

NASDAT, an Electronic Database for the Storage of the Parameters of Nasal Harmony

W. Leo WETZELS & Rob GOEDEMANS

*Paris III-CNRS Laboratoire de Phonétique Phonologie/VU University Amsterdam
& Leiden University*

1. Introduction: About NASDAT¹

NASDAT is a database designed to identify the parameters involved in Nasal Harmony (NH) systems, as well as the conditions that could explain or favour the existence of NH in a given language. In view of this aim, and bearing in mind the way in which harmony systems have been described in the current phonological literature, a number of phonetic, phonological, and other grammatical categories are listed that are known, expected, or just speculated to be involved in the operation of NH systems. Other entries in the database provide information about the phonological systems in which NH functions, in an attempt to discover possible relations between NH and other properties of the phonological grammar in which it operates.

¹ We wish to thank G. Nick Clements for his valuable advice during the development stage of the NASDAT project and Eric Akkerman for his comments on an earlier draft of this paper.

The large majority of the data that will be stored in NASDAT are taken from the public domain, mostly from published books and articles, the references of which are always listed in the relevant language record of the database together with the name of the linguist who has published the data. We will assume that no copyrights are at issue regarding the free use of such data. Data that are not published elsewhere can only be listed in NASDAT by the person who collected the data through his/her own fieldwork or by someone with that person's permission and by adding a statement that publication is done with the consent of the speech community where the data were collected. Linguists who use data taken from NASDAT in their publications are requested to acknowledge the use of NASDAT and to mention the linguist responsible for the publication and or storage of the data that were used.

NASDAT is a Microsoft Access based database which was adapted to be accessible on the Web. In practice, three different types of databases will be used. One is the input version, an empty database that will be made available to phonologists who wish to publish their NH data through NASDAT. The second type is equally Access-based, but represents an impoverished version of the input database. This version, which we call the output database, contains all the languages encoded up to a given point in time, but lacks the information that is needed to encode the NH parameters in a coherent manner. The output database is available on request by researchers who wish to do more complex searches on the stored material, some of which will be illustrated at the end of this presentation. The third version is an online version, which contains all the data of the output database. It can be consulted through the Internet and will allow researchers to do a number of elementary queries. No data can be stored directly onto this version. Authors that wish to add their data to NASDAT are invited to contact Leo Wetzels at wlm.wetzels@let.vu.nl. An input version of the database will be sent to prospective contributors as an e-mail attachment or via the NASDAT website, which should be sent back to Wetzels after completion. The new data will be uploaded to the output and Web versions of NASDAT and are then accessible and useable for anyone interested in the properties of NH. In this presentation we are mainly concerned with introducing and illustrating the use of the input database.

The copyright of the database remains with:

Eric Akkerman, (VU University) – technical maintenance of Access database and coordination of online database.

G. Nick Clements (CNRS, Laboratoire de Phonétique et Phonologie/ Paris III, Sorbonne Nouvelle) – external consultant.

Rob Goedemans (Leiden University) – technical designer of the Access database.

W. Leo Wetzels (CNRS/Paris III, Sorbonne Nouvelle & VU University, Amsterdam) – external consultant.

2. The functionality of NASDAT

Upon opening NASDAT in Access one finds the following window which represents the main record form. In this and all subsequent windows we use data from the Brazilian Indigenous language Kaingang as an example to illustrate the fields/parameters of the database.

The screenshot shows a Microsoft Access form titled "NasDat". The title bar has tabs for "Language Info", "Syllables", "Contrastive nasality", "Harmony", "Segments", "Examples", "Stress and Remarks", and "Source(s)". The main area contains the following data:

Language:	Kaingang Paraná		
Dialect of:	Kaingang		
People:	Kaingang		
Genetic Affiliation:	jé		
Region:	Paraná, Brazil. Other dialects are spoken in the states of São Paulo, Santa Catarina, and Rio Grande do Sul, Brazil.		
Latitude:	26	Longitude:	52
Number of speakers:	29000		
Quote:			

At the bottom, there is a toolbar with various icons (e.g., back, forward, search, help) and a status bar showing "Record: 1 van 1".

Do not use single quotes (') in the language name (use of single quotes will prompt an error message). All fields on this page are text fields, except latitude and longitude.

The output and online versions of the database will include a *show other dialects* button that will present all the languages in NASDAT for which the information in the field "dialect of" is identical to that in the current record. Hence, it is vital to spell the parent language exactly the same for all dialects. A description of the fields on this screen follows.

Language:

Name of the language – multiple names are permissible if appropriate.

