to-feature-set-5 mapping

Mapping from phones to their **canonical** feature values. A few notes:

- '*' indicates that a feature is unspecified; e.g. the feature **vow** for [hh] and [q] or **rd** for rhotics and palato-alveolar fricatives.
- One phone, [q], can take on either of two values for **glo**, STOP or IRR.
- All of the phones have pl2 = NONE, dg2 = VOW canonically (though [w] is arguable). However, these features are used in the transcription of non-canonical regions.
- The **notes** column records other issues that have come up in discussions about specific phones.

| phn | pl1 | dg1 | pl2 | dg2 | nasrd | glo | vow | ht | frt | examples, notes |
|-----|------|------|------|-----|-------|-----|-----|-------|-------|---|
| aa | NONE | VOW | NONE | VOW | - | VOI | aa | LOW | BK | as in "bob" |
| ae | NONE | VOW | NONE | VOW | - | VOI | ae | LOW | MID-F | as in "bat" |
| ah | NONE | VOW | NONE | VOW | - | VOI | ah | MID | MID | as in "but" |
| ao | NONE | VOW | NONE | VOW | - | VOI | ao | MID-L | BK | as in "bought" |
| aw1 | NONE | VOW | NONE | VOW | - | VOI | aw1 | LOW | MID-F | 1st articulatory configuration of the diphthong in "bout" |
| aw2 | NONE | VOW | NONE | VOW | - | VOI | aw2 | HIGH | MID-B | 2nd articulatory configuration of the diphthong in "bout" |
| ax | NONE | VOW | NONE | VOW | - | VOI | ax | MID | MID | 1st vowel in "about" |
| axr | RHO | APP | NONE | VOW | - | VOI | axr | MID | MID | as in end of "butter"; note: features same as [r], [er]? |
| ay1 | NONE | VOW | NONE | VOW | - | VOI | ay1 | LOW | BK | 1st articulatory configuration of the diphthong in "bite" |
| ay2 | NONE | VOW | NONE | VOW | - | VOI | ay2 | HIGH | MID-F | 2nd articulatory configuration of the diphthong in "bite" |
| b | LAB | FRIC | NONE | VOW | - | VOI | N/A | N/A | N/A | burst only |
| bcl | LAB | CLO | NONE | VOW | - | VOI | N/A | N/A | N/A | closure only |

| | | | | | | | | | | | |
|-----|------|------|------|-----|---|---|-----|-----|-------|-------|--|
| ch | P-A | FRIC | NONE | VOW | - | * | VL | N/A | N/A | N/A | frication part only of affricate in "chat" (closure part is [dcl]); same features as [sh]; note: rd? |
| d | ALV | FRIC | NONE | VOW | - | - | VOI | N/A | N/A | N/A | burst only ; note: same as [z]? |
| dcl | ALV | CLO | NONE | VOW | - | - | VOI | N/A | N/A | N/A | closure only |
| dh | DEN | FRIC | NONE | VOW | - | - | VOI | N/A | N/A | N/A | as in "them" |
| dx | ALV | FLAP | NONE | VOW | - | - | VOI | N/A | N/A | N/A | flap as in "butter" |
| eh | NONE | VOW | NONE | VOW | - | - | VOI | eh | MID | MID-F | as in "bet" |
| el | LAT | CLO | NONE | VOW | - | - | VOI | el | MID | MID | syllabic /l/ as in "bottle"; note: vow = N/A? |
| em | LAB | CLO | NONE | VOW | + | - | VOI | em | MID | MID | syllabic /m/ as in some productions of "bottom"; note: vow = N/A? |
| en | ALV | CLO | NONE | VOW | + | - | VOI | en | MID | MID | syllabic /n/ as in "button"; note: vow = N/A? |
| er | RHO | APP | NONE | VOW | - | * | VOI | er | MID | MID | as in "bird"; note: rd? features same as [r], [axr]? |
| ey1 | NONE | VOW | NONE | VOW | - | - | VOI | ey1 | MID-H | FRT | 1st articulatory configuration of the diphthong in "bait" |
| ey2 | NONE | VOW | NONE | VOW | - | - | VOI | ey2 | HIGH | MID-F | 2nd articulatory configuration of the diphthong in "bait" |
| f | L-D | FRIC | NONE | VOW | - | - | VL | N/A | N/A | N/A | |
| g | VEL | FRIC | NONE | VOW | - | - | VOI | N/A | N/A | N/A | burst only |
| gcl | VEL | CLO | NONE | VOW | - | - | VOI | N/A | N/A | N/A | closure only |
| hh | NONE | VOW | NONE | VOW | - | * | ASP | * | * | * | voiceless aspirant as in "he" |
| ih | NONE | VOW | NONE | VOW | - | - | VOI | ih | HIGH | MID-F | as in "bit" |
| ix | NONE | VOW | NONE | VOW | - | - | VOI | ix | MID-H | MID-F | front schwa, as in some productions of 2nd vowel in "roses" |
| iy | NONE | VOW | NONE | VOW | - | - | VOI | iy | V-HI | FRT | as in "beet" |
| jh | P-A | FRIC | NONE | VOW | - | * | VOI | N/A | N/A | N/A | frication part only of palato-alveolar affricate in "jot" (closure part is [dcl]); same features as [zh]; note: rd? |
| k | VEL | FRIC | NONE | VOW | - | - | VL | N/A | N/A | N/A | burst only ; note: for the canonical mapping, we are calling the entire burst fricated, i.e. ignoring the aspiration portion in stressed environments |
| kcl | VEL | CLO | NONE | VOW | - | - | VL | N/A | N/A | N/A | closure only |
| l | LAT | CLO | NONE | VOW | - | - | VOI | N/A | N/A | N/A | |
| m | LAB | CLO | NONE | VOW | + | - | VOI | N/A | N/A | N/A | |
| n | ALV | CLO | NONE | VOW | + | - | VOI | N/A | N/A | N/A | |
| nx | ALV | FLAP | NONE | VOW | + | - | VOI | N/A | N/A | N/A | nasal flap, as in some productions of "winter" or "winner" |
| ng | VEL | CLO | NONE | VOW | + | - | VOI | N/A | N/A | N/A | as in "bang" |
| ow1 | NONE | VOW | NONE | VOW | - | + | VOI | ow1 | MID | BK | 1st articulatory configuration of the diphthong in "boat" |
| ow2 | NONE | VOW | NONE | VOW | - | + | VOI | ow2 | HIGH | MID-B | 2nd articulatory configuration of the diphthong in "boat" |
| oy1 | NONE | VOW | NONE | VOW | - | + | VOI | oy1 | MID-L | BK | 1st articulatory configuration of the diphthong in "boy" |
| oy2 | NONE | VOW | NONE | VOW | - | - | VOI | oy2 | HIGH | MID-F | 2nd articulatory configuration of the diphthong in "boy" |
| p | LAB | FRIC | NONE | VOW | - | - | VL | N/A | N/A | N/A | burst only ; see note for [k] |

