

UPSID and PHONEME

Version 1.1

Principal Investigator: Ian Maddieson

Programmer: Kristin Precoda

January 1992

Table of Contents

I. General overview of UPSID	1
II. Instructions for installation	3
III. Database and output files	3
IV. How UPSID menus and input screens work	5
V. What UPSID actually does	6
VI. Assignment of language numbers	14
VII. Meaning of character codes	15
VIII. UPSID features	16
IX. Maximum limits in UPSID	19
APPENDICES	
1. Classified List of Languages Included in UPSID	19
2. List of Character Codes and IPA equivalents, with segment frequency	24

I. General overview of UPSID

UPSID is an acronym for *UCLA Phonological Segment Inventory Database*. An earlier version was described in the book *Patterns of Sounds* (Maddieson 1984). The current version is corrected and expanded and is provided in the form of a custom-written program to facilitate its use. UPSID is used to describe both the database and the program (for IBM PC's and compatibles). A simplified companion program is called PHONEME. The database is designed to provide a tool for studying the patterns of distribution of phonologically contrastive segments in the world's languages. It can also be used as a quick reference to find occurrences of some particular segment or to look up the segment inventory of a particular language. However, it is designed to provide data on a sample of languages based on their genetic relationships, not an exhaustive listing of all languages or all known segment types. The primary purpose is to provide a sample from which statistically valid statements concerning frequency and co-occurrence can be drawn.

For each language included in the database there is a list of the segments that seem to serve a lexically contrastive function. The segments are listed as a set of mnemonic "character codes" and each code is described in terms of a set of features that are roughly equivalent to the phonetic properties of standard international phonetic theory. The properties assigned to a segment are determined from its most characteristic phonetic form. With rare exceptions all properties are positively specified; for example, all segments have a property [voiceless] or a property [voiced]

(in other words, voiceless sounds are not defined as "not [voiced]"). The details of the features used are described in a later section and can also be studied in chapter 10 of Patterns of Sounds. Users should become familiar with the feature structure in order to avoid the possibility of misinterpreting the data.

The sample is designed to include a language from each grouping of languages at about the level of relatedness represented by, say, the West Germanic subfamily of Indo-European languages (German, English, Dutch, Frisian and their close relatives). The sample is thus a quota sample. The target is to include one and only one language from each such identifiable grouping; in this case the chosen language is German. In assembling the database detailed classifications were compared for all language families. One language was then selected from each grouping of the appropriate degree of relatedness. The choice of which language, and dialect, to include within any group is largely determined by availability of good descriptive accounts.

There are some limitations on the achievement of the quota design. These stem from two causes -- uncertainties of classification and lack of descriptive sources. For some families of languages, internal subgrouping is not well established, or is of such a nature that it is hard to determine appropriate sampling levels. In some cases no satisfactory description could be located for any language in a targeted group. In such cases a language may be included based on less than satisfactory sources, or the group may be left unrepresented. Less than 2% of the targeted groups are unrepresented. The influence of problematical subclassification on the sample cannot be easily quantified.

It is important to emphasize that the selection criterion is genetic relationship, not any phonological characteristics of a language or language family. UPSID can thus be used to determine the frequency of segments across the range of human languages. Of course, related or geographically contiguous languages will often share phonological characteristics. Users should bear this in mind when using the database and consider the extent to which shared inheritance or areal spreading of characteristics might account for a pattern of distribution. For some analyses it might be appropriate to use only a subset of the data.

Note that assumptions made about membership in larger groupings do not affect the structure of the sample. For example, it does not matter if Hmong-Mien (Miao-Yao) languages are assumed to be related to Sino-Tibetan, as here, or if they are assumed to belong with Austro-Tai languages, as in the unified classification used by Ruhlen (1988). In either case, the subgroupings within Hmong-Mien are what matters for inclusion in the sample, not the higher affiliation. Major family affiliation does, of course, affect hypotheses about possible inherited similarities between languages.

The UPSID program enables a user to add languages, change the entries for previously entered languages, and redesign the feature set used. A variant, called PHONEME, does not allow changes to be made. This version is recommended for installations where less sophisticated users will be using the program. Both programs allow considerable analysis to be carried out by selecting and subsetting the data. Other analyses can be carried out by generating an external data file suitable for reading into a standard statistics package, such as SAS.

Maddieson, Ian. 1984. *Patterns of Sounds*. Cambridge University Press, Cambridge.

Ruhlen, Merritt. 1988. *The World's Languages, Volume 1, Classification*. Stanford University Press.

II. Instructions for installation

The program is distributed on 5 1/4 or 3 1/2 inch diskettes.

You must have:

- an IBM PC or compatible running DOS
- at least 1.1 Mbytes available on your hard disk or diskette; more space will be required to install all the ancillary language information files, but the program will run with this.
- 360K available memory
- a graphics monitor (either color or monochrome)

The UPSID program itself is in a file called UPSID.EXE. Copy this to whatever directory you want it to run from. For the sake of example, assume you copy it to **c:\upsid**. Use these commands:

```
md c:\upsid          (makes a subdirectory called upsid)
copy a:\upsid.exe c:\upsid      (copies program file)
```

To install PHONEME, copy the file called PHONEME.EXE into whatever directory you want to run it from. For example, use this command:

```
copy a:\phoneme.exe c:\upsid  (copies program file)
```

The database is in four files called CHARCOD.STR, LANGS.STR, FEATS.STR, and SEGS.STR. These must be in the directory you are in when you run UPSID or PHONEME; it is suggested that they be copied to the same directory as UPSID or PHONEME, i.e. **c:\upsid** in our present example. Use this command:

```
copy a:*.str c:\upsid
```

Each language has an associated file with the extension .INF containing general information such as references, comments, and genetic classification. UPSID looks first for these language information files in a subdirectory **\lginf** inside whatever the working directory is; although you may specify a different directory name for UPSID to look in, it is most convenient just to put them in a subdirectory **\lginf** in the same directory as UPSID.EXE, i.e. copy all the files ending in .INF into **c:\upsid\lginf**. Use these commands:

```
md c:\upsid\lginf    (makes a subdirectory called lginf)
copy a:\lginf\*.inf c:\upsid\lginf  (copies info files)
```

To run UPSID, just type UPSID from within the **\upsid** directory. To run PHONEME, just type PHONEME from within the **\upsid** directory.

The UPSID database is stored in four files called LANGS.STR, CHARCOD.STR, SEGS.STR, and FEATS.STR. These files may be printed for reference (for example to review the list of languages that are in the database) but they should NOT be edited, since UPSID will be unable to cope with any changes in their format.

In addition, there is a small file for each language containing general information, such as the source of the data in UPSID on that language and any comments. Not all these files are completed in version 1.1. These are ordinary text files; they may be edited using a simple editor incorporated within UPSID. They may also be edited outside UPSID if so desired, since UPSID only reads this information in. The only restriction is that each file may be at most 6000 characters long. Longer files cannot be read from within the UPSID. These file names are of the form **LG####.INF**, where **####** is the language's identification number (see below). Thus the file for

Greek would be called LG2000.INF. These files also contain the language name and number and any alternate names. If you change the language name or number elsewhere in UPSID, they will not be changed inside the language information file; you must remember to change them there yourself by editing the information file.

The four main files are structured as follows:

LANGS.STR: This file contains records containing the language name, the four-digit UPSID language identification number (see section VI below), and the number of segments in the language. The records are stored in alphabetical order by language name. Version 1.1 contains data on 451 languages.

CHARCOD.STR: This file contains records each containing a character code name, a string of feature values, the number of occurrences of that character code in the database (i.e. in SEGS.STR), and a standard phonetic description, in words, of the segment represented. Each record describes one phonological segment, divorced from any language, and can be thought of as a phonetic symbol. Each "symbol" is described by a unique set of features chosen from a fairly standard set, encompassing manner and place of articulation, secondary articulations, vowel features, phonation type, and so on. These are described in section VII below. To save space, the feature values are not stored as 1's and 0's but as hexadecimal digits (0-9 and A-F) which can be expanded into four 1's and 0's. For example, if the first character in the feature string is 'A', this means the first four features have the values 1,0,1,0; '3' would map to 0,0,1,1. The records are stored in order of decreasing feature strings, or reverse alphabetical order (by feature string, not by character code name). In this way, character codes turn out to be ordered by manner of articulation, which, barring any great changes from UPSID's original design, is:- plosive, implosive, ejective stop, click, fricative, ejective fricative, affricate, ejective affricate, affricated click, unspecified r-sound, tap, flap, trill, approximant, nasal, simple vowel, diphthong, lateral, sibilant. Within each manner, character codes are stored by place of articulation, from front to back; within each place, by secondary articulation; then by vowel features (if any) and finally by phonation type and miscellaneous.

SEGS.STR: This file contains records each containing a language number, character code name, and anomaly value, so each record represents one segment in a particular language. Anomaly values are assigned as follows (from Patterns of Sounds, I. Maddieson, 1984, p. 170):

- 1: indicates a segment of extremely low frequency in the language (e.g. it only occurs in a handful of words or certain morphological markers, but these are well entrenched parts of the language).
- 2: indicates a segment that occurs only in foreign words or unassimilated loans but these are frequent enough to consider including the segment in the inventory.
- 3: indicates a segment which is posited in underlying forms to account for some phonological patternings but which is neutralized in surface forms (very rarely used).
- 4: indicates a segment which is treated as phonemic in UPSID but which may be regarded as derived from other underlying segments (very rarely used).
- 5: indicates a segment which although apparently a genuine member of the inventory, is described in particularly obscure or contradictory fashion (e.g. a segment in Ashuslay, 6814, described as simultaneously a (velar) stop and a lateral).

These records are in order by increasing language number, and within each language by order of entry into UPSID of the segments.

FEATS.STR: This file contains records each containing a long feature name, short feature name, and group number. The long feature name is the name normally used, e.g. "plosive"; the short feature name is an abbreviation of at most 8 characters, for example "plos", which is intended to be used as a variable name in statistics packages which import UPSID output files. The group

number describes what type a feature is and thus allows features of the same type to be stored together, and also facilitates the storing of character codes in CHARCOD.STR in order by manner, place, and so on. The feature group numbers are 1 for manner of articulation; 2 for place of articulation; 3 for secondary articulation; 4 for vowel (e.g. "high"); and 5 for phonation type and miscellaneous. Features are stored in order by increasing group number and within each group number by order of insertion into UPSID.

OUTPUT FILES: UPSID output files are structured so as to be easily imported into an outside statistics package, with one exception noted in the next paragraph. For specific information on types of output files, see section V below. In all output files containing segment records there is a list of variable names at the beginning of the file, in the same order as the fields in the rest of the file. The variable name list is followed by a semicolon and then the data records. Each record is one segment and contains the language number, the character code, the anomaly value, and 0's and 1's for the value of each feature. Each variable within a record is separated from the others by at least one space, and each record is separated from the other records by a blank line. Such files can be read directly into the SAS statistics package, and several other statistical analysis programs. This is also the file format required for files serving as input to the "Select from previous output" option. If the file's format has been changed since it was output, the "Select from previous output" routines may not be able to cope and may behave incorrectly and/or unpredictably.

The exception to the above format is a file output by the "Select subset of database - select character codes with some particular set of features" command. This file is designed to be human-readable rather than statistics-package-readable; it contains the names of the selected character codes, the number of occurrences of each in the database, and the phonetic descriptions, in words, of each. This is exactly the same information as is printed to the screen with this option. The file is provided for convenience and is not expected to be used for further analysis.

IV. How UPSID and PHONEME menus and input screens work

Menus show a list of items of which you may choose one; this is typically a list of actions. You may move around the list by using the up and down arrow keys or by typing the first letter of an item; the item that's currently selected is highlighted or in a different color from the others. To tell UPSID to carry out the selected action, press the <enter> key. If you end up in a menu or input screen where you don't want to be, you may press <esc> from almost any menu or input screen to tell UPSID to abort whatever it's doing. Note that UPSID must be asking you for input or for a choice in order for <esc> to cause an abort; if UPSID is working and the screen is blank, <esc> will not stop UPSID but will be held in a buffer as the answer to the next thing UPSID asks you, whatever that might be. The only place <esc> will not cause UPSID to abort its current action is in the editors which allow (nearly) unlimited input, where <esc> means "end of input"; the editors are used in editing general language information, in entering a list of character codes for selecting, and in entering a combination of features for either selecting or defining a new feature. As a reminder, whenever an editor is displayed on the screen, the instructions that are also displayed tell you to press <esc> to perform a normal exit.

Input screens typically have one or more fields where you may enter information. This is usually done by typing words or numbers but a few screens require special input (for instance using the spacebar to check off features when defining a new character code); in these cases instructions are given on the screen. For some fields only certain kinds of input are legal; for instance, a language number may only contain digits. These fields do not accept illegal input, so if you type a letter into a language number, UPSID will simply ignore the letter. When you are finished entering information into a field, use the up and down arrow keys or the <enter> key to move to other fields. Some input screens require all fields to be filled (for example when adding a new language you must give both a language name and number), and these screens will not go away until you

either fill all the fields and press <enter>, or press <esc> to abort. In input screens which may be only partially filled, if you press <enter> while in the last field, UPSID will take this to mean you've finished with the input screen.

A few screens contain more information than can be shown on a single page; these screens will have either a scroll bar (a fat bar on the right border which shows what part of the whole screen you're seeing) or scroll lights (small upward- or downward-pointing triangles in the left margin which tell if there's more information above or below what you're seeing). PG UP and PG DN keys move up and down in these screens.

V. What UPSID actually does

This section contains a description of every menu option in UPSID and is intended as a detailed reference.

1.0 Language

1.1 Display segments

This option displays the segments identified by their character codes currently stored in the database for the language you specified. It can display either all the segments in the language or only those with an anomaly value of other than 0. The anomaly values are only shown in the "only anomalous segments" screen; they appear as a number preceding the code they apply to.

Some languages have more segments than can be displayed on the screen at one time. For such a language the box containing the segments will have a small upward- and/or downward-pointing triangle in the left margin; the triangles indicate "more data above" or "more data below". To view segments not on the screen, use PG UP and PG DN keys (not the arrow keys). Note that UPSID will always start out by displaying the last page of segments.

1.2 General information

This option allows you to read and/or edit general language information. Information on some languages is supplied with UPSID; you may add information on other or new languages. Each language is in a file by itself, which UPSID expects to find in a subdirectory \LGINF. The file names are of the form LG####.INF, where #### is the language number. These are ordinary text files and may be edited outside UPSID if so desired, since UPSID does not use this information; the only restriction is that each file may be at most 6000 characters long.

If a file exists for the language you have specified, UPSID will open the file and print the information on the screen. If no file exists and you tell it to make a new file, UPSID will create the file and write a suggested structure for information in it. In either case, UPSID will then put you into a special, rather primitive text editor; instructions for using most of the capabilities of the editor are on the right of the screen. A few functions are not displayed, for reasons of screen space and obviousness or infrequency of use. These are:

- alt-c: turns marking by column on and off
(alt-m turns marking by row on and off)
- keypad minus: cuts marked region
- ins: inserts the cut buffer
- keypad plus: copies marked region
(possible nonfunctional)
- arrow keys: as expected
- HOME, END: moves to the beginning or end of the line
- PG UP, PG DN: moves up or down a page, or as much as possible if less than a page.

The only way to exit the editor is to press <esc>, which will save whatever is in the editor to the language file.

1.3 Add (a language). Not an option in PHONEME.

This option allows you to enter a new language name and number and then simply calls the "add segment" function repeatedly, until you press <esc>. The only caveat here is that this option should be exited in the conventional manner (i.e. <esc>. Try to avoid exiting by rebooting, power failure, and so on.) since before it begins adding segments it splits the segment database file, SEGS.STR, into two parts, and puts the parts back together after you have pressed <esc>. This was done to effect a substantial increase in the speed of adding a language. If a power failure or other untoward event should occur, all the data for languages with a higher number than the one being added will be in a file called UPSID.TMP. You may attempt to concatenate UPSID.TMP onto SEGS.STR using DOS or an editor, or simply copy an old version of SEGS.STR off your backup. UPSID.TMP is used as a temporary file for other purposes too, so do not add it to SEGS.STR unless the above occurred.

1.4 Delete (a language). Not an option in PHONEME.

This is a self-explanatory option. In case you choose this menu item by mistake, UPSID queries you again before it actually deletes the language.

1.5 Edit (a language). Not an option in PHONEME.

This brings up a submenu of three items: edit a language name or number, add a new segment, or delete a segment. This option can be used for comparatively trivial changes, such as choosing an alternate spelling of the language name, or can be used to alter the entire phoneme inventory entered for that language.