Dialect of:

Sometimes differences between speech communities are small but significant, in which case one speaks of *dialects* of a single language. If the language being described is considered a dialect of a larger group, list the name of the dialect above under 'language' and list the group (parent language) name here. If there is no name for the specific dialect, give the name of the village, town, or region where the dialect is spoken.

People:

Name of the people that speaks the language. It is also possible to enter multiple names if the people are referred to by several names in the literature.

Genetic Affiliation:

Note here whether the language is a linguistic isolate, or provide the information regarding its linguistic affiliation (group (in case of a dialect), family, phylum).

Region:

More detailed geographical information, at least including the name of the region in which the language is spoken, end with the country name. Spelling of country names should conform to English, i.e. Brazil, not *Brasil*; Italy, not *Italia*, etc.

Latitude and Longitude:

Mark latitude and longitude of the (centre of) the region where the language is spoken.

Number of speakers:

Provide here the (estimated) number of speakers, and, if available, the date of the source's count.

Quote:

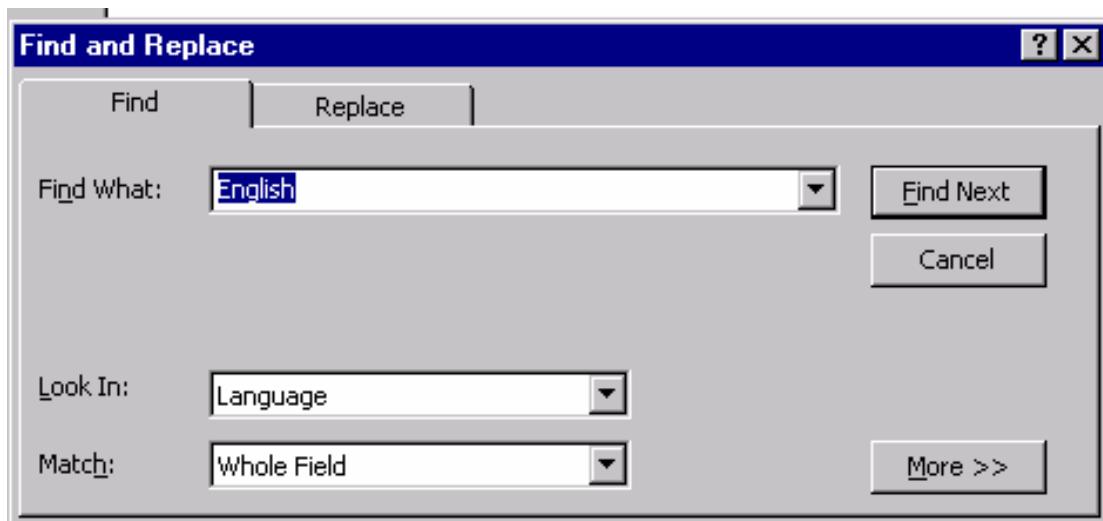
Insert a direct quote from the main primary source describing the Nasal Harmony pattern.

Navigation: there are two sets of navigation buttons. The lower one is Access innate. It is unnecessary to use this button to Navigate through NASDAT.

From left to right, the buttons in the custom made navigation bar are:

1. first record
2. previous record
3. next record
4. last record
5. create record (i.e. add new language)
6. search record
7. undo record
8. delete record
9. help
10. info

From 6 on, things are non-trivial so it is a good idea to begin there. Clicking the search button (illustrated with the binoculars icon) opens the following window:



The search option will only function in the output version of the database because it is irrelevant as a data encoding option. In the output database, if the *exact* name of a language in NASDAT is known, simply type it here in the *Find What:* field and the system will call up that record. Click on *cancel* to leave the search window. If the name of the language is unclear or ambiguous, click the *Match* option, and select *any part of field*. This will bring up records for which any part of the language name matches the input provided in the top field. Thus, typing in *Eng* will direct the encoder to English (a handy option if there is some variation in the spelling of a language). It is advisable to always use this option since one never knows whether extra information has been added to the language name. When it is still not possible to locate a specific language, use the *list* option from among the available forms (which is also only available in the output version of the database).

7. Undo is useful if something was accidentally deleted. It nullifies ALL the changes made to this record since last saved.

8. Deletes the entire record. This option should only be used to delete accidentally-created empty records.

Options 9 and 10 are self-explanatory.

The next windows ask for information about the *syllable* structure of the language.

The screenshot shows the NasDat software interface for the language Paraná Kaingang. The main window title is 'NasDat'. In the top menu bar, the 'Language Info' tab is selected, followed by 'Syllables', 'Harmony', 'Segments', 'Examples', 'Stress and Remarks', and 'Source(s)'. The 'Syllables' tab is active.

