Language Documentation & Conservation
Vol. 11 (2017), pp. 49–80
http://nflrc.hawaii.edu/ldc
http://hdl.handle.net/10125/24724
Revised Version Received: 08 Nov 2016

Earbuds: A Method for Analyzing Nasality in the Field

Jesse Stewart
University of Saskatchewan

Martin Kohlberger
Leiden University & James Cook University

Existing methods for collecting and analyzing nasality data are problematic for linguistic fieldworkers: aerodynamic equipment can be expensive and difficult to transport, and acoustic analyses require large amounts of optimally-recorded data. In this paper, a highly mobile and low-cost method is proposed. By connecting low impedance earbuds into a microphone jack of a recording device and placing one earbud immediately below one nostril while keeping the other earbud by the mouth, it is possible to capture the relative intensity of sound exiting the nasal and oral cavities. The two channels can then be normalized to assess the relative prominence of nasality and orality in a given speech sound. This method can not only be used to establish whether nasality is present in a speech signal, but it can also provide information about the timing and duration of nasal gestures. As such, it is an ideal tool for collecting high-quality nasality data in the field.

1. Introduction[1] Nasality is defined by the coupling of the nasal and oral tract, as controlled by the velum and the posterior and lateral pharyngeal walls. Nasal speech sounds are produced when the velum is lowered and the passageway connecting the oral and nasal cavities is open. Compared to other phonetic parameters (such as those relating to vowel quality and place/manner of articulation in consonants), quantitative measurements of the degree of velopharyngeal opening (and thus degree of nasality) are relatively difficult to obtain. So far, three types of methods have been employed to collect nasality data: articulatory, aerodynamic, and acoustic techniques (see Baken 1987 for a thorough overview). While all these methods have particular strengths, most of them are problematic in the context of (potentially remote) fieldwork, especially with regards to cost and mobility.

Articulatory imaging techniques, such as X-ray, MRI and fleshpoint tracking, provide a clear indication of velopharyngeal opening, but they involve expensive equipment that cannot be transported out of a laboratory setting. Other articulatory pro-

[1]We would like to thank our language consultants (alphabetical order): Natalie Clarke, Marisol Farinango, Fachrizal Halim, Léa Kimenau, Izumi Krasznai, Lucía Gonza, Magdalena Piaguaje, Gabriela Prado, Aaron Simbananiye, Bettina Spreng, Christiani Pinheiro Thompson Wagner, and a Shiwiar woman who would like to remain anonymous. We are grateful to Patricia Keating, Grant McGuire, and Heriberto Avelino (who appears to have already been using a similar earbud method as a pedagogical tool in his classes) for reaching out to us with very helpful comments in the early stages of this work. We are also indebted to two anonymous reviewers for their immensely valuable feedback and suggestions.

E-ISSN 1934-5275

cedures like nasoendoscopies (Benguerel et al. 1975) and photonasographies (Ohala 1971) are more portable, but they are problematic outside of a clinical and sterile environment because of their invasive nature. In contrast, acoustic data can be easily collected anywhere but even well-known acoustic measures of nasality – such as Chen's A1-P1 and A1-P0 (Chen 1996) – require large quantities of optimally-recorded speech and are not suited for cross-vowel comparisons. Recently Styler (2015) has improved on these methods demonstrating that, particularly, F1 bandwidth is a primary correlate in determining nasality perception. While his research provides consistent results, it is still limited to vowel segments and does not reveal gestural movements involved in velum coarticulation in surrounding segments. Aerodynamic methods involving double-chambered masks and transducers are arguably the best approach to studying nasality as they provide detailed and dynamic information about pressure and airflow from both the oral and nasal cavities. However, unless a fieldworker's primary goal is to research nasality, aerodynamic equipment is often too expensive to purchase, too sensitive to carry and, crucially, difficult to calibrate in remote areas. For this reason, field linguists continue to rely mostly on impressionistic perceptual cues of nasality, resulting in incomplete or inadequate analyses in lesser-studied languages.

In this paper, we introduce a new variant of an existing technique based on a measure known as nasalance and present it as a highly portable, low-cost and easily implementable tool for empirically measuring nasality in the field. We propose using a pair of low-impedance earbuds to assess the relative nasality of speech sounds: one earbud is placed directly under or in a nostril to record the relative amplitude of sound emanating from the nasal cavity and the other is placed at the corner of the mouth to capture the relative amplitude of oral sounds. The differences in amplitude and intensity of the two tracks can be compared in a single stereo recording. Inherent differences in loudness produced by the oral and nasal tracts can automatically be eliminated by simultaneously normalizing both tracks based on a predefined decibel value (e.g., -3 dB) using a common effects setting included in most audio editing software programs. Although this method cannot provide any information on airflow, and therefore cannot be considered a replacement for more robust aerodynamic approaches, it can provide detailed information about the timing and duration of nasal gestures (e.g., pre- and postnasalization, nasal harmony, nasal leakage, and coarticulation effects). As such, it can be used by all linguistic fieldworkers, even those with a limited background in acoustic phonetics, to more reliably describe the behavior of nasality in a given language in cases where other analytic tools are not available.