1.5.1 Edit a language name or number

This allows you to change either the name or the number of a language. It is also useful because it displays the language name, number, and number of segments, so this may be used to find out what language corresponds to a particular number. Changing a language's number is a slow process and you might therefore want to avoid doing so often.

1.5.2 Add a new segment

This option is used to add a segment to a language -- not to be confused with defining an entirely new type of segment, called a character code, which is independent of any language and is UPSID's representation of a phonetic symbol. However, if you want to add a new segment to a language and find that the character code you want to use for it is not defined, you may define it at this point.

After you have entered the language name and/or number, UPSID will ask if you want to see the segments currently stored for the language. If the language has more segments than can be displayed on the screen at one time, UPSID will start by displaying the last page of segments. A small upward- and/or downward-pointing triangle in the left margin of the box containing the segments will indicate "more data above" or "more data below". To view segments not on the screen, use PG UP and PG DN keys (not the arrow keys).

1.5.3 Delete a segment

This is similar to the "add a new segment" option described immediately above in section 1.5.2. Note that if the last remaining occurrence of some particular character code is deleted, the character code will still remain in the database, though unused. It can be deleted under the Character code menu.

1.6 Examine name and number. PHONEME only.

Allows the pairing of language name and number to be checked in PHONEME.

2.0 Character code

2.1 Add (a character code). Not an option in PHONEME.

This allows an entirely new segment type, UPSID's representation of a phonetic symbol, to be defined for later use. After you enter the name for this segment type, UPSID gives you a menu of all currently defined features; those which are to have the value 1 (or "present") should be checked off by pressing the spacebar. Keys you may use to move around the list are PG UP, PG DN, HOME, END, and up and down arrows (not right and left arrows). The thick scroll bar on the right border of the box containing the features moves to indicate what portion of the feature list you are seeing.

Several "fail-safe" properties are built into the functioning of this command. If the character code you want to add has already been associated with a different set of features, or the set of features has already been associated with a different character code, UPSID will tell you this and will not let you add the new one. If there is already a character code with the same name, you can see what its features are by using the "edit character code" option; if there is already a character code with the same set of features UPSID will tell you what the name of that code is.

2.2 Delete (a character code). Not an option in PHONEME.

This option deletes a character code if it is not used in the database. A character code that is used for a segment in some language in the database is prevented from being deleted.

2.3 Edit (a character code). Not an option in PHONEME.

This gives you the opportunity to change the name of a character code, its phonetic description, or the set of features associated with it. The display of features works exactly as for adding a character code and is described above in section 2.1.

2.4 Examine (a character code). Only in PHONEME.

This permits the phonetic description for any character code to be read, and the set of features that are positive for that segment type to be viewed. PHONEME does not permit these assignments to be changed. It also reports the number of occurrences in of the character code in the database.

3.0 Feature

3.1 Add (a feature). Not an option in PHONEME.

This option allows you to add a new feature to all the character codes stored in the database. The feature may be either entirely new (for example a new vowel height) or defined as dependent on some combination of existing features (for example a feature "lip-involvement" which is 1 for each character code which is bilabial, linguolabial, labiodental, or labialized or rounded).

For an entirely new feature, UPSID asks you for the full name of the feature ("long" name, e.g. linguolabial) and a short feature name (e.g. linglab) which will be used in output files and is intended for use as a variable name by statistics packages. Then it will ask what group of features this feature belongs with (manner of articulation, place of articulation, secondary articulation, vowel, or phonation type and miscellaneous). The new feature will be stored as the last feature within its group, hence may be somewhere in the middle of the feature strings in CHARCOD.STR. UPSID will make room for the feature, and set the feature value to 0, in every character code record. If you then want to set that feature to 1 for certain character codes, you must do so using "Edit character code".

For a feature whose value is defined by some combination of other features, UPSID requests the long and short names and group membership of the feature exactly as above. After it has made space for the feature in every character code, however, it opens a text editor which you may use to enter the defining combination of features. You may use short feature names, parentheses, and "AND", "OR", and "NOT". A lip-involvement feature, for instance, could be defined as

bil or linglab or labdent or labzd

If you wanted to define a feature for non-bilabial voiceless plosives, you could write

plos and vl and not bil

Non-bilabial, non-uvular voiceless plosives could be defined as

plos and vl and not (bil or uv)

or

not bil and not uv and plos and vl

or

(not bil and plos and vl) and (plos and vl and not uv)

or ...

When the expression you enter is evaluated, NOT takes highest precedence, then AND, and finally OR, exactly as multiplication takes precedence over addition when you write a mathematical expression without parentheses. As in math, you may also use parentheses to change precedence;

bil or (uv and vl)

(all segments which are either bilabial or are voiceless uvulars)

is the same as

bil or uv and vl

but not the same as

(bil or uv) and vl

(all segments which are voiceless and either bilabial or uvular).

Remember that the expression you enter into the editor may be preceded by "Set the new feature to 1 in all segments which are ..." -- thus, it doesn't make sense to write

bil and labdent

since there is no segment which is both bilabial and labiodental.

If there is a syntax error in the expression you enter, UPSID will tell you as best it can what went wrong and set the feature to 0 for all character codes. To try again, use the "Delete feature" option to get rid of it and then "Add feature" to start over.

3.2 Delete (a feature). Not an option in PHONEME.

This removes a feature from the feature strings stored with each character code and from the feature file. It does not, however, check to see if any distinctions are destroyed when the feature is removed; thus if the feature "long" were removed, the character codes p and p: would have exactly the same features.

3.3 Edit (a feature). Not an option in PHONEME.

Using this option you may change the long or short name of a feature. If you want to change the feature group it belongs to, we suggest you add a new feature with some temporary name, in the feature group you want, and define it to be always the same as the old feature (just enter the old feature name in the text editor for the defining combination of features). Then delete the old feature and rename the new one. This is different from just deleting the old feature and adding it again because the latter method would cause that feature's value to be set to 0 for all character codes.

3.4 Examine (a feature). Only in PHONEME.

PHONEME does not allow the feature set to be altered or edited, but the group to which a feature is assigned and the pairing of long and short names can be examined using this option.

4.0 Select subset of database

This option is the main tool within UPSID and PHONEME for investigating patterns of segment occurrence in the database. There are a number of ways of selecting a subset of the database, which appear on a submenu when this option is chosen: you may select :-

(1) languages which (a) have one or more segments with some particular combination of features, (b) do not have any segments with some particular combination of features, (c) have one or more particular segments specified by the user, (d) do not have any of a particular set of segments, (e) have particular language numbers, or (f) have other than particular language numbers.

(2) segments which (a) have some particular combination of features, (b) do not have some particular combination of features, (c) are one of a set specified by the user, or (d) are not one of a set specified by the user.

(3) character codes with some particular combination of features.

Note that the output will always go to a new file, with a name you specify. If you want to concatenate output files this may be done in an outside editor, DOS, or other program.

For the "select languages by number" options, a count is kept of how many languages are (or are not) in the set of language numbers you defined (more on this in section 4.5). For the "select character codes by features" option, the number of character codes selected will be counted and displayed to the user at the end of a list of the character codes. The "select segments - display only" options keep count of the number of segments selected, and display the total on the last data screen. For the remaining selection options, a count will be made of both the number of character codes you used to select the subset and either the number of occurrences of those (or of other) character codes or the number of languages containing one or more (or not containing any) of those character codes. This is useful for instance if you want to know how many distinct types of plosives there are in UPSID, as opposed to the total number of plosive segments. (Note, however, that character codes with no actual occurrences in the database will be included in the "number of character codes" count, if there are any such codes. Non-occurring codes result when all their occurrences have been deleted, by deleting either particular segments or entire languages. There are none in UPSID version 1.1 as supplied.)

4.1 (Selecting) languages which have segments with some particular combination of features

This option will count the number of segment types (character codes) which are consistent with some particular combination of features and the number of languages which contain one or more of these segments, and output the entire inventory of those languages if desired. The combination of features is entered into a simple editor in UPSID, using short feature names, parentheses, and the words "AND", "OR", and "NOT". See section 3.1, Adding a feature, for a discussion of how a combination of features may be expressed. If a syntax error occurs, UPSID will diagnose the error as best it can and then return to the main menu. Remember that the expression you enter may be preceded by "Select those languages which have some segment which is ..."

4.2 (Selecting) languages which do not have any segments with some particular combination of features

This option will count the number of segment types (character codes) which are consistent with some particular combination of features and the number of languages which do not contain any of these segments, and output the entire inventory of those languages if desired. The specification of features is as described above in section 4.1. Remember that the expression you enter may be preceded by "Select those languages which have no segment which is ..."

4.3 (Selecting) languages which have one or more particular segments

This is similar to the option described above in section 4.1, "Selecting languages which have segments with some particular combination of features". What you enter into the UPSID editor, however, is just a list of character codes, separated by at least one space, and what will be output is the inventories of those languages containing at least one of the character codes specified. You may not directly select, for instance, those languages which have /t/ but not /s/. To do this, pick whichever condition is more restrictive (currently UPSID has 178 /t/s and 195 /s/s, so in this case pick /t/) and select all languages containing that segment. Then the subset of those languages which also have /s/ may be extracted using the "Select subset from previous output - select languages with particular segments" option.

4.4 (Selecting) languages which do not have any of a set of particular segments

This is similar to the above option in section 4.3 except that only those languages which do not have any of the character codes you entered into the UPSID editor will be counted and output.

4.5 (Selecting) languages which have particular language numbers

You may choose this option to select particular languages, language families, or continents. There are four blanks which may be filled in, each of which must either be full (all four digits) or empty. If you don't care what digit occurs in a particular position in a language number, use 'x', e.g. 6xxx will select any language whose language number has the first digit 6. Count will be kept of how many such languages are selected.

4.6 (Selecting) languages which have other than particular language numbers

This option is identical to that in section 4.5 except that only those languages whose numbers you did not specify are output.

4.7 (Selecting) segments with some particular combination of features

The specification of features works exactly as in the option in section 4.1. This "select segment" option presents you with two choices as to what to actually do with the selected segments: you may

- (1) count the number of character codes defined by some combination of features and the number of occurrences of those segment types, and output the required segment records (not the entire languages) to a file, along with complete feature specifications for each segment, or
- (2) count the number of occurrences of the defined character codes, and display (on the screen) each segment, the name of the language it occurs in, and its anomaly value. Only a screenful of data is displayed at a time, but while one screen of data is being displayed UPSID continues searching for the next screenful, and will tell you when the next screen of data is ready. Note that you can choose to view the next screen or to exit to the select menu, but you can't view a previous screen without restarting the search.

These two options take approximately the same amount of computer time, but the second takes less of your time if you are only interested in finding out what languages have a particular segment and not in doing any further statistics.

The feature combination expression you enter into the editor may be preceded by "Select those segments which are ..."

4.8 (Selecting) segments which do not have some particular combination of features

This option is the same as that in section 4.7 except that only those segments which do not have the specified combination of features are counted and displayed or output. The feature combination expression you enter may be thought of as being preceded by "Select those segments which are not ..."

4.9 (Selecting) segments which are one of a set you specify

This option is analogous to that in section 4.5 above, except that the output choices are the same as in section 4.7, i.e. either the segment records with their full feature specifications, language

numbers, and anomaly values may be output to a file, or the segments with their language names and anomaly values may be displayed, one screen at a time.

4.10 (Selecting) segments which are not one of a set you specify

Analogous to the option immediately above in section 4.9, except only those segments you did not specify are counted and displayed or output.

4.11 (Selecting) character codes with some particular combination of features

This works similarly to the options above in sections 4.1 and 4.7. Differences are that only those character codes (not any actual occurrences of them) meeting the feature conditions specified are output; and count is kept only of the total number of such character codes. The character codes are always displayed, and may be output to a file as well. The display and file contain the names of the character codes, the number of occurrences of each in the database, and the phonetic description of each. If there are too many character codes to be displayed on the screen at once, they will be displayed one screen at a time as described in section 4.7.

The feature expression you enter into the editor may be thought as being preceded by "Select those character codes which are ..."

5.0 Select from previous output

The important feature of this option is that it allows you to choose restricted subsets of UPSID data which cannot be chosen using the "select subset of database" options alone. For instance, if you're interested in the vowels in a particular language family, you can select (using "select subset of database") all the languages in that family first, and then select (using "select from previous output") the vowels from that output file. Another example would be to select ("select subset of database") all the languages with voiceless nasals, then ("select from previous output") all the languages without voiced nasals, which would result in a file you could use to study the occurrence of voiceless nasals without their voiced counterparts. The "select from previous output" option can in fact be applied repeatedly, so you could then select languages from a particular family or continent if so inclined.

One thing not to do if planning to use this option is edit the previously output file which you are going to use as an input file to this step. (Note also that the input file **MUST** be an UPSID output file.) UPSID has no real checks to make sure the file is in the right format -- the only landmarks it looks for are the semicolon at the end of the variable list at the beginning of the file, and the blank lines between segment records -- and if the file is not in exactly the same format as it was when output, unpredictable and incorrect things may happen.

It should also be noted that the most restrictive selection should be performed first, using "select subset of database", since searching time will be reduced as the size of the file to search in is reduced.

This option presents you with a submenu of choices which is the same as that available under "select subset of database" except that you may select either languages or segments but not character codes. The choices behave in the same manner as those under "select subset of database", and therefore will not be discussed again here, except for two differences. The first difference is in the "select languages which ..." options: the only data available to select from in this step is that in the input file; therefore, it follows that if not all the segments in a language are in the input file, not all the segments in the language will be in the output file. The second difference is in the "select segments which ..." options: you may count the selected segments and either output them in statistics-package-readable format to a file or not; but you may not choose to display the segments on the screen with the names of the languages they occur in.

5.1 (Selecting) languages which have segments which have some particular combination of features

See section 4.1.

5.2 (Selecting) languages which do not have any segments with some particular combination of features

See section 4.2.

5.3 (Selecting) languages which have one or more particular segments you specify

See section 4.3.

5.4 (Selecting) languages which do not have any of a set of particular segments

See section 4.4.

5.5 (Selecting) languages which have particular language numbers

See section 4.5.

5.6 (Selecting) languages which have other than particular language numbers

See section 4.6.

5.7 (Selecting) segments which have some particular combination of features

See section 4.7, with the difference that the results may either be counted and output to a file or just counted, but not displayed on the screen.

5.8 (Selecting) segments which do not have some particular combination of features

See section 4.8, with the difference that the results may either be counted and output to a file or just counted, but not displayed on the screen.

5.9 (Selecting) segments which are one of a set you specify

See section 4.9, with the difference that the results may either be counted and output to a file or just counted, but not displayed on the screen.

5.10 (Selecting) segments which are not one of a set you specify

See section 4.10, with the difference that the results may either be counted and output to a file or just counted, but not displayed on the screen.

6.0 Miscellaneous counts

6.1 Number of occurrences of a particular character

The number of occurrences of a particular character code can also be obtained using the "edit character code" option, but these options are not identical. "Edit character code" simply reads the number of occurrences stored in the character code record and assumes it is correct. This option reads through the entire database and actually counts the number of occurrences and writes this number to the character code record. If you suspect that for some reason the number of occurrences associated with some character code is incorrect, this option may be used to find the actual number and to correct the stored data.

6.2 Number of languages in database

This option doesn't actually count the number of languages in the database; it divides the length of the file LANGS.STR by the size of a language record and prints that as the number of languages. This could, therefore, be incorrect if LANGS.STR was corrupted during editing or some other process.

6.3 Number of segments in a language

This option bears exactly the same relationship to "edit language" as the option immediately above in section 6.1 bears to "edit character code". Use this option to check and if necessary correct the number of segments associated with a language.

7.0 UPSID output files

With one exception, UPSID output files are structured so as to be easily imported into an outside statistics package. At the beginning of the output files is a list of variable names, in the same order as the fields in the rest of the file. The variable name list is followed by a semicolon and then the data records. Each record is one segment and contains the language number, the character code, the anomaly value, and 0's and 1's for the value of each feature. Each variable within a record is separated from the others by at least one space, and each record is separated from the other records by a blank line. This is also the file format required for files input to the "Select from previous output" option. If the file's format has been changed since it was output, the "Select from previous output" routines may not be able to cope and may behave incorrectly and/or unpredictably.

The exception to this structure is the file output by "Select subset of database - select character codes with some particular set of features". This file is designed to be human-readable rather than statistics-package-readable; it contains the names of the selected character codes, the number of occurrences of each in the database, and the phonetic descriptions, in words, of each. This is exactly the same information as is printed to the screen with this option and the file is merely provided for convenience.