Syllable types: V, VC, CV, CVC, CCV, CCVC

Ambisyllabicity:

Restrictions on single onsets: No, all consonants and glides may be onsets

Restrictions on onset clusters: Only non-coronal nasal and oral stops followed by /r/ can be onsets.

Restrictions on single codas: Only nasal consonants, /r/, and glides

Restrictions on coda clusters: No clusters in coda. The word template licenses one extra position at the right edge, which can only be occupied by a single consonant, which must be an independent

Syllable types:

This is basically a summation of the syllable structures possible in the language given in the format CV, CVV, CCV, etc. Common onset and coda restrictions can normally be included in this type of notation (i.e. coda is always glottal stop CV?), but this will complicate searches (a search for languages with codas would include CVC+CV?+CVN, etc.). Separate fields are therefore added for such restrictions, so that it is possible to stick to a very broad description in this field. Collapsed syllable structure notation of the type CV(C) is to be avoided, since this also hampers searches. All syllable types should be separated by commas.

Ambisyllabicity:

This field is designed to designate whether the language has segments that belong to two syllables simultaneously. It may not be easy to determine this at face value. Structural considerations play a role in the decision of the status of, for instance, the /s/ in the Dutch noun *kassa* or the /s/ in the English noun *concert*.

Restrictions on single onsets:

If the language allows CV, CVV, CVC, CVVC (possibly with coda clusters), can the single onset then be any consonant (including glides) in the segment inventory of the language, or are certain (classes of) segments ruled out? A well-known restriction, for instance, is that English words cannot start with a velar nasal /ŋ/, while it is prolific as a word-initial segment in numerous Australian languages.

Restrictions on onset clusters:

Describe which types of onset clusters are allowed. In a CC onset sequence, dependencies will exist between the first and the second consonant. If the first is a stop, many languages will require the second consonant to be a liquid. Another dependency might require the first to be /s/ if the second is a stop. It is impossible to give exhaustive lists here. Shorthand descriptions of the dependencies are preferred: [Cstop]{l, r}, s[Cstop], etc.

Restrictions on single codas:

If the language allows VC, VVC, CVC, CVVC (possibly with onset clusters), can the single coda then be any consonant (including glides) in the segment inventory of the language, or are certain (classes of) segments ruled out? A well-known restriction, for instance, is that Dutch words with long vowels cannot end with a velar nasal /ŋ/ *ooŋ *aaŋ. This example shows that restrictions on single codas (and onsets) need not apply across the board. If restrictions apply to a subset of the possible syllables, they must be listed here.

Restrictions on coda clusters:

Which types of coda clusters are allowed? In a CC coda sequence, dependencies will exist between the first and the second consonant. If the first is a stop, for instance, a language may require the second consonant to be /s/. Another dependency might require the first to be a liquid if the second is a stop. It is impossible to give exhaustive lists here. Shorthand descriptions of the dependencies are preferred: [Cstop]s, {l, r}[Cstop], etc.

The next window is headed *Contrastive Nasality*. Here questions are asked about the contrastive properties of consonants and vowels of the

sound system within which nasal harmony functions. In particular, questions focus on the nature of contrasts present in the set of non-continuant consonants. Together with the parameter setting for the nasal/oral opposition on vowels, these facts could reveal interdependencies between the presence/absence of specific contrasts and the presence/absence of specific types of contour segments in the language and possibly suggest explanations for the occurrence of the latter.

Voiced-voiceless stop contrast: Yes/No

Does the segment inventory of the language show a contrast between voiced and voiceless stops? **Nasal contrast on consonants:** Yes/No

Does the segment inventory of the language show consonants that are only differentiated by their nasality?

Nasal-voiced stop contrast: Yes/No

Does the segment inventory of the language show a contrast between nasal and voiced stops?

Nasal-voiced stop contrast restrictions:

A language may have an opposition between voiced and nasal stops for some places of articulation only, either across the board, or in a given environment. If this is the case, state these restrictions.

Nasal-voiceless stop contrast: Yes/No

Does the segment inventory of the language show a contrast between nasal and voiceless stops?

Nasal-voiceless stop contrast restrictions:

A language may have an opposition between voiceless and nasal stops for some places of articulation only, either across the board, or in a given environment. If this is the case, state these restrictions.

Nasal-oral vowel contrast: Yes/No

Does the segment inventory of the language show vowels that are only differentiated by their nasality?

Nasal-oral vowel contrast restrictions:

Often the system of nasal vowels is more restricted than the system of oral vowels. Does this apply to the language in question? One need not state the restrictions here. Below in the window headed *Examples*, there is a special field for entering the nasal vowel system of the language.