| | | | | | | | | | | | |
|-----|------|------|------|-----|---|---|-----------|-----|------|-------|--|
| pcl | LAB | CLO | NONE | VOW | - | - | VL | N/A | N/A | N/A | closure only |
| q | GLO | CLO | NONE | VOW | - | - | STOP/IRR* | * | * | * | glottal stop; note: voi? also, unspecified vow? [Used to be ST/IRR; changed it here and above. -AB, 7/14/06] |
| r | RHO | APP | NONE | VOW | - | * | VOI | N/A | N/A | N/A | rd? |
| s | ALV | FRIC | NONE | VOW | - | - | VL | N/A | N/A | N/A | |
| sh | P-A | FRIC | NONE | VOW | - | * | VL | N/A | N/A | N/A | rd? |
| t | ALV | FRIC | NONE | VOW | - | - | VL | N/A | N/A | N/A | burst only ; see note for [k]; also: same features as [s]? |
| tcl | ALV | CLO | NONE | VOW | - | - | VL | N/A | N/A | N/A | closure only |
| th | DEN | FRIC | NONE | VOW | - | - | VL | N/A | N/A | N/A | as in "thin" |
| uh | NONE | VOW | NONE | VOW | - | + | VOI | uh | HIGH | MID-B | as in "book" |
| uw | NONE | VOW | NONE | VOW | - | + | VOI | uw | V-HI | BK | as in "boot" |
| ux | NONE | VOW | NONE | VOW | - | + | VOI | ux | V-HI | FRT | fronted /uw/ as in some productions of "too"; same as German ü or vowel in French "sur" |
| v | L-D | FRIC | NONE | VOW | - | - | VOI | N/A | N/A | N/A | |
| w | LAB | APP | NONE | VOW | - | + | VOI | N/A | N/A | N/A | pl2 = VEL, dg2 = APP? |
| y | P-A | APP | NONE | VOW | - | - | VOI | N/A | N/A | N/A | as in "yet" |
| z | ALV | FRIC | NONE | VOW | - | - | VOI | N/A | N/A | N/A | |
| zh | P-A | FRIC | NONE | VOW | - | * | VOI | N/A | N/A | N/A | as in "measure"; note: rd? |
| sil | SIL | SIL | SIL | SIL | - | - | VL | N/A | N/A | N/A | Used to have pl2 = NONE, dg2 = VOW; why, I have no idea. So I changed it to SIL for all 4 pl/dg features. -KL, 7/11/06 |

Transcription tools

- Wavesurfer can be downloaded from <http://www.speech.kth.se/wavesurfer/download.html>
- Wavesurfer config files: [feature set 4](#), [feature set 5](#), [feature set 5 v2](#), [feature set 5 v3](#), [feature set 5 v3 comparison version](#)
- The [IPA chart](#) (included mainly for the vowel chart, for quick reminders about front/back and high/low)
- [TranscriptionNotes](#) to be used as guidelines when transcribing
- [Screen dumps](#) of transcriptions using different feature sets/setups.

Analysis

- Ongoing work on analysis of these transcriptions is at [FeatureTranscriptionAnalysis](#).


Discussion

Enter any comments, questions, or discussion regarding transcriptions in the comment box below. New comments will be appended below existing ones and will be signed with your user name.

(Edit - Preview)

Add comment

-- [KarenLivescu](#) - 25 Jan 2006

| Attachment | Action | Size | Date | Who | Comment |
|--|------------------------|----------------------|----------------------|------------------------------|-------------------------|
|  main.pdf | manage | 177.9 K | 23 Jul 2006 - 03:24 | KarenLivescu | |

Release edit lock ([help](#)) Minor changes, don't notify ([help](#))

Save Changes