VI. Assignment of language numbers

Each language receives a unique four-digit number when it is first entered into UPSID. The language number reflects genetic and gross geographic classification and thus may be used when, for example, a particular family is to be output for further analysis.

The first digit represents the continent on which the language is spoken: 2 for Eurasia, 4 for Africa, 6 for the Americas, and 8 for Australasia. This however forces all of each language family to be on a single continent; so for instance Austro-Tai is classed as a Eurasian language family despite its spread into Australasia.

The first and second digits together describe the family a language belongs to; those combinations of first and second digits in use in UPSID are shown below:

20xx Indo-European
 21xx Ural-Altaic
 23xx Austro-Asiatic
 24xx Austro-Tai
 25xx Sino-Tibetan
 26xx Caucasian
 27xx Other Eurasian minor families
 29xx Dravidian

41xx Niger-Kordofanian
 42xx Nilo-Saharan
 43xx Afro-Asiatic
 49xx Khoisan

61xx Na-Dene
 67xx N. Amerind
 68xx S. Amerind
 69xx Eskimo-Aleut

83xx Australian
 86xx Papuan

The third and fourth digits have no particular meaning and serve only to distinguish languages within a family.

VII. Meaning of character codes

Character codes use only basic ASCII symbols. The symbols are chosen where possible to be reminiscent of the characters that are recommended for use in the International Phonetic Alphabet. However, with over 900 segment types to be disintegrated a certain amount of arbitrary coding is almost unavoidable. The following is a list of secondary or diacritic marks, used in association with a base character shown here as 'x'. This list is not exhaustive. Note in particular the special use of quote marks. A full listing of all the character codes used in UPSID version 1.1 is provided as an appendix to this document. This appendix provides the phonetic description and a more standard phonetic transcription equivalent for each character code.

x-	velarized
xW	labialized
xJ	palatalized
x*	laryngealized
x?	glottalized (of clicks)
x9	pharyngealized
x~	nasalized (of flaps, approximants, vowels)
nx	nasalized (of clicks)
gx	voiced (of clicks)
xh	aspirated or breathy
xhh	with breathy release
hx	preaspirated (of stops) or voiceless (of nasals, approximants, vowels, clicks)
xD	dental
"x	unspecified dental or alveolar for consonant; undifferentiated 'mid' for vowels
x_	palato-alveolar or backed (of vowels)
x,	palatal
xj	palatal
x.	retroflex
x<	implosive
x'	ejective
xF	fricative
xA	approximant
x(unrounded
x)	rounded
x+	fronted (of vowels)
x:	long
xS	overshort (of vowels)

VIII. UPSID features

The following features are UPSID's basic (non-user-defined) set; this description is modified from Patterns of Sounds, I. Maddieson, 1984. The long feature names are given first, followed by the short feature names in parentheses and any necessary explanation.

A. Features indicating manner of articulation

Conventional phonetic labels, such as "plosive", "click", "vowel", generally combine information on aperture and airstream. This set of variables does likewise. All segments, except /h/ and its variants, are marked for at least one of these; rarely, a segment may be marked for more than one, e.g. a fricative trill.

Plosive (plos) : Pulmonic egressive stops, including glottal stops, have this feature. Note that nasals are not considered to be stops of any sort.

Implosive (imp) : Glottalic ingressive stops, whether voiced or voiceless, have this feature.

Ejective stop (ejstop)

Click (click) : Non-affricated clicks have this feature.

Fricative (fric) : Pulmonic egressive fricatives have this feature. Note that /h/ and its variants are not considered fricatives.

Ejective fricative (ejfric)

Affricate (aff) : Pulmonic egressive affricates have this feature.

Ejective affricate (ejaff)

Affricated click (affclick)

Unspecified r-sound (unspecr) : This feature is used for segments which are simply identified as some kind of "r-sound" (e.g. they are transcribed /r/ or called a "vibrant") but cannot be further classified as a trill, tap, flap, approximant, etc.

Tap (tap)

Flap (flap)

Trill (trill)

Approximant (approx) : This feature is used for all consonants produced with open approximation, e.g. "semi-vowels", r-glides, nonfricative laterals, etc.

Nasal (nas) : Nasal consonants (with complete oral closure) have this feature. Nasalized segments, including consonants with nasal onsets, are identified with a different feature (nasalized).

Simple vowel (vowl) : Monophthongal vowels have this feature. A segment which is not identified as either a simple vowel or a diphthong is thereby classified as a consonant.

Diphthong (diph) : This feature is used for unit diphthongs. Phonetic diphthongs which are phonologically analyzed as the result of a juxtaposition of simple vowels or a vowel and an approximant are, of course, not units but sequences and their components will be entered as separate phonemes.

Lateral (lat) : This feature is used for all lateral segments, e.g. it indicates lateral release in lateral affricates.

Sibilant (sib) : This feature serves to identify the class of sibilants within the fricative/affricate group. In many languages this phonetic property is functionally redundant as place of articulation will distinguish the same class. However, it has occasional distinctive function in differentiating between fricatives and/or affricates with similar place of articulation but distinct acoustic characteristics (e.g. /s/ vs. theta /θ/).

B. Features indicating place of articulation

All consonants except /h/ and its variants are specified for at least one place of articulation. Places are listed in the conventional front to back sequence. Double articulations are indicated by specifying two places of articulation, but secondary articulations are indicated by use of a separate

set of variables. /h/ and its variants, having place determined by environment, do not receive any place specification.

Bilabial (bil)

Labiodental (labdent)

Linguolabial (linglab)

Dental (dent) : This feature is used for true dental consonants. Specific indication that a segment is dental must be provided in the source, for example if it says "the tongue touches the teeth" or "/t/ is like French t , not like English t ". The description "dental" is often applied to a segment which is more accurately described as alveolar, hence, unless such added evidence is available, segments that are merely called dental are not assumed to be true dentals.

Unspecified dental or alveolar (unspdent) : Segments simply indicated by transcriptions using /t,d,s/ etc. may be either dental or alveolar. In order not to falsify the data, this feature is used to indicate such segments with an incompletely specified place of articulation. Segments described simply as "dentals" are also included here.

Alveolar (alv)

Palato-alveolar (pal_al) : Palato-alveolar and alveo-palatal are not distinguished.

Retroflex (ret)

Palatal (pal)

Velar (vel)

Uvular (uv)

Pharyngeal (phar)

Glottal (glot) : This feature is only used for glottal stops; the characterization of /h/ and its variants as "glottal fricatives" is rejected.

C. Features indicating secondary articulations or accompanying features

Labialized (labzd)

Palatalized (palzd) : This feature is used only for true palatalized consonants, i.e. those with a secondary palatal articulation. Thus a segment /c/ which occurs in a language as part of a "palatalized" series of stops /p^j, t^j/ etc., will not be coded with this feature. Instead it will be reported as a palatal stop.

Velarized (velzd) : This feature is only used for true velarized segments. This feature is also used to characterize vowels with velar stricture, reported in Siriono (6829).

Pharyngealized (pharzd) : Unlike the preceding secondary articulation features, which usually are used only for consonants, this feature is used equally for consonants and vowels.

Nasalized (naszd) : This feature is used for nasalized consonants and vowels, i.e. those with simultaneous nasal and oral escape. It is also used to characterize prenasalized stops (when these are clearly units). Thus this feature in combination with any stop feature (except click or affricated click) indicates a nasal onset to the stop.

Nasal release (nasrel) : Used for postnasalized segments only.

Prestopped (pstop) : Used for sonorant segments which begin with a brief stop segment.

Lateral release (latrel) : Used for unitary segments with a non-fricative lateral release.

D. Vowel features

All simple vowels are specified by one vowel height feature, one vowel backness feature, and one lip position feature. They may also have other features to indicate other distinctions. Diphthong segments are specified by assigning all vowel quality features needed to describe both their beginning and end points. A set of diphthong features, described below, indicates the order of conflicting specifications.

High (hi)

Higher mid (himid)

Mid (mid) : This feature is used with systematic ambiguity for both those vowels which are indicated as "mid" without further particularization and those which are true mid vowels (i.e. lie between higher mid and lower mid on a height scale).

Lower mid (lomid) : Note that /E/ and /O/ (open o) are considered lower mid vowels, not low vowels.

Low (lo)

Front (fr)

Central (cen) : Note that /a/ in most languages is considered a central vowel, not a back vowel.

Back (bk)

Nonperipheral (nonperif) : This feature is used for "laxed" noncentral vowels which are produced away from the periphery of the vowel space, for example, /I/ and /U/. It may on occasion serve a mainly diacritical function where other features fail to distinguish vowels.

Rounded (rnd)

Unrounded (unrnd)

Lip-compressed (lipcomp) : This feature is used for "labial" vowels that are produced with vertical compression of the lips but no drawing in and forward of the corners of the mouth ("rounding").

R-colored (rcol) : This feature is used for retroflex or r-colored vowels.

The three features for characterizing movement in diphthongs differ from most other features in the inventory, which indicate only presence or absence of the attribute named by the feature. For diphthong features the non-use of a feature may indicate a specific property of a diphthong, namely movement in the opposite direction to that indicated by the feature name. Also, unlike most of the features, they require reference to other marked features for their interpretation. Their function is to indicate which value precedes when conflicting specifications of vowel height, backness, or lip position are given to a single segment. This method of coding diphthongs was adopted in order to avoid a very large number of features.

Backing (backing) : This feature is used when the end point of a diphthong is more back than the beginning, as in /ia/, /eu/, /^hu/, etc. The feature is not used when the end point is either more front than or has the same degree of backness as the beginning, as in /oi/, /ae/, /ou/, etc. Note that only three degrees of backness are considered (front, central, back).

Lowering (lowering) : This feature is used for diphthongs that have an end point lower than their beginning, such as /i^h/, /ea/, etc. It is not used when the endpoint is higher than or equal to the beginning on the five-point vowel height scale used in UPSID, for example /oi/, /ou/, /ai/, etc.

Rounding (rounding) : This feature is used when the endpoint of a diphthong is rounded but the beginning is unrounded, as in /eu/, /ao/, etc. It is not used when the endpoint is unrounded or both the beginning and endpoints are rounded, as in /oi/, /ai/, or /ou/.

Note that a pair of diphthongs such as /oi/ and /io/ receive the same vowel features. This will involve having conflicting features for height, backness, and roundedness; they will be specified as being both high and mid, front and back, rounded and unrounded. The diphthong features interpret these conflicts. Thus /oi/ will have none of the three diphthong features since there is no backing, lowering or rounding movement in this diphthong. Yet because the segment has conflicting features, there must be movement; it has to be in the opposite direction from the feature names, i.e. fronting, raising and unrounding. On the other hand, /io/ is marked with all the diphthong features, indicating that the mid portion of the diphthong follows the high portion, the back portion follows the front portion, and the rounded portion follows the unrounded portion.

E. Features for phonation types, etc.

Voiceless (vl) : This feature is used for all segments in which the vocal cords do not vibrate. The segment /h/ has this feature but no others.

- Voiced (vd)** : This feature is used for all segments in which the vocal cords vibrate, whether as "regular" voicing or in some other mode (creaky, breathy, etc.). The segment /voiced h/ has this feature but no others.
- Aspirated (asp)** : This feature is used for all voiceless aspirated segments. So-called "voiced aspirates" do not use this feature, but instead the feature "breathy".
- Laryngealized, creaky, glottalized (largzd)** : This feature is used for all segments with a laryngeal constriction in which that constriction is not serving as an airstream initiator or primary place of articulation. Thus, "glottalized" consonants, laryngealized vowels, etc. use this feature. Various distinctions which have been made in the literature between "pre-glottalized", "postglottalized", "creaky", etc. were judged to be inconsistently applied and hence it was decided not to represent such distinctions in the inventory. Some inaccuracies regarding the distinction between a segment which is glottalized and one with a glottalic airstream (an implosive or ejective) have undoubtedly remained.
- Long (long)** : This feature is used for contrastively long vowels or geminate consonants which are single units but are not generated as part of a general series of such long segments either from a suprasegmental length feature or because adjacent identical segments occur. (See "overshort" below.)
- Breathy, voiced aspirated, murmured (breathy)** : This feature is used for all segments characterized by breathiness.
- Overshort (ovrshort)** : In certain languages, the basic series of vowels is longer than a subset which may occur contrastively short. The feature "overshort" is used for such contrastively short segments.
- Preaspirated (preasp)**

IX. Maximum limits in UPSID

Number of features: a maximum of 110 features
 Feature long name: up to 34 characters
 Feature short name: 8 characters
 Character code name: 11 characters
 Number of occurrences of a character code: 9999
 Phonetic description of a character code: 80 characters
 Language name: 39 characters
 Number of segments in a language: 9999
 Size of a language information file: 6000 characters, or bytes

APPENDICES

1. Classified List of Languages Included in UPSID

2. EURASIA

Indo-European (20xx)

Greek:	Greek 2000
Celtic:	Irish 2001, Breton 2002
Germanic:	German 2004, Norwegian 2006
Baltic:	Lithuanian 2007
Slavic:	Russian 2008, Bulgarian 2009
Romance:	French 2010, Spanish 2011, Romanian 2012
Iranian:	Farsi 2013, Pashto 2014, Kurdish 2015, Ormuri 2025

- Indic: Hindi 2016, Kashmiri 2018, Sinhalese 2020, Bengali 2017, Nepali 2023, Konkani 2024
- Albanian: Albanian 2021
- Armenian: E. Armenian 2022
- Ural-Altai (21xx)**
- Yukaghir: Yukaghir 2172
- Ugric: Hungarian 2154, Khanty (Ostyak) 2150
- Finnic: Finnish 2153, Komi 2152, Mari (Cheremis) 2151, Saami (Lappish) 2155
- Samoyed: Nenets (Yurak) 2156, Nganasan (Tavgy) 2157, Selkup 2164
- Turkic: Chuvash 2160, Turkish 2158, Azerbaijani 2159, Uzbek 2175, Bashkir 2163, Kirghiz 2162, Yakut 2161, Tuva 2165
- Mongolian: Khalkha (Mongolian) 2166, Moghol 2174, Dagur 2176, Monguor (Mongor) 2173
- Tungus: Even 2167, Manchu 2169, Nanai (Goldi) 2168
- Korean: Korean 2170
- Japanese: Japanese 2171
- Ainu: Ainu 2140
- Austro-Asiatic (23xx)**
- Munda: Mundari 2300, Kharia 2301
- Khasi: Khasi 2302
- Palaungic: Parauk (Wa) 2309, Khmu? 2312
- Vietmuong: Vietnamese 2303
- Katuic: Pacoh 2314, Bruu 2316
- Bahnaric: Sedang 2304, Sre 2310, Brao 2311
- Khmer: Khmer (Cambodian) 2306
- Pearic:
- Mon: Nyah Kur 2315
- Aslian:
- Nicobarese: Nicobarese 2313
- Austro-Tai (24xx)**
- Lati-Gelao: Gelao 2436
- Li-Kam-Tai: Thai 2400, Lakkia 2401, Yay 2402, Sui 2403, Po-ai 2405, Lungchow 2406, Lue 2427, Lai 2432, Kam 2404
- Atayalic: Atayal 2407
- Tsouic: Tsou 2418, Rukai 2417
- Paiwanic: Paiwan 2434
- W. Malayo-Polynesian: Chamorro 2416, Tagalog 2414, Sa'ban 2415, Malagasy 2410, Javanese 2409, Batak 2413, Cham 2411, Ivatan 2428, Maranao 2445, Tiruray 2431, Sama 2412, Iban 2435
- Central M-P.: Tetun (Timorese) 2437
- E. Malayo-P.: Adzera 2419, Roro 2420, Kaliai 2421, Iai (Iaai) 2422, Hawaiian 2424, Mor 2429, Tigak 2425, Kwaio 2433, Lenakel 2426, Fijian 2439, Pohnpeian 2430, Irarutu (Iritu) 2440
- Sino-Tibetan (25xx)**
- Sinitic: Mandarin 2500, Taishan 2501, Hakka (Kejia) 2502, Changzhou 2503, Xiamen (Amoy) 2504, Fuzhou 2505, Bai (Minkia, Minjia) 2506
- Himalayic: Tamang 2507, Dafla (Nisi) 2508, Newari 2518
- Lolo-Burmese: Burmese 2509, Lahu 2510, Naxi 2525
- Kachin: Jingpho 2511
- Kuki-Naga: Ao 2512, Tiddim Chin (Kamhau) 2513
- Baric: Bodo (Boro) 2515,