The next window, *Harmony* contains the parameters involved in nasal assimilation and long-distance nasal harmony. Languages may have assimilation, harmony, or both. Some languages have multiple assimilation or nasal harmony processes. In this case, open a new language entry for each different process and number the entries L 1, L 2, L 3, etc., where L represents the name of the language. Under 'quote', in the Language window, stipulate for all the language entries created that L has more than one assimilation and/or harmony process, provide a concise characterization of the different types and state which type is being described in the present entry. Only the entry in L 1 contains the information that is characteristic of the language as a whole.

The upper half of the Harmony window appears below:

The screenshot shows the NasDat software interface for the Paraná Kaingang language. The title bar reads "Paraná Kaingang" and "NasDat". Below the title bar is a menu bar with tabs: Language Info, Syllables, Contrastive nasality, Harmony (which is currently selected), Segments, Examples, Stress and Remarks, and Source(s). The Harmony tab contains two main sections: Nasal assimilation and Nasal harmony.

Nasal assimilation:

- Trigger segment class: nasal vowels
- Trigger segment list: ï, æ, ə, ã, ɔ
- Target segment class: voiceless stops (become prenasalized)
- Target segment list: p, t, k

Prosodic conditions on triggers: [empty text area]

Prosodic conditions on targets: [empty text area]

Nasal harmony:

- Trigger segment class: all vowels
- Trigger segment list: ï, æ, ə, ã, ɔ, i, e, ε, ɿ, ɻ, a, u, o, ɔ
- Target segment class: sonorant segments
- Target segments list: w, j, r, m, n, ɺ, ɳ

Prosodic conditions on triggers: triggers and targets are tautosyllabic

Prosodic conditions on targets: triggers and targets are tautosyllabic

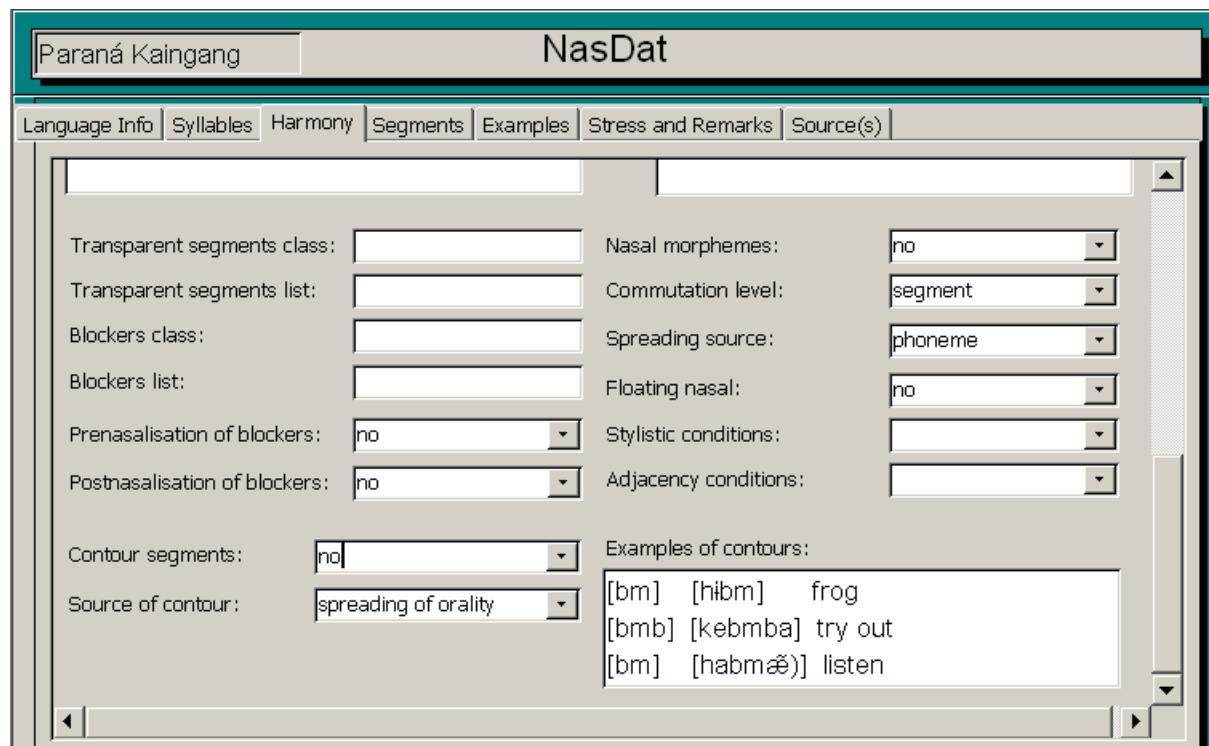
The *Harmony* window has a number of fields with preselection lists. Click the black arrow and select the relevant option. If some lists appear incomplete, please contact the database developers. It is possible to enter values in these fields oneself, but this is not advisable. In later stages we will want to query these fields, and it is much preferred to have a limited set of values for usage in such queries.