2. Background

2.1 Nasalance Nasalance refers to a measure which compares acoustic energy at the nostrils to acoustic energy at the mouth, thereby giving an indication of the degree of nasality in speech sounds. In the 40 years since the term was coined by Fletcher, Sooudi and Frost (1974), several different variations of this measure have been proposed. In this paper we will deal with the most basic calculation of nasalance, as shown in Equation 1.

$$\frac{A_{nasal}}{A_{nasal} + A_{oral}}\text{(Ratio)}\frac{100 * A_{nasal}}{A_{nasal} + A_{oral}}\text{(Percentage ratio)}$$

Equation 1. This calculation (summarized in Rothenberg, unpublished MS) expresses the ratio or percentage ratio between nasal and oral amplitude. The two calculations are equivalent but the resulting numbers are in different scales.

Nasalance data has most commonly been collected by means of a so-called nasometer. This setup was first proposed by Fletcher, Adams and McCutcheon (1989) and involves a sound-separating plate (also known as an acoustic baffle) being placed between the nose and the mouth. Microphones are then mounted on either side of the baffle so that the sound exiting the nasal cavity can be recorded separately to that exiting the oral cavity. Given the relative simplicity of this technique, it is not surprising that nasalance measures and nasometry have been used extensively in linguistic and clinical research on nasality since they were first introduced. However, even though the method has been available for four decades, the use of nasalance as a tool for understanding nasality is still extremely limited in the context of descriptive and documentary fieldwork.

2.2 Field studies on nasality There has long been an interest within the subdiscipline of phonetics to devote attention to the phonetic structures of underdocumented languages (Ladefoged 1999, 2003). Nevertheless, field studies on nasality have lagged behind other areas of phonetics. Although there have been several publications on the aerodynamic and acoustic properties of nasality in endangered languages in recent years (see Brenner et al. 2011; Demolin 2011; de Paula 2016; Riehl 2008, *inter alia*), the number of these studies is still very limited and restricted to a handful of languages.

In the domain of descriptive and documentary linguistics, the situation is even more inadequate. The empirical basis for claims about nasality in recently published grammars is virtually non-existent. This conclusion is the result of an assessment of 372 grammars listed in the Grammar Watch List of the Association of Linguistic Typology.[2] None of the grammars consulted provided resolvable and quantifiable data to back up claims regarding the existence and behavior of nasality. In all cases, judgments about nasality were impressionistic.

The limited number of studies on nasality in lesser-studied languages and the lack of empirical basis in describing nasality from a descriptive and documentary perspective is a cause for concern. This is especially true given the growing awareness of the importance of data transparency and reproducibility in language documentation

[2]The Grammar Watch List of the Association for Linguistic Typology is the most complete list of recently published grammars available. It includes grammars published worldwide between 1993 and 2007. There are 574 publications on the list that refer to spoken languages, but the authors only had access to 372 of those.

(Berez 2015). In the following sections, we address these shortcomings by introducing a simple, portable and low-cost tool based on the concept of nasalance which will allow fieldworkers to measure and empirically document the presence of nasality in speech.

3. Methods

3.1 Materials The technique presented here builds on the concepts of nasalance and nasometry detailed in §2.1. This method requires the same standard recording setup that a field researcher uses for capturing oral speech. However, instead of a standard microphone, a pair of low impedance earbud headphones (~27 ohms) are used to record sound as it exits the oral and the nasal tracts respectively. An immediate advantage of this technique is that low quality earbuds that can typically be purchased at very low cost work fine for this method. Only earbuds with silicon attachments should be used, as trials with flat surface earbuds show that too much sound is captured from the oral tract (see Figure 1).

Figure 1. Earbuds with silicon tips (left) are preferred over the flat earbuds on the right.

Instead of using the earbuds as an output device, they are used to record sound as it exits the nasal and oral cavities. Because earphones work on the same basic principle as microphones (i.e. a diaphragm vibrates when agitated by a medium and a copper coil around a magnet creates an electromagnetic field which interprets the signal), earbuds function as small microphones when attached to the input jack of a recording device. It should be emphasized that in order for this method to work properly, the digital recording device must be set to record in *stereo* and not *mono*.

The benefit of using earbuds over a standard microphone is their low cost and their high directionality. The silicon buds serve two purposes: (1) by surrounding the filter, they act as a buffer reducing oral feedback and (2) they function similarly to a cardioid microphone picking up audio input primarily from the direction of focus. The fact that cheap earbuds are often designed for low amplitude output also means input power from any direction other than that of the focus is further reduced when used as a recording device.