Karenic: Karen (Sgaw) 2516, Phlong (Pwo, Pho) 2520
 Hmong-Mien: Mien (Yao) 2517, Hmong (Miao) 2519

Caucasian (26xx)

Caucasian: Georgian (Kartvelian) 2600, Kabardian 2601, Lak 2602, Bats 2604, Avar 2606, Archi 2605, Rutul 2603

Dravidian (29xx)

Brahui: Brahui 2917
 Dravidian: Telugu 2902, Kota 2903, Kurukh 2904, Koya 2905, Tulu 2906

Other Eurasian Families (27xx)

Chukchi-Kamchatkan: Chukchi 2708, Koryak 2709, Itelmen (Kamchadal) 2710
 Paleo-Siberian: Nivkh (Gilyak) 2700
 Ket: Ket (Yenisei) 2706
 Basque: Basque 2714
 Burushaski: Burushaski 2715

4. AFRICA

Niger-Kordofanian (41xx)

Kordofanian: Kadugli 4102, Moro 4101, Jomang 4142
 Mande: Kpelle 4103, Bisa 4104, Bambara (Mandekan) 4105, Dan 4106, Bobo-Fing 4138
 West Atlantic: Wolof 4107, Diola 4108, Temne 4109, Ndut 4149, Konyagi 4145
 Voltaic: Dagbani 4110, Senadi 4111, Tampulma 4112, Bariba 4113
 Kru: Bete 4144, Klao 4133, Aizi 4131
 Kwa: Ewe 4114, Akan 4115, Igbo 4116, Gã 4117, Alladian 4152, Gwari 4140, Yoruba 4137, Isoko 4147
 Ijo: Kolokuma Ijo 4150
 Togo Remnant: Lelemi 4118
 Jukunoid: Kpan 4135
 Cross River: Efik 4119, Ogbia 4124, Kohumono 4136
 Plateau: Birom 4120, Tarok (Yergam) 4121, Amo 4122
 Bantoid: Beembe 4123, Zulu 4126, Teke (Kukuya) 4127, Mambila 4153, Ejagham 4125, Ewondo 4141, Fe'fe' 4148, Aghem 4134, Noni 4139
 Adamawa: Doayo 4128, Mumuye 4132, Mbum 4146, Lua (Niellim) 4151
 Ubangi (Eastern): Gbeya 4129, Azande (Zande) 4130, Sango 4143, Mba 4160
 Dogon: Dogon 4170

Nilo-Saharan (42xx)

Songhai: Songhai 4200
 Saharan: Kanuri 4201
 Maban: Maba (Mabang) 4202
 Fur: Fur 4203
 E. Sudanic: Maasai 4204, Luo 4205, Nubian (Mahas, Nobiin) 4206, Nyangi 4207, Ik (Teuso) 4208, Sebei (Sapiny) 4209, Tama 4210, Temain 4211, Nera 4212, Tabi 4213, Mursi (Murle) 4214, Dinka 4226, Nyimang 4227, Daju 4225
 C. Sudanic: Logbara 4215, Yulu 4216
 Berta: Berta 4218
 Kunama: Kunama 4219
 Koman: Koma 4220

Afro-Asiatic (43xx)

- Semitic: Arabic 4350, Tigre 4351, Amharic 4352, Socotri 4354, Neo-Aramaic 4355
 Berber: Shilha 4356, Tamasheq (Tuareg) 4357
 Cushitic: Somali 4358, Awiya (Awngi) 4359, Iraqw 4360, Beja 4361, Dahalo 4372
 Omotic: Kullo 4362, Dizi 4363, Kefa 4364, Hamar 4365
 Chadic: Hausa 4366, Angas 4367, Margi 4368, Ngizim 4369, Kanakuru 4370, Kera 4371, Dangaleat (Dangla) 4374, Tera 4375, Kotoko 4373, Lame 4376

Khoisan (49xx)

- Hadza: Hadza (Kindiga) 4900
 Sandawe: Sandawe 4902
 Northern Kh.: !X~u 4904
 Central Khoisan: Nama 4906
 Southern Kh.:

6. AMERICAS**Na-Dene (61xx)**

- Haida: Haida 6100
 Tlingit: Tlingit 6102
 Eyak: Eyak 6109
 Athabaskan: Navajo 6104, Chipewyan 6106, Hupa 6105, Ahtna 6108

American I (Northern) (67xx)

- Almosan: Wiyot 6753, Ojibwa 6750, Quileute 6732, Kwak'w'ala 6731, Tseshah (Nootka) 6730, Lushootseed (Puget Sound Salish) 6734, Upper Chehalis 6773, Shuswap 6778, Bella Coola 6771
 Keresiouan: Acoma (W. Keres) 6749, Yuchi 6757, Dakota 6756, Wichita 6755, Seneca 6754, Caddo 6775, Cherokee 6781
 Penutian: Klamath 6707, Nez Perce 6706, Wintu 6709, Maidu 6708, Zuni 6748, Tunica 6758, Wappo 6760, Alabama 6759, Totonac 6713, Mixe 6715, Zoque 6711, Tzeltal 6712, K'ekchi 6714, Tsimshian 6774, S. Sierra Miwok 6779, Huave 6716, Huastec 6776, Yucatec 6777, Jacaltec 6780
 Hokan: Karok 6741, Shasta 6746, Achumawi 6744, Yana 6745, Pomo 6742, Diegueño 6743, Tonkawa 6752, Tol (Jicaque) 6747
 Kiowa-Tanoan: Picuris (Tiwa) 6740
 Uto-Aztecan: Luiseño 6737, Hopi 6738, Pima-Papago 6736, Kawaiisu 6763, Nahuatl 6762
 Oto-Manguean: Mazahua 6717, Mazatec 6727, Chatino 6729, Mixtec 6728, Amuzgo 6770, Tlapanec 6772, Highland Chinantec (Quiotepec d.) 6786

American II (Southern) (68xx)

- Chibchan: Tarascan 6847, Bribri 6801, Paya 6854, Shiriana 6855, Cuna 6857
 Paezan: Itonama 6800, Pirahã 6802, Paez 6804, Cayapa 6803, Warao 6852, Epena Pedee (Saijia) 6811, Guambiano 6862
 Andean: Auca (Waorani) 6818, Arabela 6817, Quechua 6819, Jaqaru (Aymara d.) 6820, Tehuelche (Gununa-Kena) 6821, Mapudungu (Araucanian) 6837, Jebero 6844, Qawasqar (Alakaluf) 6859
 M.-Tucanoan: S. Nambiquara 6816, Ticuna 6831, N. Barasano 6832, Siona 6833, Movima 6860, Cacua 6864, Cubeo 6846, Huari 6841, Iranshe 6845
 Equatorial: Cofan 6836, Jivaro (Huambisa) 6835, Guarani 6828, Siriono 6829, Cayuvava 6840, Camsa 6863, Trumai 6865, Saliba 6861, Ache (Guayaki) 6866

- M.-Arawakan: Guahibo 6830, Amuesha 6824, Campa (Machiguenga) 6825, Wapishana 6822, Moxo (Ignaciano) 6827, Island Carib 6823, Guajiro 6826, Resigaró 6838, Yucuna 6843
- Macro-Carib: Muinane 6806, Ocaina 6805, Carib (Galibi) 6807, Andoke 6851, Yagua 6839, Japreria 6848, Panare 6849, Akawaio 6853, Hixkaryana 6842, Bakairi 6856
- Macro-Panoan: Abipon 6815, Axluxlay (Ashushlay) 6814, Amahuaca 6810, Tacana 6812
- Macro-Ge: Apinaye 6809, Iate (Fulnio) 6870, Bororo 6850, Maxakali 6858, Kaingang 6808

Eskimo-Aleut (69xx)

- Aleut: Aleut 6900
- Eskimo: Inuit (Greenlandic) 6901, Yupik (Chaplino) 6902

8. AUSTRALASIA

Australian (83xx)

- Yiuidjan: Maung 8350
- Gunwinyuan: Burarra (Burerá) 8352
- Maran: Alawa 8354
- W. Barkly: Garawa 8345
- Garawan: Malakmalak 8356
- Daly: Djeragan: Bunaban: Nyulnyulan: Bardi (Nyulnyul) 8357
- Wororan: (U)Ngarinjin 8374
- Pama-Nyungan: Wik-Munkan 8358, W. Desert 8360, Arrente (Aranda) 8362, Gugu-Yalandyi 8364, Kala Lagaw Ya (Mabuiag) 8365, Diyari 8367, Bandjalang 8368, Yolngu (Dhangu) 8375, Yidiny 8370, Dyirbal 8371, Kalkatungu 8369, Ngiyambaa 8372, Mbabaram 8373
- Ungrouped: Nunggubuyu 8353, Tiwi 8351, Waray 8348, Murinpatha 8349, Yanyuwa (Anyula, Yanyula) 8347

Papuan (86xx)

- Andamanese: Andamanese 8600
- Trans-New Guinea: Selepet 8607, Wantoat 8615, Gadsup 8608, Yagaria (Kamano, Kami) 8609, Chuave 8611, Wahgi 8628, Kewa 8610, Fasu 8617, Asmat 8601, Sentani 8603, Dani 8613, Suena 8618, Kunimaipa 8620, Koiari 8622, Yareba 8621, Dadibi (Daribi) 8616, Taoripi 8623, Dera 8619, Nimbora 8604, Angaatiha 8627, Ekari (Ekagi) 8650, Usan 8630, Southern Kiwai 8636, Woisika 8633, Amele 8635, Waris 8637
- West Papuan: W. Makian 8647
- E. Vogelkop: Geelvink Bay: Yawa (Yava) 8629
- Skó: Vanimo 8640
- Kwomtari: Torricelli: Sepik-Ramu: Iwam 8605, Kwoma (Washkuk) 8602, Yessan-Mayo 8632, Alambalak 8634
- East Papuan: Rotokas 8625, Nasioi 8624, Nambakaengo (Santa Cruz) 8626, Savosavo 8645, Baining 8631

List of character codes with phonetic definitions and IPA transcriptions

Plosives

kpW	n = 1	labialized voiceless labialized labial-velar plosive	[kp ^w]
gbW	n = 1	labialized voiced labialized labial-velar plosive	[gb ^w]
Nmgb	n = 6	prenasalized voiced labial-velar plosive	[nmgb]
kph	n = 1	voiceless aspirated labial-velar plosive	[kp ^h]
kp	n = 35	voiceless labial-velar plosive	[kp]
gb	n = 39	voiced labial-velar plosive	[gb]
pW-h	n = 1	labialized velarized voiceless aspirated bilabial plosive	[p ^{wh}]
bW-	n = 1	labialized velarized voiced bilabial plosive	[b ^w]
mbW	n = 2	prenasalized labialized voiced bilabial plosive	[mb ^w]
pW	n = 4	labialized voiceless bilabial plosive	[p ^w]
bW	n = 1	labialized voiced bilabial plosive	[b ^w]
mbJ	n = 1	prenasalized palatalized voiced bilabial plosive	[mbj]
pJh	n = 3	palatalized voiceless aspirated bilabial plosive	[pj ^h]
pJ	n = 13	palatalized voiceless bilabial plosive	[pj]
bJh	n = 1	palatalized breathy voiced bilabial plosive	[bj]
bJ	n = 8	palatalized voiced bilabial plosive	[bj]
mph	n = 2	prenasalized voiceless aspirated bilabial plosive	[mp ^h]
mp	n = 6	prenasalized voiceless bilabial plosive	[mp]
mb	n = 48	prenasalized voiced bilabial plosive	[mb]
bm	n = 1	nasally-released voiced bilabial plosive	[bm]
ph	n = 101	voiceless aspirated bilabial plosive	[ph]
p*	n = 5	laryngealized voiceless bilabial plosive	[p]
p:	n = 3	long voiceless bilabial plosive	[p:]
phh	n = 2	voiceless bilabial plosive with breathy release	[p ^h]
hp	n = 1	voiceless preaspirated bilabial plosive	[h ^p]
p	n = 375	voiceless bilabial plosive	[p]
b*	n = 15	laryngealized voiced bilabial plosive	[b]
b:	n = 1	long voiced bilabial plosive	[b:]
bh	n = 9	breathy voiced bilabial plosive	[b]
b	n = 287	voiced bilabial plosive	[b]
bD	n = 1	voiced labiodental plosive	[ɸ]
tDJ	n = 2	palatalized voiceless dental plosive	[tj]
dDJ	n = 2	palatalized voiced dental plosive	[dj]
tD9	n = 2	pharyngealized voiceless dental plosive	[t ^ɣ]
dD9	n = 2	pharyngealized voiced dental plosive	[d ^ɣ]
ntDh	n = 2	prenasalized voiceless aspirated dental plosive	[nt ^h]
ntD	n = 3	prenasalized voiceless dental plosive	[nt]
ndD	n = 6	prenasalized voiced dental plosive	[nd]
dDn	n = 1	nasally released voiced dental plosive	[d ⁿ]
tDh	n = 30	voiceless aspirated dental plosive	[t ^h]
tD*	n = 1	laryngealized voiceless dental plosive	[t]
tDhh	n = 1	voiceless dental plosive with breathy release	[t ^h]
tD	n = 106	voiceless dental plosive	[t]
dD*	n = 1	laryngealized voiced dental plosive	[d]
dDh	n = 5	breathy voiced dental plosive	[d]
dD	n = 80	voiced dental plosive	[d]
"ndW	n = 1	prenasalized labialized voiced dental/alveolar plosive	["nd ^w "]
p"tWh	n = 1	labialized voiceless aspirated dental/alveolar plosive	["t ^{wh} "]
"tW	n = 2	labialized voiceless dental/alveolar plosive	["t ^w "]

"dW	n = 1	labialized voiced dental/alveolar plosive	["dʷ"]
l"ndJ	n = 1	prenasalized palatalized voiced dental/alveolar plosive	["ndʝ"]
"tʃh	n = 1	palatalized voiceless aspirated dental/alveolar plosive	["tʃʰ"]
"tʃ	n = 11	palatalized voiceless dental/alveolar plosive	["tʃ"]
"dʃ	n = 5	palatalized voiced dental/alveolar plosive	["dʃ"]
"t-h	n = 1	velarized voiceless aspirated dental/alveolar plosive	["t̠ʰ"]
"tʁ	n = 1	pharyngealized voiceless dental/alveolar plosive	["tʁ"]
"dʁ	n = 1	pharyngealized voiced dental/alveolar plosive	["dʁ"]
"nt	n = 2	prenasalized voiceless dental/alveolar plosive	["nt̚"]
"nd	n = 17	prenasalized voiced dental/alveolar plosive	["nd̚"]
"th	n = 49	voiceless aspirated dental/alveolar plosive	["tʰ"]
"t*	n = 2	laryngealized voiceless dental/alveolar plosive	["t̟"]
"t:	n = 2	long voiceless dental/alveolar plosive	["tː"]
"thh	n = 1	voiceless dental/alveolar plosive with breathy release	["tʰ̤"]
"t	n = 152	voiceless dental/alveolar plosive	["t"]
"d*	n = 3	laryngealized voiced dental/alveolar plosive	["d̟"]
"dh	n = 4	breathy voiced dental/alveolar plosive	["d̤"]
"d	n = 91	voiced dental/alveolar plosive	["d"]
ndJ	n = 1	prenasalized palatalized voiced alveolar plosive	[ndʝ]
tʃ	n = 3	palatalized voiceless alveolar plosive	[tʃ]
dʃ	n = 2	palatalized voiced alveolar plosive	[dʃ]
t-	n = 1	velarized voiceless alveolar plosive	[t̠]
d-	n = 1	velarized voiced alveolar plosive	[d̠]
nt	n = 3	prenasalized voiceless alveolar plosive	[nt̚]
nd	n = 26	prenasalized voiced alveolar plosive	[nd̚]
dn	n = 1	nasally-released voiced alveolar plosive	[dn̚]
th	n = 34	voiceless aspirated alveolar plosive	[tʰ]
t:	n = 1	long voiceless alveolar plosive	[tː]
thh	n = 1	voiceless alveolar plosive with breathy release	[tʰ̤]
ht	n = 1	voiceless preaspirated alveolar plosive	[h̟t]
t	n = 181	voiceless alveolar plosive	[t]
d*	n = 7	laryngealized voiced alveolar plosive	[d̟]
d	n = 120	voiced alveolar plosive	[d]
nt_	n = 1	prenasalized voiceless palato-alveolar plosive	[nt̚]
nd_	n = 4	prenasalized voiced palato-alveolar plosive	[nd̚]
d_n	n = 1	nasally-released voiced palato-alveolar plosive	[dn̚]
t_h	n = 1	voiceless aspirated palato-alveolar plosive	[tʰ]
t_	n = 20	voiceless palato-alveolar plosive	[t̟]
d_*	n = 1	laryngealized voiced palato-alveolar plosive	[d̟]
d_	n = 6	voiced palato-alveolar plosive	[d̟]
nt.	n = 2	prenasalized voiceless retroflex plosive	[nt̚]
nd.	n = 1	prenasalized voiced retroflex plosive	[nd̚]
d.n	n = 1	nasally-released voiced retroflex plosive	[dn̚]
t.h	n = 10	voiceless aspirated retroflex plosive	[tʰ]
t.*	n = 1	laryngealized voiceless retroflex plosive	[t̟]
t.	n = 34	voiceless retroflex plosive	[t̟]
d.*	n = 1	laryngealized voiced retroflex plosive	[d̟]
d.h	n = 6	breathy voiced retroflex plosive	[d̤]
d.	n = 27	voiced retroflex plosive	[d̟]
nc	n = 1	prenasalized voiceless palatal plosive	[nc̚]
ndj	n = 6	prenasalized voiced palatal plosive	[ndʝ]