As regards the presented analysis of the nasal harmony system of Kaingang, some explanation is required. The Kaingang phonemic system lacks a triple opposition between voiced, voiceless, and nasal consonants, a characteristic that it shares with a considerable number of indigenous languages of the Americas. Instead, it opposes a series of voiceless consonants to a series of consonants that have nasal [m], pre-oralized [^bm], post-oralized [m^b], and circum-oralized [^bm^b] allophones in complementary distribution. It is possible in principle to analyze this consonant system as opposing voiceless (p, t, k) to non-sonorant voiced sounds (b, d, ɺ, g) in which case the target class for tautosyllabic nasal spreading would be 'voiced segments' (including the redundantly voiced glides and liquids) rather than 'sonorant segments'. Here we have chosen for an underlying system in which voiceless stops (p, t, k) contrast with nasal consonants (m,

n, ŋ, ɳ). Since nasal consonants are realized as contour consonants when tautosyllabic with an oral nucleus, the definition of the class of sonorants as the targets for the spreading of a binary [±nasal] feature allows for treating the spreading of nasality as well as the spreading of the orality of a vowel to the tautosyllabic sonorant consonants as a single operation (see for discussion and different analyses of this phenomenon, Anderson 1976, Steriade 1993, Wetzels 2008, 2009).

To introduce IPA symbols into the database the IPA Console can be used. The IPA console, which was developed by Menzo Windhouwer and Rob Goedemans, is available at the website <http://languagelink.let.uu.nl/tds/ipa/index.html>. It allows encoders to select the symbols from the IPA, compose segment classes or examples and copy them into NASDAT. MS Unicode must be available on the computer being used.

The lower half of the *Harmony* window appears below:



Although most of the parameters that appear in the Harmony window are self-explanatory, please find a succinct explanation below.

Nasal harmony: Yes/No

Nasal harmony refers to cases of spreading whereby a feature spreads further than the neighboring segment, when the harmony is obtained from a floating nasal feature, or when the domain of spreading is a prosodic or morphological domain, such as the syllable (as in the case of Kaingang). This excludes cases of one-directional assimilation (one segment to the left or to the right), as in languages where the nasal feature of an onset stop spreads to the preceding vowel or bi-directional assimilation (one segment to the left and to the right), as in languages where the nasal feature of an onset consonant spreads to the preceding and following vowel only.

Nasal assimilation: Yes/No

Nasal assimilation is understood here as the local spreading of a nasal feature, to maximally one segment at one or each side of the triggering segment (but see above).

Trigger segments class:

Note here which segments trigger nasal harmony/assimilation. Provide the set of trigger segments as a class, if possible.

Trigger segments list:

Give here the set of trigger segments as a list.

Conditions on trigger segments:

If there are conditions under which the trigger segments function as such, list them here.

Domain of spreading (both phonological and morphological):

Provide here the largest morphological or prosodic unit within which spreading takes place, for example: root, stem, word, syllable, foot, phonological word, clitic group, etc.

Directionality:

Does nasality spread from the trigger to the left (regressive), to the right (progressive), or does it move in both directions (bidirectional)?

Valence:

Specify unary or binary. Valence is unary if the single value [nasal] is sufficient to account for the data. Valence is binary if [\pm nasal] is required, as is proposed here for Kaingang, and which, for different reasons, can be

argued for Desano and Barasana (for Desano, see the contribution of Wilson Silva to this volume).

Target segments class:

Identify which segments receive nasality as the result of the harmony/assimilation rule. Provide the set of target segments as a class, if possible.

Target segments list:

List the set of target segments.

Conditions on target segments:

If there are conditions under which the target segments become nasal, register them in this field.

Transparent segments class:

Provide here the class of segments that let nasality pass, but that do not become nasal as part of the spreading process.

Transparent segments list:

Provide here the list of segments that let nasality pass, but that do not become nasal as part of the spreading process

Blockers class:

Provide here the class of segments that block the spreading process, if possible.

Blockers list:

Provide here the list of segments that block the spreading process

Prenasalization of blockers: Yes/No

In some languages the blocking segments are prenasalized. If this happens in the language in question, state whether or not this applies to the class (or a subclass) of blocking segments.

Postnasalization of blockers: Yes/No

In some languages the blocking segments are postnasalized. If this happens in the language being entered, state whether or not this applies to the class (or a subclass) of blocking segments.

Contour segments:

Does the language have contour segments that cannot be considered variants of an underlying segment that is not a contour segment? This means that the contour property is itself contrastive in the language. Notice that it is possible to choose 'No' in this entry and still have contour segments in the language, as in Kaingang, where contour segments are derived.

Source of contour segments:

Languages may have contour segments that are derived from plain segments. Here we list three possible sources.

a. Spreading of nasality.

Contours can arise through the contextual influence of nasal segments. In a typical case $\hat{V}^N C$ arise as variants of underlying sequences $\hat{V} C$ through progressive nasal spreading from the nasal vowel to the consonant (/əp/ > [ə^mp]).

b. Spreading of orality.

Contours can arise through the contextual influence of oral segments. In a typical case $V^C N$ arise as variants of underlying sequences $V N$ sequence through progressive spreading of orality from the nasal vowel to the consonant (/əm/ > [ə^bm]).

c. spontaneous contours (phonetic enhancement of non-sonorant voice). These contour segments typically occur as variants of underlying voiced oral stops, especially in the syllable onset (see for example discussion by Ohala 1983, Iverson & Salmons 1996, Keyser & Stevens 2006, Wetzel 2008)

It may sometimes be difficult to decide whether derived contour consonants are caused by the spreading of orality or by phonetic enhancement of non-sonorant voice.

Nasal morphemes: Yes/No

Sometimes the nasal feature represents some morphological category in the language, in which case it must be listed in the lexicon as a morpheme by itself, as, for example, in Terena, where /N/ represents the 1person singular morpheme.

Commutation level:

Establish the level at which nasality is contrastive; specify segment/syllable/morpheme/other.

Spreading source:

Indicate phoneme or nasal morpheme

Floating nasal: Yes/No

In many languages exhibiting nasal harmony, words are either ‘entirely’ nasal or entirely oral. In such cases nasality is usually represented as a floating feature that is linked by the harmonic segments by association.

Stylistic conditions:

Assert whether the rule is optional, if it only applies in fast speech, etc.

Adjacency conditions:

Specify here any conditions on the adjacency of nasal(ized) segments, such as segment adjacency, syllable adjacency, or none.

The *Segments* window is the most complex section of this database in all respects. It links the main sections of NASDAT to a table that contains all the segment information stored in the UCLA Phonological Segment Inventory Database (UPSID, which can be found at <http://www.linguistics.ucla.edu/faciliti/sales/software.htm#upsid>). Note that only the segments with their properties are included. UPSID contains phoneme inventories for hundreds of languages based on selections from the present list of segments. These inventories are not included in NASDAT. On the segment tab, one can select which segments from the UPSID range are present in the phoneme inventory of the language being inputted.

The screenshot shows the NasDat software interface for the language Paraná Kaingang. The top menu bar includes Language Info, Syllables, Contrastive nasality, Harmony, Segments, Examples, Stress and Remarks, and Source(s). The main window is divided into three main sections: Segment Type, Segment, and Segment inventory.

- Segment Type:** A list of segment categories including Plosives, Implosives, Ejective Stops, Clicks, Fricatives, Ejective Fricatives, and Affricates. A sub-list below shows articulation types: bilabial, labiodental, dental, unspecified dental, alveolar, palato alveolar, and retroflex.
- Segment:** A list of segments: p:, p̚, ʰp, and b.
- Segment inventory:** A table titled "segment_IPA" listing IPA symbols: p, t, k, m, n, ɿ, ɳ, ?, ɸ, ʃ, h, r, and i.

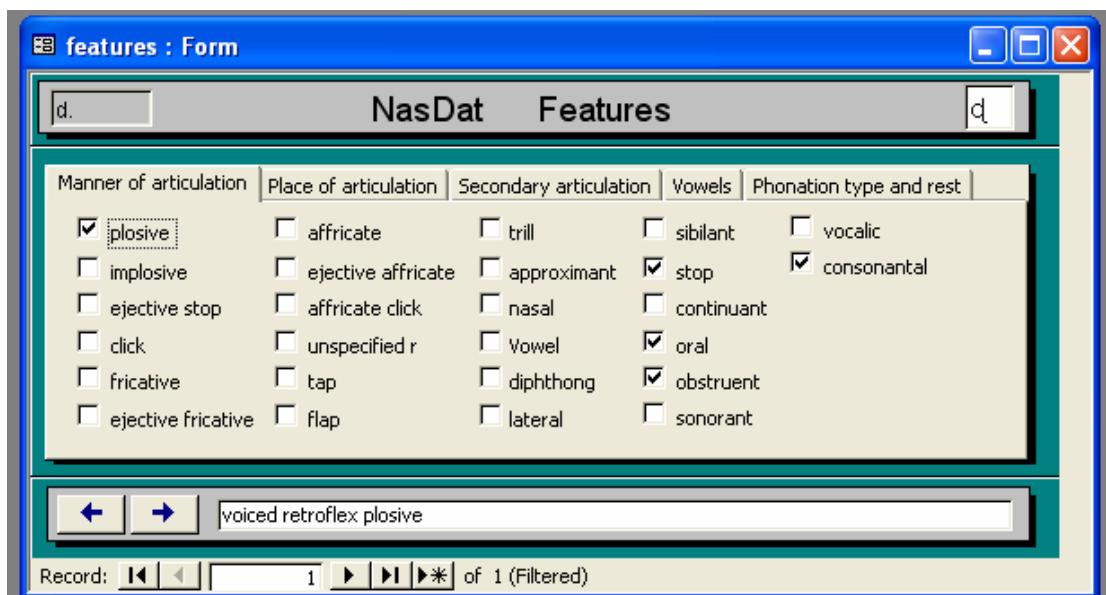
An anomalous segments section is also present.