ch	n = 17	voiceless aspirated palatal plosive	[c ^h]
c:	n = 1	long voiceless palatal plosive	[c:]
c	n = 54	voiceless palatal plosive	[c]
dj*	n = 1	laryngealized voiced palatal plosive	[ʤ]
dj:	n = 1	long voiced palatal plosive	[ʤ:]
dj	n = 43	voiced palatal plosive	[ʤ]
NkW	n = 1	prenasalized labialized voiceless velar plosive	[ŋk ^w]
NgW	n = 10	prenasalized labialized voiced velar plosive	[ŋg ^w]
kWh	n = 21	labialized voiceless aspirated velar plosive	[k ^w h]
kW*	n = 2	laryngealized labialized voiceless velar plosive	[k ^w]
kW:	n = 2	long labialized voiceless velar plosive	[k ^w :]
kW	n = 60	labialized voiceless velar plosive	[k ^w]
gWh	n = 1	labialized breathy voiced velar plosive	[g ^w]
gW	n = 24	labialized voiced velar plosive	[g ^w]
kJh	n = 3	palatalized voiceless aspirated velar plosive	[k ^j h]
kJ	n = 13	palatalized voiceless velar plosive	[k ^j]
gJ	n = 7	palatalized voiced velar plosive	[g ^j]
k9	n = 1	pharyngealized voiceless velar plosive	[k ^ʕ]
Nkh	n = 1	prenasalized voiceless aspirated velar plosive	[ŋk ^h]
Nk	n = 6	prenasalized voiceless velar plosive	[ŋk]
Ng	n = 45	prenasalized voiced velar plosive	[ŋg]
gn	n = 1	nasally-released voiced velar plosive	[g ^ɰ]
gL	n = 1	laterally-released voiced velar plosive	[g ^l]
kh	n = 103	voiceless aspirated velar plosive	[k ^h]
k*	n = 6	laryngealized voiceless velar plosive	[k]
k:	n = 3	long voiceless velar plosive	[k:]
khh	n = 2	voiceless velar plosive with breathy release	[k ^h]
hk	n = 1	voiceless preaspirated velar plosive	[^h k]
k	n = 403	voiceless velar plosive	[k]
gh	n = 10	breathy voiced velar plosive	[g]
g	n = 253	voiced velar plosive	[g]
qW9h	n = 1	labialized pharyngealized voiceless aspirated uvular plosive	[q ^{wʕh}]
GW9	n = 1	labialized pharyngealized voiced uvular plosive	[g ^{wʕ}]
qWh	n = 4	labialized voiceless aspirated uvular plosive	[q ^{wh}]
qW*	n = 1	labialized laryngealized aspirated uvular plosive	[q ^w]
qW:	n = 1	long labialized voiceless uvular plosive	[q ^w :]
qW	n = 12	labialized voiceless uvular plosive	[q ^w]
GW	n = 5	labialized voiced uvular plosive	[g ^w]
q9h	n = 1	pharyngealized voiceless aspirated uvular plosive	[q ^{ʕh}]
G9	n = 1	pharyngealized voiced uvular plosive	[g ^ʕ]
Nqh	n = 1	prenasalized voiceless aspirated uvular plosive	[nq ^h]
Nq	n = 1	prenasalized voiceless uvular plosive	[nq]
qh	n = 17	voiceless aspirated uvular plosive	[q ^h]
q*	n = 1	laryngealized voiceless uvular plosive	[q]
q:	n = 1	long voiceless uvular plosive	[q:]
q	n = 52	voiceless uvular plosive	[q]
G	n = 14	voiced uvular plosive	[g]
99	n = 3	voiced pharyngeal plosive	[ʔ]
?W	n = 1	labialized glottal plosive	[ʔ ^w]
?9	n = 2	pharyngealized glottal plosive	[ʔ ^ʕ]
?	n = 216	glottal plosive	[ʔ]
??	n = 1	voiced glottal plosive	[ʔ]

Implosives, ejective stops and unaffricated clicks

bJ<	n = 1	palatalized voiced bilabial implosive	[β]
p<	n = 2	voiceless bilabial implosive	[p<]
b<	n = 49	voiced bilabial implosive	[β]
dD<	n = 3	voiced dental implosive	[ɖ]
"d<	n = 14	voiced dental/alveolar implosive	["d"]
t<	n = 2	voiceless alveolar implosive	[t<]
d<	n = 23	voiced alveolar implosive	[ɖ]
d_<	n = 1	voiced palato-alveolar implosive	[ɖ]
d.<	n = 1	voiced retroflex implosive	[ɖ]
dj<	n = 10	voiced palatal implosive	[ɟ]
g<	n = 5	voiced velar implosive	[g]
q<	n = 1	voiceless uvular implosive	[q<]
G<	n = 0	voiced uvular implosive	[ɢ]
p'	n = 44	voiceless bilabial ejective stop	[p']
b'	n = 1	voiced bilabial ejective stop	[β']
tD'	n = 16	voiceless dental ejective stop	[t']
"t'	n = 24	voiceless dental/alveolar ejective stop	["t'"]
t'	n = 24	voiceless alveolar ejective stop	[t']
d'	n = 1	voiced alveolar ejective stop	[ɖ']
c'	n = 9	voiceless palatal ejective stop	[ç']
kW'	n = 23	labialized voiceless velar ejective stop	[kʷ']
kJ'	n = 3	palatalized voiceless velar ejective stop	[kʲ']
k'	n = 63	voiceless velar ejective stop	[k']
g'	n = 1	voiced velar ejective stop	[g']
qW9'	n = 1	labialized pharyngealized voiceless uvular ejective stop	[qʷʕ']
qW'	n = 11	labialized voiceless uvular ejective stop	[qʷ']
q9'	n = 1	pharyngealized voiceless uvular ejective stop	[qʕ']
q'	n = 26	voiceless uvular ejective stop	[q']
hn/x?	n = 1	glottalized nasalized velarized voiceless alveolar click	[ŋxʔ]
/xh	n = 1	velar-fricated voiceless aspirated alveolar click	[xʰ]
/x	n = 1	velar-fricated voiceless alveolar click	[x]
g/x?	n = 1	glottalized velar-fricated voiced alveolar click	[g xʔ]
g/x	n = 1	velar-fricated voiced alveolar click	[g x]
hn/h	n = 2	nasalized voiceless aspirated alveolar click	[ŋʰ]
hn/?	n = 2	glottalized nasalized voiceless alveolar click	[ŋʔ]
n/h	n = 1	nasalized breathy voiced alveolar click	[ŋ̤]
n/	n = 2	nasalized voiced alveolar click	[ŋ]
/h	n = 1	voiceless aspirated alveolar click	[hʰ]
/	n = 2	voiceless alveolar click	[ʔ]
g/h	n = 1	breathy voiced alveolar click	[g̤]
g/	n = 1	voiced alveolar click	[g̥]
!xh	n = 1	velar-fricated voiceless aspirated palato-alveolar click	[!xʰ]
hn!h	n = 1	nasalized voiceless aspirated palato-alveolar click	[ŋʰ!]
hn!?	n = 1	glottalized nasalized voiceless palato-alveolar click	[ŋʔ!]
n!	n = 3	nasalized voiced palato-alveolar click	[ŋ!]
!h	n = 2	voiceless aspirated palatal-alveolar click	[!hʰ]
!?	n = 1	glottalized voiceless palatal-alveolar click	[!ʔ]
!	n = 4	voiceless palato-alveolar click	[!]
g!	n = 1	voiced palatal-alveolar click	[g!]

hn/=x?	n = 1	glottalized nasalized velar-fricated voiceless palatal click	[ŋɥxʔ]
/=x	n = 1	velar-fricated voiceless palatal click	[ɥx]
g/=x?	n = 1	glottalized velar-fricated voiced palatal click	[gɥxʔ]
g/=x	n = 1	velar-fricated voiced palatal click	[gɥx]
hn/=h	n = 1	nasalized voiceless aspirated palatal click	[ŋɥ ^h]
hn/=?	n = 1	glottalized nasalized voiceless palatal click	[ŋɥʔ]
n/=h	n = 1	nasalized breathy voiced palatal click	[ŋɥ̤]
n/=	n = 1	nasalized voiced palatal click	[ŋɥ]
/=h	n = 1	voiceless aspirated palatal click	[ɥ ^h]
/=	n = 1	voiceless palatal click	[ɥ]
g/=h	n = 1	breathy voiced palatal click	[gɥ̤]
g/=	n = 1	voiced palatal click	[gɥ]

Fricatives

r[F	n = 1	voiced alveolar fricative flap	[ɾ]
"rF	n = 1	voiced dental/alveolar fricative trill	["r̥"]
iF	n = 1	fricative high front unrounded vowel	[i]
uF	n = 1	fricative high back rounded vowel	[u]
uuF	n = 2	fricative high back unrounded lip-compressed vowel	[ɯ]
hIDFJ	n = 1	palatalized voiceless dental lateral fricative	[ɬ]
IDFJ	n = 1	palatalized voiced dental lateral fricative	[ɮ]
hIDF:	n = 1	long voiceless dental lateral fricative	[ɬ:]
hIDF	n = 1	voiceless dental lateral fricative	[ɬ]
IDF	n = 1	voiced dental lateral fricative	[ɮ]
"hIFW:	n = 1	long labialized voiceless dental/alveolar lateral fricative	["ɬw̥"]
"hIFW	n = 1	labialized voiceless dental/alveolar lateral fricative	["ɬw̥"]
"hIFJ	n = 1	palatalized voiceless dental/alveolar lateral fricative	["ɬj"]
"hIF:	n = 2	long voiceless dental/alveolar lateral fricative	["ɬ:"]
"hIF	n = 24	voiceless dental/alveolar lateral fricative	["ɬ"]
"IF	n = 3	voiced dental/alveolar lateral fricative	["ɮ"]
hxIF	n = 1	voiceless velar-alveolar lateral fricative	[xɬ]
hIFJ	n = 1	palatalized voiceless alveolar lateral fricative	[ɬj]
IFJ	n = 1	palatalized voiced alveolar lateral fricative	[ɮj]
hIF	n = 22	voiceless alveolar lateral fricative	[ɬ]
IF	n = 8	voiced alveolar lateral fricative	[ɮ]
l.F	n = 1	voiced retroflex lateral fricative	[ɮ̠]
hLF	n = 1	voiceless velar lateral fricative	[ɬ]
sDJ	n = 2	palatalized voiceless dental sibilant fricative	[ɬj]
zDJ	n = 2	palatalized voiced sibilant dental fricative	[ɮj]
sD9	n = 1	pharyngealized voiceless dental sibilant fricative	[ɬˤ]
zD9	n = 2	pharyngealized voiced dental sibilant fricative	[ɮˤ]
nzD	n = 1	prenasalized voiced dental sibilant fricative	[nɮ]
sD*	n = 1	laryngealized voiceless dental sibilant fricative	[ɬ̥]
sD	n = 42	voiceless dental sibilant fricative	[ɬ]
zD	n = 17	voiced dental sibilant fricative	[ɮ]
"sW:	n = 1	long labialized voiceless dental/alveolar sibilant fricative	["s̥w̥"]
"sW	n = 3	labialized voiceless dental/alveolar sibilant fricative	["s̥w̥"]
"zW	n = 1	labialized voiced dental/alveolar sibilant fricative	["zw̥"]
"sJ	n = 7	palatalized voiceless dental/alveolar sibilant fricative	["sj̥"]
"zJ	n = 1	palatalized voiced dental/alveolar sibilant fricative	["zj̥"]
"s9	n = 1	pharyngealized voiceless dental/alveolar sibilant fricative	["s̥ˤ"]
"z9	n = 1	pharyngealized voiced dental/alveolar sibilant fricative	["z̥ˤ"]

"ns	n = 1	prenasalized voiceless dental/alveolar sibilant fricative	["ns"]
"nz	n = 1	prenasalized voiced dental/alveolar sibilant fricative	["nz"]
"sh	n = 3	voiceless aspirated dental/alveolar sibilant fricative	["sh"]
"s:	n = 4	long voiceless dental/alveolar sibilant fricative	["s:"]
"hs	n = 1	voiceless preaspirated dental/alveolar sibilant fricative	["hs"]
"s	n = 135	voiceless dental/alveolar sibilant fricative	["s"]
"z	n = 43	voiced dental/alveolar sibilant fricative	["z"]
sJ	n = 2	palatalized voiceless alveolar sibilant fricative	[sʲ]
s9	n = 1	pharyngealized voiceless alveolar sibilant fricative	[sˤ]
nz	n = 1	prenasalized voiced alveolar sibilant fricative	[nz]
s*	n = 3	laryngealized voiceless alveolar sibilant fricative	[s̥]
s:	n = 1	long voiceless alveolar sibilant fricative	[s:]
s	n = 196	voiceless alveolar sibilant fricative	[s]
z	n = 62	voiced alveolar sibilant fricative	[z]
SW:	n = 2	long labialized voiceless palato-alveolar sibilant fricative	[ʃʷ]
SW	n = 3	labialized voiceless palato-alveolar sibilant fricative	[ʃʷ]
ZW	n = 1	labialized voiced palato-alveolar sibilant fricative	[ʒʷ]
SJ	n = 3	palatalized voiceless palato-alveolar sibilant fricative	[ʃʲ]
ZJ	n = 1	palatalized voiced palato-alveolar sibilant fricative	[ʒʲ]
S-	n = 1	velarized voiceless palato-alveolar sibilant fricative	[ʃ̠]
Z-	n = 1	velarized voiced palato-alveolar sibilant fricative	[ʒ̠]
nZ	n = 3	prenasalized voiced palato-alveolar sibilant fricative	[nz̥]
S:	n = 4	long voiceless palato-alveolar sibilant fricative	[ʃ:]
hS	n = 1	voiceless preaspirated palato-alveolar sibilant fricative	[ʃʰ]
S	n = 187	voiceless palato-alveolar sibilant fricative	[ʃ]
Z	n = 61	voiced palato-alveolar sibilant fricative	[ʒ]
s.	n = 23	voiceless retroflex sibilant fricative	[ʂ]
z.*	n = 1	laryngealized voiced retroflex sibilant fricative	[ʐ̥]
z.	n = 9	voiced retroflex sibilant fricative	[ʐ]
C,	n = 9	voiceless palatal sibilant fricative	[ç]
z,	n = 7	voiced palatal sibilant fricative	[ʒ̟]
PW-	n = 1	labialized velarized voiceless bilabial fricative	[ϕʷ]
PW	n = 1	labialized voiceless bilabial fricative	[ϕʷ]
PJ	n = 2	palatalized voiceless bilabial fricative	[ϕʲ]
BJ	n = 1	palatalized voiced bilabial fricative	[βʲ]
P	n = 39	voiceless bilabial fricative	[ϕ]
B	n = 54	voiced bilabial fricative	[β]
fW	n = 1	labialized voiceless labiodental fricative	[fʷ]
vW	n = 1	labialized voiced labiodental fricative	[vʷ]
fJ	n = 3	palatalized voiceless labiodental fricative	[fʲ]
vJ	n = 5	palatalized voiced labiodental fricative	[vʲ]
mv	n = 4	prenasalized voiced labiodental fricative	[mv̥]
f:	n = 1	long voiceless labiodental fricative	[f:]
f	n = 180	voiceless labiodental fricative	[f]
vh	n = 2	breathy voiced labiodental fricative	[v̤]
v	n = 95	voiced labiodental fricative	[v]
6DJ	n = 1	palatalized voiced dental fricative	[dʲ]
0D	n = 18	voiceless dental fricative	[θ]
6D	n = 22	voiced dental fricative	[ð]
"6	n = 2	voiced dental/alveolar sibilant fricative	["ð"]
6	n = 1	voiced dental/alveolar fricative	[ð]
0_	n = 1	voiceless palato-alveolar fricative	[θ]
6_	n = 1	voiceless palato-alveolar fricative	[θ]

0.	n = 1	voiceless retroflex fricative	[θ]
6.	n = 4	voiced retroflex fricative	[ð]
CW	n = 1	labialized voiceless palatal fricative	[ç ^w]
C	n = 11	voiceless palatal fricative	[ç]
jF	n = 12	voiced palatal fricative	[j]
xW:	n = 1	long labialized voiceless velar fricative	[x ^w ː]
xW	n = 25	labialized voiceless velar fricative	[x ^w]
gFW	n = 6	labialized voiced velar fricative	[ɣ ^w]
xJ	n = 1	palatalized voiceless velar fricative	[xʲ]
gFJ	n = 1	palatalized voiced velar fricative	[ɣʲ]
x:	n = 4	long voiceless velar fricative	[xː]
x	n = 94	voiceless velar fricative	[x]
gF*	n = 1	laryngealized voiced velar fricative	[ɣ̠]
gF	n = 55	voiced velar fricative	[ɣ̃]
XW9:	n = 1	long labialized pharyngealized voiceless uvular fricative	[χ ^w ː]
XW9	n = 2	labialized pharyngealized voiceless uvular fricative	[χ ^w]
RFW9	n = 2	labialized pharyngealized voiced uvular fricative	[ʁ ^w]
XW:	n = 2	long labialized voiceless uvular fricative	[χ ^w ː]
XW	n = 14	labialized voiceless uvular fricative	[χ ^w]
RFW	n = 4	voiced uvular fricative	[ʁ ^w]
X9:	n = 1	long pharyngealized voiceless uvular fricative	[χ ^h ː]
X9	n = 2	pharyngealized voiceless uvular fricative	[χ ^h]
RF9	n = 2	pharyngealized voiced uvular fricative	[ʁ ^h]
X:	n = 4	long voiceless uvular fricative	[χː]
X	n = 44	voiceless uvular fricative	[χ]
RF	n = 22	voiced uvular fricative	[ʁ]
H	n = 19	voiceless pharyngeal fricative	[ħ]
9	n = 10	voiced pharyngeal fricative	[ʕ]
hIDFJ'	n = 1	palatalized voiceless dental lateral ejective fricative	[tʃ']
"hIF'	n = 2	voiceless dental/alveolar lateral ejective fricative	["tʃ'"]
sD'	n = 1	voiceless dental sibilant ejective fricative	[s']
"s'	n = 7	voiceless dental/alveolar sibilant ejective fricative	["s'"]
s'	n = 1	voiceless alveolar sibilant ejective fricative	[s']
S'	n = 4	voiceless palato-alveolar sibilant ejective fricative	[ʃ']
s.'	n = 1	voiceless retroflex sibilant ejective fricative	[ʂ']
C.'	n = 1	voiceless palatal sibilant ejective fricative	[ç']
P'	n = 1	voiceless bilabial ejective fricative	[p']
f'	n = 1	voiceless labio-dental ejective fricative	[f']
xW'	n = 1	labialized voiceless velar ejective fricative	[x ^w ']
x'	n = 2	voiceless velar ejective fricative	[x']
XW'	n = 1	labialized voiceless uvular ejective fricative	[χ ^w ']
X'	n = 1	voiceless uvular ejective fricative	[χ']

Affricates, ejective affricates and affricated clicks

t.r	n = 1	voiceless retroflex affricated trill	[tr]
d.r	n = 1	voiced retroflex affricated trill	[dr]
"tIFWh	n = 1	labialized voiceless aspirated dental/alveolar lateral affricate	["tʃ ^{wh} "]
"tIFh	n = 3	voiceless aspirated dental/alveolar lateral affricate	["tʃ ^h "]
"tIF*	n = 1	laryngealized voiceless dental/alveolar lateral affricate	["tʃ̠"]
"tIF:	n = 1	long voiceless dental/alveolar lateral affricate	["tʃː"]
"tIF	n = 6	voiceless dental/alveolar lateral affricate	["tʃ"]
"dIF	n = 5	voiced dental/alveolar lateral affricate	["dʒ"]
klF	n = 1	voiceless velar plosive with alveolar lateral fricative release	[kɬ]

tʃʰ	n = 3	voiceless aspirated alveolar lateral affricate	[tʃʰ]
tʃ	n = 5	voiceless alveolar lateral affricate	[tʃ]
dʃ	n = 3	voiced alveolar lateral affricate	[dʃ]
tʃs	n = 1	voiceless palatalized dental sibilant affricate	[tʃ̟]
ntʃʰ	n = 1	prenasalized voiceless aspirated dental sibilant affricate	[n̩tʃʰ]
ntʃs	n = 1	prenasalized voiceless dental sibilant affricate	[n̩tʃs]
ndʃ	n = 2	prenasalized voiced dental sibilant affricate	[n̩dʃ]
tʃʰ	n = 8	voiceless aspirated dental sibilant affricate	[tʃʰ]
tʃs	n = 16	voiceless dental sibilant affricate	[tʃs]
dʃ	n = 10	voiced dental sibilant affricate	[dʃ]
"tsʰ	n = 3	labialized voiceless aspirated dental/alveolar sibilant affricate	["tsʰʷ]
"tsʰ:	n = 1	long labialized voiceless dental/alveolar sibilant affricate	["tsʰʷ:]
"tsʃ	n = 3	palatalized voiceless dental/alveolar sibilant affricate	["tsʃ]
"dzʃ	n = 1	palatalized voiced dental/alveolar sibilant affricate	["dzʃ]
"nts	n = 1	prenasalized voiceless dental/alveolar sibilant affricate	["nts]
"tʃʰ	n = 25	voiceless aspirated dental/alveolar sibilant affricate	["tʃʰ]
"ts*	n = 1	laryngealized voiceless dental/alveolar sibilant affricate	["ts̟]
"ts:	n = 3	long voiceless dental/alveolar sibilant affricate	["ts:]
"tʃʰh	n = 1	voiceless dental/alveolar sibilant affricate with breathy release	["tʃʰʰ]
"ts	n = 45	voiceless dental/alveolar sibilant affricate	["ts]
"dzʰ	n = 1	breathy voiced dental/alveolar sibilant affricate	["dzʰ]
"dz	n = 20	voiced dental/alveolar sibilant affricate	["dz]
ts-	n = 1	velarized voiceless alveolar sibilant affricate	[ts̟]
dz-	n = 1	velarized voiced alveolar sibilant affricate	[dz̟]
ndz	n = 5	prenasalized voiced alveolar sibilant affricate	[n̩dz̟]
tʃʰ	n = 17	voiceless aspirated alveolar sibilant affricate	[tʃʰ]
ts*	n = 1	laryngealized voiceless alveolar sibilant affricate	[ts̟]
tʃʰh	n = 1	voiceless aspirated alveolar sibilant affricate with breathy release	[tʃʰʰ]
ts	n = 62	voiceless alveolar sibilant affricate	[ts]
dzʰ	n = 2	breathy voiced alveolar sibilant affricate	[dzʰ]
dz	n = 24	voiced alveolar sibilant affricate	[dz]
tʃʰʷ	n = 3	labialized voiceless aspirated palato-alveolar sibilant affricate	[tʃʰʷ]
tʃʰʷ:	n = 1	labialized voiceless palato-alveolar sibilant affricate long	[tʃʰʷ:]
tʃʰʷ	n = 1	labialized voiceless palato-alveolar sibilant affricate	[tʃʰʷ]
dʃʰʷ	n = 2	labialized voiced palato-alveolar sibilant affricate	[dʃʰʷ]
tʃʰʰ	n = 2	palatalized voiceless aspirated palato-alveolar sibilant affricate	[tʃʰʰ]
tʃʰʃ	n = 1	palatalized voiceless palato-alveolar sibilant affricate	[tʃʰʃ]
dʃʰʃ	n = 1	palatalized voiced palato-alveolar sibilant affricate	[dʃʰʃ]
tʃ-	n = 1	velarized voiceless palato-alveolar sibilant affricate	[tʃ̟]
dʃ-	n = 1	velarized voiced palato-alveolar sibilant affricate	[dʃ̟]
ntʃʰ	n = 1	prenasalized voiceless aspirated palato-alveolar sibilant affricate	[n̩tʃʰ]
ntʃ	n = 3	prenasalized voiceless palato-alveolar sibilant affricate	[n̩tʃ]
ndʃ	n = 10	prenasalized voiced palato-alveolar sibilant affricate	[n̩dʃ]
tʃʰ	n = 51	voiceless aspirated palato-alveolar sibilant affricate	[tʃʰ]
ts*	n = 3	laryngealized voiceless palato-alveolar sibilant affricate	[ts̟]
ts:	n = 2	long voiceless palato-alveolar sibilant affricate	[ts̟:]
htʃ	n = 1	voiceless preaspirated palato-alveolar sibilant affricate	[h̩tʃ]
tʃ	n = 188	voiceless palato-alveolar sibilant affricate	[tʃ]
dʃʰ	n = 7	breathy voiced palato-alveolar sibilant affricate	[dʃʰ]
dʃ	n = 113	voiced palato-alveolar sibilant affricate	[dʃ]
nt.ʃʰ	n = 1	prenasalized voiceless aspirated retroflex sibilant affricate	[n̩tʃʰ]
nt.ʃ	n = 1	prenasalized voiceless retroflex sibilant affricate	[n̩tʃ]

nd.z	n = 2	prenasalized voiced retroflex sibilant affricate	[ⁿ dz]
t.sh	n = 9	voiceless aspirated retroflex sibilant affricate	[tʂ ^h]
t.s	n = 16	voiceless retroflex sibilant affricate	[tʂ]
d.z	n = 4	voiced retroflex sibilant affricate	[dʂ]
ncC,	n = 1	prenasalized voiceless palatal sibilant affricate	[ⁿ cç]
ndjz,	n = 1	prenasalized voiced palatal sibilant affricate	[ⁿ ʃʒ]
cC,h	n = 3	voiceless aspirated palatal sibilant affricate	[cç ^h]
cC,	n = 7	voiceless palatal sibilant affricate	[cç]
djz,	n = 2	voiced sibilant palatal affricate	[ʃʒ]
pfh	n = 1	voiceless aspirated labio-dental affricate	[pf ^h]
pf	n = 3	voiceless labio-dental affricate	[pf]
bv	n = 3	voiced labio-dental affricate	[bv]
tD0h	n = 1	voiceless aspirated dental affricate	[t θ ^h]
tD0	n = 2	voiceless dental affricate	[t θ]
dD6	n = 1	voiced dental affricate	[dð]
t0h	n = 1	voiceless aspirated alveolar affricate	[tθ ^h]
t0	n = 2	voiceless alveolar affricate	[tθ]
t.0	n = 1	voiceless retroflex affricate	[tθ]
cCW	n = 1	labialized voiceless palatal affricate	[cç ^w]
djjFW	n = 1	labialized voiced palatal affricate	[ʃʒ ^w]
ndjjF	n = 1	prenasalized voiced palatal affricate	[ⁿ ʃʒ]
cCh	n = 1	voiceless aspirated palatal affricate	[cç ^h]
cC	n = 12	voiceless palatal affricate	[cç]
djjF	n = 8	voiced palatal affricate	[ʃʒ]
kxWh	n = 1	labialized voiceless aspirated velar affricate	[kx ^{wh}]
kxh	n = 2	voiceless aspirated velar affricate	[kx ^h]
kx:	n = 1	long voiceless velar affricate	[kx:]
kx	n = 1	voiceless velar affricate	[kx]
qXW9	n = 1	labialized pharyngealized voiceless uvular affricate	[qχ ^{wɿ}]
qXW	n = 2	labialized voiceless uvular affricate	[qχ ^w]
qX9	n = 1	pharyngealized voiceless uvular affricate	[qχ ^ɿ]
qX:	n = 1	long voiceless uvular affricate	[qχ:]
qX	n = 4	voiceless uvular affricate	[qχ]
"tFW'	n = 1	labialized voiceless dental/alveolar lateral ejective affricate	[ⁿ t ^w ']
"tF':	n = 1	long voiceless dental/alveolar lateral ejective affricate	[ⁿ t [;] ']
"tF'	n = 11	voiceless dental/alveolar lateral ejective affricate	[ⁿ t'']
tF'	n = 9	voiceless alveolar lateral ejective affricate	[t']
klF'	n = 1	voiceless velar lateral ejective affricate	[kL']
tDs'	n = 6	voiceless dental sibilant ejective affricate	[tʂ']
"tsW'	n = 3	labialized voiceless dental/alveolar sibilant ejective affricate	[ⁿ ts ^w '']
"ts':	n = 2	long voiceless dental/alveolar sibilant ejective affricate	[ⁿ ts:']
"ts'	n = 17	voiceless dental/alveolar sibilant ejective affricate	[ⁿ s'']
ts'	n = 26	voiceless alveolar sibilant ejective affricate	[ts']
dz'	n = 1	voiced alveolar sibilant ejective affricate	[dz']
tSW'	n = 3	labialized voiceless palato-alveolar sibilant ejective affricate	[tʃ ^w ']
tS':	n = 2	long voiceless palatal-alveolar sibilant ejective affricate	[tʃ:']
tS'	n = 44	voiceless palato-alveolar sibilant ejective affricate	[tʃ']
dZ'	n = 1	voiced palato-alveolar sibilant ejective affricate	[dʒ']
t.s'	n = 3	voiceless retroflex sibilant ejective affricate	[tʂ']
cC,'	n = 1	voiceless palatal sibilant ejective affricate	[cç']
pf'	n = 1	voiceless labiodental ejective affricate	[pf']
tD0'	n = 2	voiceless dental ejective affricate	[tθ']
kx':	n = 1	voiceless velar ejective affricate long	[kx':]

qXW9'	n = 1	labialized pharyngealized voiceless uvular ejective affricate	[qχ ^{wʕ'}]
qXW'	n = 1	labialized voiceless uvular ejective affricate	[qχ ^{w'}]
qX9':	n = 1	long pharyngealized voiceless uvular ejective affricate	[qχ ^{ʕ'} :]
qX9'	n = 1	pharyngealized voiceless uvular ejective affricate	[qχ ^{ʕ'}]
qX':	n = 2	long voiceless uvular ejective affricate	[qχ':]
qX'	n = 1	voiceless uvular ejective affricate	[qχ']
#xh	n = 1	velar-fricated voiceless aspirated alveolar lateral affricated click	[x ^h]
hn#h	n = 1	nasalized voiceless aspirated alveolar lateral affricated click	[ŋ ^h]
hn#?	n = 2	glottalized nasalized voiceless alveolar lateral affricated click	[ŋ ʔ]
n#	n = 3	nasalized voiced alveolar lateral affricated click	[ŋ]
#h	n = 2	voiceless aspirated alveolar lateral affricated click	[^h]
#?	n = 1	glottalized voiceless alveolar lateral affricated click	[ʔ]
#	n = 4	voiceless alveolar lateral affricated click	[]
g#	n = 1	voiced alveolar lateral affricated click	[g]
hn#jx?	n = 1	glottalized nasalized velar-fricated voiceless palatal lateral affricated click	[ŋ xʔ]
#jx	n = 1	velar-fricated voiceless palatal lateral affricated click	[x]
g#jx?	n = 1	glottalized velar-fricated voiced palatal lateral affricated click	[g xʔ]
g#jx	n = 1	velar-fricated voiced palatal lateral affricated click	[g x]
hn#jh	n = 1	nasalized voiceless aspirated palatal lateral affricated click	[ŋ ^h]
hn#j?	n = 1	glottalized nasalized voiceless palatal lateral affricated click	[ŋ ʔ]
n#jh	n = 1	nasalized breathy voiced palatal lateral affricated click	[ŋ]
n#j	n = 1	nasalized voiced palatal lateral affricated click	[ŋ]
#jh	n = 1	voiceless aspirated palatal lateral affricated click	[^h]
#j	n = 1	voiceless palatal lateral affricated click	[]
g#jh	n = 1	breathy voiced palatal lateral affricated click	[g]
g#j	n = 1	voiced palatal lateral affricated click	[g]
hnlx?	n = 1	glottalized nasalized velar-fricated voiceless dental affricated click	[ŋ xʔ]
lxh	n = 1	velarized voiceless aspirated dental affricated click	[x ^h]
lx	n = 1	velar-fricated voiceless dental affricated click	[x]
glx?	n = 1	glottalized velar-fricated voiced dental affricated click	[g xʔ]
glx	n = 1	velar-fricated voiced dental affricated click	[g x]
hnlh	n = 2	nasalized voiceless aspirated dental affricated click	[ŋ ^h]
hn ?	n = 3	glottalized nasalized voiceless dental affricated click	[ŋ ʔ]
nlh	n = 1	nasalized breathy voiced dental affricated click	[ŋ]
nl	n = 5	nasalized voiced dental affricated click	[ŋ]
lh	n = 3	voiceless aspirated dental affricated click	[^h]
l?	n = 1	glottalized voiceless affricated dental click	[ʔ]
l	n = 5	voiceless dental affricated click	[]
glh	n = 1	breathy voiced dental affricated click	[g]
gl	n = 2	voiced dental affricated click	[g]
/s	n = 1	voiceless alveolar affricated click	[ʃ]