Because there are 920 different segments represented in UPSID data entry is far from trivial. We have chosen for a list and sub-list presentation to guide encoders through the selection of segments. Upon clicking *plosives*, for instance, a sub-list opens below it:

A zoomed-in view of the Segment Type panel. The "Plosives" category is selected, and its sub-list is displayed: bilabial, labiodental, dental, unspecified dental, alveolar, palato alveolar, and retroflex.

Select the articulation type from this list, after which a third list will open containing segments that are plosives of the articulation type selected. In

this third list click the desired segment, and with the buttons below it choose the relevant to take regarding this segment. "Add Segment" puts it in the phoneme inventory of the language, which is shown in the window to the right of the buttons and lists. "Delete Segment" removes it from that inventory. It will be updated in real time. When the *show features* button is clicked after a segment is select, encoders can see the relevant UPSID info, as below:



Note that only 1 of the segments in the segment table is displayed. Clicking on 'By Category' will yield an overview of all of the segments in the language, thus providing three lists of phonemes for that language: one for consonants, one for vowels and one for glottals and approximants (use the "back" button to go back to the main form). To view all 920 segments from UPSID, choose the *Segments* form in the main form window (fig 1 above). This will again give the UPSID segment form, but now one can navigate through the list of 920 segments.

The small white field on the top right corner of this feature form holds the IPA equivalent of the UPSID code in the left hand corner. These characters are entered in Unicode format. The font used in NASDAT is Arial Unicode MS.

Segments:

Select all segments that occur in the language using these tables.

Anomalous Segments:

In this field the rare segments of the language are listed. Note that these segments are also present in the inventory that can be found in the list on the right hand side of the segments page. This field simply conveys which of the segments in that list are anomalous according to the UPSID definition (rare segments, loans, etc.). For 317 languages the status of segments can be checked in UPSID. For the rest, the decision is up to the encoder.

The next field is a space to insert examples that illustrate the contrastive segments of the language. We have not bothered to do this for all the phonemes of Kaingang. Remember that these examples can be entered easily in phonetic transcription using the IPA Console mentioned previously.

Language Info Syllables Contrastive nasality Harmony Segments Examples Stress and Remarks Source(s)	
Consonant examples:	/po/ stone, /ta/ rain, etc.
Oral vowel examples:	/ti/ he, /ŋi/ very, etc.
Nasal vowel examples:	/t̄i/ fly, /tǣ/ go, etc.

The next window requests encoders to describe the stress system of the language in question. There is also space for special remarks that one might wish to make relating to the language

Paraná Kaingang

NasDat

Language Info | Syllables | Contrastive nasality | Harmony | Segments | Examples | Stress and Remarks | Source(s)

Stress: Word-final, some exceptions in function words, which may have initial stress

Stress source: Wiesemann, 1972:56 and 79-81

Remarks
first record, check carefully!

Stress:

Describe the stress rule(s) here. As such descriptions may often be quite prolix, and thus not conducive to easy searching, stress patterns should be coded. To facilitate easy encoding, we use the scheme used in the StressTyp database (where it has had 15 years to mature) Stress codes for 510 languages can be found there. For stress systems that are not yet represented in StressTyp, new codes need to be created. To do so consult the StressTyp manual on line <http://www.unileiden.net/stresstyp/> or Rob Goedemans (who will be glad to do this provided the stress information of the language). His email is r.w.n.goedemans@hum.leidenuniv.nl.

Stress Source:

Provide the descriptive source for the data in the previous field. Only use primary sources. Second hand descriptions in theoretical works on stress may be inaccurate or biased.