Trills, taps, flaps, and lateral and central approximants

rrD	n = 1	voiced dental r-sound	[rr̥]
"hrr	n = 1	voiceless dental/alveolar r-sound	[ʰrr̥]
"rr*	n = 3	laryngealized voiced dental/alveolar r-sound	[ʰrr̥]
"rr	n = 36	voiced dental/alveolar r-sound	[rr̥]
rrJ	n = 1	palatalized voiced alveolar r-sound	[rr̥j]

rr	n = 10	voiced alveolar r-sound	[r]
rDT*	n = 1	laryngealized voiced dental tap	["D"]
rDT	n = 1	voiced dental tap	[d]
rT	n = 7	voiced alveolar tap	[d]
"l[n = 3	voiced dental/alveolar lateral flap	["l"]
l[J	n = 1	palatalized voiced alveolar lateral flap	[j]
l[n = 9	voiced alveolar lateral flap	[l]
l.[*	n = 1	laryngealized voiced retroflex lateral flap	[ɖ]
l.[n = 8	voiced retroflex lateral flap	[ɖ]
v[n = 1	voiced labio-dental flap	[ʋ]
rD[n = 1	voiced dental flap	[ɖ]
"r[J	n = 1	palatalized voiced dental/alveolar flap	["rj"]
"r[n = 26	voiced dental/alveolar flap	["r"]
hr[J	n = 1	palatalized voiceless alveolar flap	[ç]
r[J	n = 1	palatalized voiced alveolar flap	[ɟ]
hr[-	n = 1	velarized voiceless alveolar flap	[ɕ]
r[-	n = 1	velarized voiced alveolar flap	[ɟ]
r[*	n = 1	glottalized voiced alveolar flap	[ɖ̚]
r[n = 91	voiced alveolar flap	[r]
r_[n = 1	voiced palato-alveolar flap	[ɹ]
r.[~	n = 1	nasalized voiced retroflex flap	[ɖ̃]
r.[n = 14	voiced retroflex flap	[ɖ]
rDJ	n = 1	palatalized voiced dental trill	[ɖj]
rD	n = 9	voiced dental trill	[ɖ]
"rJ	n = 2	palatalized voiced dental/alveolar trill	["rj"]
"r-	n = 1	velarized voiced dental/alveolar trill	["ɕ"]
"r9	n = 1	pharyngealized voiced dental/alveolar trill	["rʕ"]
"hr	n = 1	voiceless dental/alveolar trill	["t̪"]
"r*	n = 1	laryngealized voiced dental/alveolar trill	["ɖ̚"]
"r	n = 51	voiced dental/alveolar trill	["r"]
rJ	n = 2	palatalized voiced alveolar trill	[ɟ]
nr	n = 1	prenasalized voiced alveolar trill	[nɹ]
hr	n = 2	voiceless alveolar trill	[t̪]
r	n = 95	voiced alveolar trill	[r]
r.	n = 1	voiced retroflex trill	[ɖ]
rj	n = 1	voiced palatal trill	[ɟ]
R	n = 4	voiced uvular trill	[ʀ]
lDJ	n = 2	palatalized voiced dental lateral approximant	[ɟ]
lD-	n = 3	velarized voiced dental lateral approximant	[ɕ]
dID	n = 1	voiced prestopped dental lateral approximant	[d̪̚]
hID	n = 1	voiceless dental lateral approximant	[t̪]
lD:	n = 1	long voiced dental lateral approximant	[l:]
lD	n = 34	voiced dental lateral approximant	[l]
"lJ	n = 6	palatalized voiced dental/alveolar lateral approximant	["lj"]
"l-	n = 5	velarized voiced dental/alveolar lateral approximant	["ɕ"]
"l9	n = 1	pharyngealized voiced dental/alveolar lateral approximant	["lʕ"]
"hl	n = 10	voiceless dental/alveolar lateral approximant	["t̪"]
"l*	n = 7	laryngealized voiced dental/alveolar lateral approximant	["l̚"]
"lh	n = 3	breathy voiced dental/alveolar lateral approximant	["l̥"]
"l	n = 136	voiced dental/alveolar lateral approximant	["l"]
hIJ	n = 1	palatalized voiceless alveolar lateral approximant	[ç]
IJ	n = 4	palatalized voiced alveolar lateral approximant	[ɟ]
hl-	n = 1	velarized voiceless alveolar lateral approximant	[ɕ]

l-	n = 2	velarized voiced alveolar lateral approximant	[ɭ]
l9	n = 1	pharyngealized voiced alveolar lateral approximant	[ɭʕ]
l~	n = 1	nasalized voiced alveolar lateral approximant	[ɭ̃]
hl	n = 2	voiceless alveolar lateral approximant	[ɭ̥]
l*	n = 5	laryngealized voiced alveolar lateral approximant	[ɭ̠]
l	n = 174	voiced alveolar lateral approximant	[ɭ]
l_h	n = 1	breathy voiced palato-alveolar lateral approximant	[ɭ̥]
l_	n = 12	voiced palato-alveolar lateral approximant	[ɭ̥]
hl.	n = 1	voiceless retroflex lateral approximant	[ɭ̥]
l.	n = 27	voiced retroflex lateral approximant	[ɭ̥]
lj	n = 20	voiced palatal lateral approximant	[ɭ̟]
L	n = 1	voiced velar lateral approximant	[ɭ̠]
wj	n = 6	voiced labial-palatal approximant	[ɭ̟]
w~	n = 2	nasalized voiced labial-velar approximant	[ɭ̟̃]
hw	n = 15	voiceless labial-velar approximant	[ɭ̟̥]
w*	n = 17	laryngealized voiced labial-velar approximant	[ɭ̟̠]
w	n = 332	voiced labial-velar approximant	[ɭ̟]
BAJ	n = 1	palatalized voiced bilabial approximant	[β̟]
BA-	n = 1	velarized voiced bilabial approximant	[β̠]
PA	n = 1	voiceless bilabial approximant	[β̥]
BA:	n = 1	long voiced bilabial approximant	[β̟:]
BA	n = 19	voiced bilabial approximant	[β]
vA	n = 6	voiced labio-dental approximant	[β̟]
"hrA	n = 1	voiceless dental/alveolar approximant	["ɾ"]
"rA	n = 2	voiced dental/alveolar approximant	["ɾ"]
rA	n = 11	voiced alveolar approximant	[ɾ]
j_	n = 1	voiced palatal-alveolar approximant	[ʒ]
r.A	n = 17	voiced retroflex approximant	[ɻ]
j~	n = 2	nasalized voiced palatal approximant	[ʒ̃]
hj	n = 9	voiceless palatal approximant	[ʒ̥]
j*	n = 17	laryngealized voiced palatal approximant	[ʒ̠]
j	n = 378	voiced palatal approximant	[ʒ]
gA*	n = 2	laryngealized voiced velar approximant	[ɰ̠]
gA	n = 12	voiced velar approximant	[ɰ]
RAW*	n = 1	laryngealized labialized voiced uvular approximant	[ɣ̠]
RAW	n = 1	labialized voiced uvular approximant	[ɣ̟]
RA	n = 2	voiced uvular approximant	[ɣ]

Nasals

hNm	n = 1	voiceless labial-velar nasal	[ŋ̥]
Nm	n = 7	voiced labial-velar nasal	[ŋ]
mW-	n = 1	labialized velarized voiced bilabial nasal	[m̟̠]
mW	n = 5	labialized voiced bilabial nasal	[m̟]
mJ	n = 10	palatalized voiced bilabial nasal	[m̟]
hm	n = 17	voiceless bilabial nasal	[m̥]
m*	n = 14	laryngealized voiced bilabial nasal	[m̠]
m:	n = 3	long voiced bilabial nasal	[m̟:]
mh	n = 3	breathy voiced bilabial nasal	[m̥]
m	n = 425	voiced bilabial nasal	[m]
mD	n = 1	voiced labio-dental nasal	[m̟]
nDJ	n = 1	palatalized voiced dental nasal	[ɲ̟]
hnD	n = 2	voiceless dental nasal	[ɲ̥]

nD*	n = 2	laryngealized voiced dental nasal	[ṅ]
nDh	n = 1	breathy voiced dental nasal	[ṅ̃]
nD	n = 83	voiced dental nasal	[ṅ]
"nW	n = 2	labialized voiced dental/alveolar nasal	["nʷ"]
"nJ	n = 4	palatalized voiced dental/alveolar nasal	["nʲ"]
"hn	n = 9	voiceless dental/alveolar nasal	["n̥"]
"n*	n = 8	laryngealized voiced dental/alveolar nasal	["ṅ"]
"n:	n = 1	long voiced dental/alveolar nasal	["nː"]
"nh	n = 2	breathy voiced dental/alveolar nasal	["ṅ̃"]
"n	n = 160	voiced dental/alveolar nasal	["n"]
hnJ	n = 1	palatalized voiceless alveolar nasal	[ṅ̥ʲ]
nJ	n = 4	palatalized voiced alveolar nasal	[ṅʲ]
hn-	n = 1	velarized voiceless alveolar nasal	[ṅ̠]
n-	n = 1	velarized voiced alveolar nasal	[ṅ̠]
hn	n = 4	voiceless alveolar nasal	[n̥]
n*	n = 4	[laryngealized voiced alveolar nasal	ṅ]
n:	n = 2	long voiced alveolar nasal	[nː]
n	n = 202	voiced alveolar nasal	[n]
hn_	n = 2	voiceless palatal-alveolar nasal	[ṅ̥]
n_ h	n = 1	breathy voiced palato-alveolar nasal	[ṅ̃]
n_	n = 45	voiced palato-alveolar nasal	[ṅ̠]
hn.	n = 1	voiceless retroflex nasal	[ṅ̠]
n.	n = 24	voiced retroflex nasal	[ṅ̠]
njW	n = 1	labialized voiced palatal nasal	[ɲʷ]
hnj	n = 8	voiceless palatal nasal	[ɲ̥]
nj*	n = 3	laryngealized voiced palatal nasal	[ṅ̠]
nj:	n = 1	long voiced palatal nasal	[ɲː]
nj	n = 141	voiced palatal nasal	[ɲ]
hNW	n = 2	labialized voiceless velar nasal	[ŋ̠ʷ]
NW	n = 17	labialized voiced velar nasal	[ŋ̠ʷ]
hNJ	n = 1	palatalized voiceless velar nasal	[ɲ̥ʲ]
NJ	n = 4	palatalized voiced velar nasal	[ɲʲ]
N9	n = 1	pharyngealized voiced velar nasal	[ŋ̠]
hN	n = 9	voiceless velar nasal	[ŋ̠]
N*	n = 3	laryngealized voiced velar nasal	[ṅ̠]
N	n = 237	voiced velar nasal	[ŋ̠]

Vowels

i~	n = 1	[nasalized high front unrounded vowel with velar stricture	[i̠]
i-	n = 1	high front unrounded vowel with velar stricture	[i̠]
"o9~	n = 1	nasalized pharyngealized mid back rounded vowel	["o̠̠"]
O9~:	n = 1	long nasalized pharyngealized lower mid back rounded vowel	[o̠̠ː]
a9~:	n = 1	long nasalized pharyngealized low central unrounded vowel	[a̠̠ː]
a9~	n = 1	nasalized pharyngealized low central unrounded vowel	[a̠̠]
I9	n = 2	pharyngealized lowered high front unrounded vowel	[i̠]
i9:	n = 1	long pharyngealized high front unrounded vowel	[i̠ː]
i9	n = 2	pharyngealized high front unrounded vowel	[i̠]
U9	n = 2	pharyngealized lowered high back rounded vowel	[u̠]
u9:	n = 1	long pharyngealized high back rounded vowel	[u̠ː]
u9	n = 2	pharyngealized high back rounded vowel	[u̠]
"o/9	n = 1	pharyngealized mid front rounded vowel	["ø̠"]

"e9	n = 4	pharyngealized mid front unrounded vowel	["eʕ"]
"o9:	n = 1	long pharyngealized mid back rounded vowel	["oʕ:"]
"o9	n = 3	pharyngealized mid back rounded vowel	["oʕ"]
O9	n = 2	pharyngealized lower mid back rounded vowel	[ɔʕ]
aa9	n = 1	pharyngealized raised low front unrounded vowel	[æʕ]
49	n = 1	pharyngealized raised low central unrounded vowel	[eʕ]
a9:	n = 2	long pharyngealized low central unrounded vowel	[aʕ:]
a9	n = 3	pharyngealized low central unrounded vowel	[aʕ]
I~	n = 10	nasalized lowered high front unrounded vowel	[ĩ]
y~	n = 1	nasalized high front rounded vowel	[ỹ]
i~*	n = 1	laryngealized nasalized high front unrounded vowel	[i̠]
i~:	n = 8	long nasalized high front unrounded vowel	[i̠:]
i~	n = 82	nasalized high front unrounded vowel	[i̠]
I_~	n = 1	nasalized lowered high central unrounded vowel	[ɸ]
i_~:	n = 1	long nasalized high central unrounded vowel	[ɸ:]
i_~	n = 11	nasalized high central unrounded vowel	[ɸ]
U~	n = 9	nasalized lowered high back rounded vowel	[ũ]
u~:	n = 7	long nasalized high back rounded vowel	[ũ:]
u~	n = 74	nasalized high back rounded vowel	[ũ]
uu~	n = 7	nasalized high back unrounded vowel	[u̠]
e~:	n = 2	long nasalized higher mid front unrounded vowel	[ẽ:]
e~	n = 18	nasalized higher mid front unrounded vowel	[ẽ]
@~	n = 1	nasalized higher mid central rounded vowel	[ə̠]
@~:	n = 1	long nasalized higher mid central unrounded vowel	[ə̠:]
@~	n = 2	nasalized higher mid central unrounded vowel	[ə̠]
o~:	n = 4	long nasalized higher mid back rounded vowel	[õ:]
o~	n = 27	nasalized higher mid back rounded vowel	[õ]
"e~*	n = 1	laryngealized nasalized mid front unrounded vowel	["ẽ̠"]
"e~	n = 33	nasalized mid front unrounded vowel	["ẽ̠"]
"@~	n = 12	nasalized mid front unrounded vowel	["ə̠"]
"o(+~	n = 1	nasalized fronted mid back unrounded vowel	["ɸ̠"]
"o~*	n = 1	laryngealized nasalized mid back rounded vowel	["õ̠"]
"o~	n = 35	nasalized mid back rounded vowel	["õ̠"]
"o(~	n = 2	nasalized mid back unrounded vowel	["ɸ̠"]
E)~	n = 1	nasalized lower mid front rounded vowel	[æ̠]
E~:	n = 4	long nasalized lower mid front unrounded vowel	[ɛ̠:]
E~	n = 35	nasalized lower mid front unrounded vowel	[ɛ̠]
3~	n = 3	nasalized lower mid central unrounded vowel	["ə̠"]
O~:	n = 2	long nasalized lower mid back rounded vowel	[ɔ̠:]
O~	n = 32	nasalized lower mid back rounded vowel	[ɔ̠]
^~	n = 1	nasalized lower mid back unrounded vowel	[ɔ̠]
aa~	n = 10	nasalized raised low front unrounded vowel	[æ̠]
a+~	n = 5	nasalized low front unrounded vowel	[a̠]
a~*	n = 1	laryngealized nasalized low central unrounded vowel	[ã̠]
a~:	n = 7	long nasalized low central unrounded vowel	[ã̠:]
a~	n = 83	nasalized low central unrounded vowel	[ã̠]
a_~	n = 3	nasalized low back rounded vowel	[ɔ̠]
a_~:	n = 2	long nasalized low back unrounded vowel	[ɑ̠:]
a_~	n = 6	nasalized low back unrounded vowel	[ɑ̠]
Y	n = 4	lowered high front rounded vowel	[ɣ]
IS	n = 1	overshort lowered high front unrounded vowel	[ɪ]
I	n = 74	lowered high front unrounded vowel	[ɪ]
y:	n = 4	long high front rounded vowel	[y:]
y	n = 24	high front rounded vowel	[y]

hi	n = 3	voiceless high front unrounded vowel	[i̥]
i*	n = 4	laryngealized high front unrounded vowel	[i̠]
i:	n = 40	long high front unrounded vowel	[i:]
ih	n = 5	breathy voiced high front unrounded vowel	[i̪]
iS	n = 2	overshort high front unrounded vowel	[i̟]
i	n = 393	high front unrounded vowel	[i]
U+	n = 1	lowered high central rounded vowel	[ɘ]
I_	n = 4	lowered high central unrounded vowel	[ɪ]
u+:	n = 1	long high central rounded vowel	[u:]
u+S	n = 1	overshort high central rounded vowel	[ʊ]
u+	n = 6	high central rounded vowel	[ɤ]
i_.	n = 1	retroflexed high central unrounded vowel	[i̠ʲ]
i_:	n = 3	long high central unrounded vowel	[i:]
i_S	n = 2	overshort high central unrounded vowel	[i̟]
i_	n = 61	high central unrounded vowel	[ɪ]
U:	n = 1	long lowered high back rounded vowel	[u:]
US	n = 1	overshort lowered high back rounded vowel	[ʊ]
U	n = 66	lowered high back rounded vowel	[ɘ]
UUS	n = 1	overshort lowered high back unrounded vowel	[ʊ̟]
UU	n = 5	lowered high back unrounded vowel	[ɯ]
hu	n = 3	voiceless high back rounded vowel	[h̥]
u*	n = 3	laryngealized high back rounded vowel	[u̠]
u:	n = 36	long high back rounded vowel	[u:]
uh	n = 5	breathy voiced high back rounded vowel	[u̪]
uS	n = 2	overshort high back rounded vowel	[ʊ]
u	n = 369	high back rounded vowel	[u]
uu:	n = 1	long high back unrounded vowel	[u:]
uuh	n = 3	breathy voiced high back unrounded vowel	[u̪]
uu	n = 41	high back unrounded vowel	[ɯ]
o/_	n = 1	higher mid retracted front rounded vowel	[ɔ̠]
e_	n = 2	higher mid retracted front unrounded vowel	[ɛ]
o/:	n = 2	long higher mid front rounded vowel	[o:]
o/S	n = 1	overshort higher mid front rounded vowel	[ɔ̟]
o/	n = 12	higher mid front rounded vowel	[ɔ]
e*	n = 2	laryngealized higher mid front unrounded vowel	[ɛ̠]
e:	n = 21	long higher mid front unrounded vowel	[e:]
eh	n = 5	breathy voiced higher mid front unrounded vowel	[ɛ̪]
e	n = 124	higher mid front unrounded vowel	[ɛ]
@)	n = 4	higher mid central rounded vowel	[ɞ]
@:	n = 2	long higher mid central unrounded vowel	[ə:]
@	n = 20	higher mid central unrounded vowel	[ə]
o+	n = 1	fronted higher mid back rounded vowel	[ɔ̟]
o(+	n = 2	fronted higher mid back unrounded vowel	[ɤ]
o*	n = 2	laryngealized higher mid back rounded vowel	[o̠]
o:	n = 24	long higher mid back rounded vowel	[o:]
oh	n = 5	breathy voiced higher mid back rounded vowel	[ɔ̪]
oS	n = 1	higher mid back rounded vowel	[ɔ̟]
o	n = 131	higher mid back rounded vowel	[ɔ]
o(h	n = 3	breathy voiced higher mid back unrounded vowel	[ɤ̪]
o(n = 12	higher mid back unrounded vowel	[ɤ]
"e_	n = 1	retracted mid front unrounded vowel	["ɛ"]
"o/:	n = 2	long mid front rounded vowel	["o:]
"o/	n = 1	mid front rounded vowel	["ɔ"]

"he	n = 1	voiceless mid front unrounded vowel	[^h e]
"e*	n = 1	laryngealized mid front unrounded vowel	[^h e̥]
"e:	n = 11	long mid front unrounded vowel	[e:]
"eS	n = 1	overshort mid front unrounded vowel	[e̚]
"e	n = 169	mid front unrounded vowel	[e]
"@)S	n = 2	overshort mid central rounded vowel	[^h ø̚]
"@)	n = 5	mid central rounded vowel	[^h ø]
"@.	n = 3	retroflexed mid central unrounded vowel	[^h ɐ̠]
"@:	n = 2	long mid central unrounded vowel	[^h ə:]
"@S	n = 4	overshort mid central unrounded vowel	[^h ɐ̚]
"@	n = 76	mid central unrounded vowel	[^h ə]
"o+	n = 1	fronted mid back rounded vowel	[^h o̟]
"ho	n = 1	voiceless mid back rounded vowel	[^h o̥]
"o*	n = 1	laryngealized mid back rounded vowel	[^h o̥]
"o:	n = 12	long mid back rounded vowel	[o:]
"oS	n = 1	overshort mid back rounded vowel	[o̚]
"o	n = 181	mid back rounded vowel	[o]
"o(n = 1	long mid back unrounded vowel	[y:]
"o(n = 8	mid back unrounded vowel	[y]
E)	n = 8	lower mid front rounded vowel	[œ]
E*	n = 3	laryngealized lower mid front unrounded vowel	[ɛ̥]
E:	n = 14	long lower mid front unrounded vowel	[ɛ:]
Eh	n = 4	breathy voiced lower-mid front unrounded vowel	[ɛ̃]
ES	n = 1	overshort lower mid front unrounded vowel	[ɛ̚]
E	n = 186	lower mid front unrounded vowel	[ɛ]
3	n = 1	lower mid central rounded vowel	[ɔ̞]
3:	n = 1	long lower mid central unrounded vowel	[ɜ:]
3	n = 15	lower mid central unrounded vowel	[ɜ]
O*	n = 3	laryngealized lower mid back rounded vowel	[ɔ̥]
O:	n = 8	long lower mid back rounded vowel	[ɔ:]
Oh	n = 4	breathy voiced lower-mid back rounded vowel	[ɔ̃]
OS	n = 2	overshort lower mid back rounded vowel	[ɔ̚]
O	n = 162	lower mid back rounded vowel	[ɔ]
^h	n = 1	breathy voiced lower mid back unrounded vowel	[ʌ̃]
^	n = 10	lower mid back unrounded vowel	[ʌ]
aa:	n = 8	long raised low front unrounded vowel	[æ:]
aa	n = 39	raised low front unrounded vowel	[æ]
a+:	n = 5	long low front unrounded vowel	[a:]
a+	n = 26	low front unrounded vowel	[a]
4)S	n = 1	overshort raised low central rounded vowel	[ɛ̚]
4S	n = 1	overshort raised low central unrounded vowel	[ɛ̚]
4	n = 14	raised low central unrounded vowel	[ɛ]
a.	n = 1	retroflexed low central unrounded vowel	[a̠]
ha	n = 1	voiceless low central unrounded vowel	[ḁ]
a*	n = 4	laryngealized low central unrounded vowel	[ḁ]
a:	n = 34	long low central unrounded vowel	[a:]
ah	n = 5	breathy low central unrounded vowel	[ã]
aS	n = 3	overshort low central unrounded vowel	[a̚]
a	n = 392	low central unrounded vowel	[a]
4)_	n = 1	raised low back rounded vowel	[ɔ̞]
4_	n = 1	raised low back unrounded vowel	[ɔ̞]
a_):	n = 1	long low back rounded vowel	[ɔ:]
a_)S	n = 1	overshort low back rounded vowel	[ɔ̚]
a_)	n = 19	low back rounded vowel	[ɔ]

a_:	n = 8	long low back unrounded vowel	[ɑ:]
a_	n = 25	low back unrounded vowel	[ɑ]
<u>diphthongs</u>			
oi9~	n = 1	nasalized pharyngealized mid back rounded to high front unrounded diphthong	[oi̯ ^ɰ]
ae9~	n = 1	nasalized pharyngealized low central unrounded to mid front unrounded diphthong	[æe̯ ^ɰ]
ao9~	n = 1	nasalized pharyngealized low central unrounded to mid back rounded diphthong	[ãõ ^ɰ]
oa9~	n = 1	nasalized pharyngealized mid back rounded to low central unrounded diphthong	[õã ^ɰ]
oi9	n = 1	pharyngealized mid back rounded to high front unrounded diphthong	[oi̯ ^ɰ]
ae9	n = 1	pharyngealized low central unrounded to mid front unrounded diphthong	[æe̯ ^ɰ]
ao9	n = 1	pharyngealized low central unrounded to mid back rounded diphthong	[ao̯ ^ɰ]
oa9	n = 1	pharyngealized mid back rounded to low central unrounded diphthong	[oa̯ ^ɰ]
eu~	n = 1	nasalized mid front unrounded to high back rounded diphthong	[ẽũ ^ɰ]
oi~	n = 2	nasalized mid back rounded to high front unrounded diphthong	[õi̯]
ie~	n = 1	nasalized high front unrounded to mid front unrounded diphthong	[iẽ]
ei~	n = 2	nasalized mid front unrounded to high front unrounded diphthong	[ēi̯]
ou~	n = 1	nasalized mid back rounded to high back rounded diphthong	[õũ]
Oi~	n = 2	nasalized lower mid back rounded to high front unrounded diphthong	[õi̯]
ai~	n = 1	nasalized low central unrounded to high front unrounded diphthong	[āi̯]
a+i~	n = 1	nasalized low front unrounded to high front unrounded diphthong	[a̯i̯]
i_i~	n = 1	nasalized high central unrounded to high front unrounded diphthong	[i̯i̯]
ui~	n = 2	nasalized high back rounded to high front unrounded diphthong	[ũi̯]
oa~	n = 1	nasalized mid back rounded to low central unrounded diphthong	[õã]
a_O~	n = 1	nasalized low back unrounded to lower mid back rounded diphthong	[ãõ]
EO~	n = 1	nasalized lower mid front unrounded to lower mid back rounded diphthong	[ẽõ]
o(ih	n = 1	breathy voiced higher mid front unrounded to high front unrounded diphthong	[ʏi̯]
o(i	n = 1	higher mid front unrounded to high front unrounded diphthong	[yi̯]
o/y	n = 1	higher mid front rounded to high front rounded diphthong	[øy]
i@h	n = 1	breathy voiced high front unrounded to mid central unrounded diphthong	[i̯ə]
i@	n = 4	high front unrounded to mid central unrounded diphthong	[iə]
@i	n = 2	mid central unrounded to high front unrounded diphthong	[əi̯]

io	n = 1	high front unrounded to mid back rounded diphthong	[io]
eu	n = 3	mid front unrounded to high back rounded diphthong	[eu]
oi	n = 6	mid back rounded to high front unrounded diphthong	[oi]
le	n = 1	lowered high front unrounded to mid front unrounded diphthong	[le]
yo/	n = 1	high front rounded to mid front rounded diphthong	[yø]
ie	n = 6	high front unrounded to mid front unrounded diphthong	[ie]
ei	n = 7	mid front unrounded to high front unrounded diphthong	[ei]
@u	n = 4	mid central unrounded to high back rounded diphthong	[əu]
u@h	n = 1	breathy voiced high back rounded to mid central unrounded diphthong	[uə̥]
u@	n = 5	high back rounded to mid central unrounded diphthong	[uə]
@uu	n = 1	mid central unrounded to high back unrounded diphthong	[əu]
uu@h	n = 1	breathy voiced high back unrounded to mid central unrounded diphthong	[uə̥]
uu@	n = 2	high back unrounded to mid central unrounded diphthong	[uə]
i_@	n = 1	high central unrounded to mid central unrounded diphthong	[iə]
uo	n = 4	high back rounded to mid back rounded diphthong	[uo]
ou	n = 5	mid back rounded to high back rounded diphthong	[ou]
3i	n = 1	lower mid central unrounded to high front unrounded diphthong	[ɜi]
Oih	n = 1	breathy voiced lower mid back rounded to high front unrounded diphthong	[ɔi]
Oi	n = 6	lower mid back rounded to high front unrounded diphthong	[ɔi]
Oy	n = 1	lower mid back rounded to high front rounded diphthong	[ɔy]
i^	n = 1	high front unrounded to lower mid back unrounded diphthong	[iʌ]
Euu	n = 1	lower mid front unrounded to high back unrounded diphthong	[ɛu]
iEh	n = 1	breathy voiced high front unrounded to lower mid front unrounded diphthong	[iɛ̥]
iE	n = 1	high front unrounded to lower mid front unrounded diphthong	[iɛ]
Ei	n = 3	lower mid front unrounded to high front unrounded diphthong	[ɛi]
u^	n = 1	high back rounded to lower mid back unrounded diphthong	[uʌ]
uu^	n = 1	high back unrounded to lower mid back unrounded diphthong	[uʌ]
aau+	n = 1	raised low front unrounded to high central rounded diphthong	[æ̃]
iah	n = 1	breathy voiced high front unrounded to low central unrounded diphthong	[iʌ̥]
ia	n = 5	high front unrounded to low central unrounded diphthong	[iʌ]
aih	n = 1	breathy voiced low central unrounded to high front unrounded diphthong	[ai̥]
ai	n = 19	low central unrounded to high front unrounded diphthong	[ai]
aa	n = 2	raised low front unrounded to high front unrounded diphthong	[æi]
ia+	n = 1	high front unrounded to low front unrounded diphthong	[iʌ]
a+i	n = 1	low front unrounded to high front unrounded diphthong	[ʌi]
auh	n = 1	breathy voiced low central unrounded to high back rounded diphthong	[aʊ̥]
au	n = 18	low central unrounded to high back rounded diphthong	[aʊ]
uah	n = 1	breathy voiced high back rounded to low central unrounded diphthong	[uʌ̥]
ua	n = 3	high back rounded to low central unrounded diphthong	[uʌ]
auuh	n = 1	breathy voiced low central unrounded to high back unrounded diphthong	[aʊ̥]
auu	n = 3	low central unrounded to high back unrounded diphthong	[aʊ]
uua	n = 2	high back unrounded to low central unrounded diphthong	[uʌ]
i_a	n = 1	high central unrounded to low central unrounded diphthong	[iʌ]

ai	n = 2	low central unrounded to high central unrounded diphthong	[ai]
i_i	n = 2	high central unrounded to high front unrounded diphthong	[ii]
Ui	n = 1	lowered high back rounded to high front unrounded diphthong	[ui]
iuh	n = 1	breathy voiced high front unrounded to high back rounded diphthong	[iu]
iu	n = 2	high front unrounded to high back rounded diphthong	[iu]
uuh	n = 1	breathy voiced high back rounded to high front unrounded diphthong	[ui]
ui	n = 8	high back rounded to high front unrounded diphthong	[ui]
uuih	n = 1	breathy voiced high back unrounded to high front unrounded diphthong	[ui]
uui	n = 1	high back unrounded to high front unrounded diphthong	[ui]
Eo	n = 1	lower mid front unrounded to mid back rounded diphthong	[eo]
ea	n = 1	mid front unrounded to low central unrounded diphthong	[ea]
ae	n = 3	low central unrounded to mid front unrounded diphthong	[ae]
ao	n = 4	low central unrounded to mid back rounded diphthong	[ao]
oa	n = 2	mid back rounded to low central unrounded diphthong	[oa]
e@	n = 1	mid front unrounded to mid central unrounded diphthong	[eə]
eo	n = 3	mid front unrounded to mid back rounded diphthong	[eo]
oe	n = 1	mid back rounded to mid front unrounded diphthong	[oe]
aE	n = 1	low central unrounded to lower mid front unrounded diphthong	[æ]

"h" sounds (variable place)

hW	n = 5	labialized voiceless "h"	[h ^w]
hJ	n = 1	palatalized voiceless "h"	[h ^j]
h*	n = 1	laryngealized voiceless "h"	[h ^l]
h	n = 279	voiceless "h"	[h]
hh	n = 16	voiced "h"	[ɦ]