Remarks:

Any remarks pertinent to the complete entry for this language must go here. This field typically contains information that the encoder could not put into one of the regular fields, but is relevant nonetheless, or information regarding alternative analyses of the data. If there is doubt about certain data, that can be expressed here also. Lastly, if the language has nasal morphemes, the morphological function of the nasal morpheme may be stipulated in this area. The field ends with the name of the person who entered the remark, if it is not the author of the record.

The final window is designed to provide information about the *Sources* of the data that entered in the database.

Data collected: by the author/from secondary sources/both

In this field the notion 'secondary source' means that the data have been encoded from published sources that were not personally collected by the person who enters the data.

Primary source(s):

List here the literature that describes the primary linguistic data of the language being entered. These are usually descriptive grammars, dissertations, or papers, but may be theoretical papers that are based upon data gathered by the author(s).

Secondary source(s):

List here the literature that uses the data of this language gathered by others for theoretical, comparative, typological, or other purposes.

3. Queries

Since enabling queries is the main reason for building a database like this in the first place, it is a good idea to review some basic examples by way of exposition. As was made clear above, queries can be done by individual researchers in the Access output database and, although to a lesser degree, in the Web version of NASDAT.

Go to the other Access window on the task bar, or select window 1 under "window" in the menu to see a window in which "queries" is a menu option on the left of the panel. Once that option has been clicked, users will see a list of simple example queries.

1. Consonants: lists all the consonants in NasDat together with the language they are found in. Just as in all the queries outline here, it is possible to right click the query name and select "Design view" to see how this query was created.
2. Feature true for segment: this query lists the segments for which the fields "fricative" and "pharyngeal" are both true, hence which segments have the features fricative and pharyngeal (it selects those segments from the entire UPSID list). Double click this to see a list of those segments. To change the query, go to the design view. In this window users can click the features, which both head a column, so that a drop down list appears. From this list select a different feature. It is also possible to change the value that feature should have from true to false. Close the window when done and check the result by double clicking the query.
3. Languages having features xy: this query is only slightly more complicated than 2. It combines the main table with that query to show which languages have phonemes that have, or do not have, certain features. At this point checks for languages that have no "labial" and "voiceless" segments. Click *design* for this one, and notice the way the tables above the columns are joined by two new ones, and linked to them. It is possible to select fields from these tables (NB, do **not** alter the *link* table ever!) and add them to the columns to build more complicated queries.
4. Main query: Leave this one alone. It drives the *list* form.
5. Search IPA char: Use this to test whether it is possible to search for IPA characters. If a Unicode representation is entered in the pop-up window that appears, the query will look for this segment in the UPSID table.
6. Vowels: lists the vowels from all NASDAT languages.

4. To conclude

NASDAT will soon be available on-line, through the website [ww2.let.vu.nl/nasdat](http://www2.let.vu.nl/nasdat). There will be links to this Web address from several other places, as for example <http://www2.let.vu.nl/amazonianlinguistics/>. Authors who are inspired by this article to start encoding their language for its nasal assimilation and or nasal harmony characteristics are invited to write to Leo Wetzels, in which case they will be sent an Access version of NASDAT, which will be uploaded to the Web version upon return. Feel free to pass the word to colleagues who work with nasal assimilation/harmony languages. It is obvious that the use of NASDAT will be interesting and useful only after a considerable number of languages are encoded.

References

- ANDERSON, Stephen
1976 Nasality and the Internal Structure of Segments. *Language* 52:
326-344.
- IVERSON, Gregory K. & SALMONS, Joseph C.
1996 Mixtec prenasalization as hypervoicing. *International Journal of
American Linguistics* 62.2: 165-75.
- KEYSER, Samuel J. & STEVENS, Kenneth N.
2006 Enhancement and Overlap in the Speech Chain. *Language* 82.1:
33-63.
- OHALA, John
1983 The Origin of Sound Patterns in Vocal Tract Constraints. *The
Production of Speech*, Peter MacNeilage (ed.), pp. 189-21. New
York: Springer-Verlag.

STERIADE, Donca

1993 Closure, Release and Nasal Contours. *Nasal, Nasalization and the Velum. Phonetics and Phonology*, M. Huffman & R. Krakow (eds), pp. 401-470. Academic Press.

WETZELS, W. Leo

2008 Thoughts on the Phonological Definition of Nasal/Oral Contour Consonants in Some Indigenous Languages of South-America. *Revista ALFA* 52(2): *Abordagens em Fonética e Fonologia: estudos auditivos, acústicos e perceptivos; modelos de análise fonológica de ontem e de hoje*. Universidade Estadual Paulista (UNESP).

2008 Word Prosody and the Distribution of Oral/Nasal Contour Consonants in Kaingang. *Linguistics and Archaeology in the Americas. The historization of language and society*, Eithne, B. Carlin & Simon van de Kerke (eds.). Leiden: Brill Publishers