3.2 Data collection Once the earbuds have been connected to a digital recording device, the consultant/participant is asked to hold one earbud just below or in a nostril, with the silicon tip facing upwards, and to hold the second one at the corner of the mouth, with the silicon tip facing forward.[3] With the recorder set to stereo, both earbuds will independently pick up output from their respective tracts. It should be noted that, without our asking, some consultants thought it was easier to simply insert the nasal earbud in their nostril instead of holding it. Preliminary analyses show little variation in the results based on whether a participant held the earbud below the nostril or inserted it, suggesting both methods serve for data collection. It was initially thought, however, that by holding the nasal earbud just below the nostril, air would be allowed to escape naturally, which would improve results, while inserting the earbud into the nostril might simulate a velum response similar to that of a speaker with nasal/sinus congestion. It is worth revisiting this line of thought in the future when a larger dataset is available, in case the different methods indeed yield significant differences. It can also be surmised that the comfort level of some participants may decrease if the nasal earbud is inserted into the nostril. For the comfort of the participants and to avoid any ethical complications, we strongly suggest asking participants to hold the earbud below the nostril unless the participants take the initiative to insert it themselves (see §5 for a discussion on ethics). The data presented here was recorded in stereo wave format with a sampling frequency of 44.1 kHz.

This study provides example data from the following 10 languages:

- Banjar (Indonesia; Austronesian)
- English (US; Indo-European)
- French (France; Indo-European)
- Japanese (Japan; Japonic)
- Kirundi (Burundi; Niger-Congo)
- Portuguese (Brazil; Indo-European)
- Media Lengua (Ecuador; Quichua-Spanish mixed language)
- Quichua (Ecuador; Quechuan)
- Siona (Ecuador, Tucanoan)
- Shiwiar (Ecuador; Chicham)

[3] Note that, as opposed to traditional nasometry approaches, no acoustic baffle is used between the nose and the mouth. This approach, which requires directional microphones to be placed immediately next to the nostril and mouth, was already proposed by Audibert and Amelot (2011).

3.3 Implementation While there are surely multiple approaches for analyzing the earbud data, here we detail three methods of visualization for interpreting the results. The first involves using Praat (Boersma and Weenink 2013) to juxtapose nasal and oral wave forms and to plot their opposing intensity measurements over them. By superimposing the opposing intensity curves (i.e., the nasal intensity curve over the oral wave form and vice versa), an additional layer of analysis is provided for determining the nasality of a given segment. The second and third methods involve using R (R Development Core Team 2008) to extract wave form amplitude in pascals and intensity in decibels. These analyses provide us with a basis for quantifying the results using several phonetic measures with different variables extracted from the signal. Scripts written in R, provided in the appendix, have been written to automate these processes by extracting the amplitude and intensity data to graph line plots of the results based on the independent tracts.

Due to the inherent differences in loudness between nasal and oral speech sounds, the sound file should undergo amplification normalization as a first step. This is an automated process which is available in most audio editing software packages. Amplification normalization increases or decreases the overall amplitude of a sound file based on a predefined decibel value (e.g., -3 dB). In Adobe Audition CC v.6, each sound track should be selected one at a time and normalization should be run on each track with the same value: Effects (located in the main menu) →Amplification and compression → Normalization process…→ *value* dB → OK. In Audacity v.2.1.0, each track should also be selected one at a time and normalization with the same value should be run separately on each track: Effect (located in the main menu) → Amplify…→ New peak amplitude → *value* db → OK.

In the second step, Praat is used to open the stereo recording and to extract both nasal and oral speech channels (*Convert → Extract all channels*). For analysis, a carefully annotated Praat TextGrid with IPA transcriptions of the utterance is necessary to help identify the type of airflow (nasal or oral) emanating from the vocal tract during a given sound segment. Annotations are best achieved by using the oral track as the nasal track should sound muffled. If one can clearly make out the speech in the nasal track, this is an indication that the nasal earbud picked up too much sound from the oral tract. If this is the case, consider lowering the gain and re-recording the stimuli.

Praat and other audio editing software typically normalize wave form amplitude based on the surrounding information in the view window. Therefore, if one zooms in on a perceived area of low amplification in Praat, the highest/lowest observable amplitude spike will always reach the top/bottom of the view window.[4] This attribute is beneficial for visualizing the tracks independently using the Praat Picture window as the wave forms of both tracks automatically adjust to their designated windows (making the amplitude normalization optional for simple viewing). However, while this typically reveals robust trends in the data by allowing us to visually identify instances of nasality, orality, and closure, amplification normalization (see §3.3.2) is

[4]To measure the actual pressure differences, maximum and minimum pascals (found on the left hand side of the wave form) also adjust with the view window in Praat.

still required before attempting any type of quantitative analysis. Therefore, it is always advisable to proceed with the amplification normalization prior to conducting any type of analysis (see Figure 10 and Figure 11 for differences in the raw data normalized and graphed in Praat and the data with amplification normalization graphed in R).

One disadvantage in observing normalized data is that if little to no information is found in a specific track, audio software such as Praat will still allow the loudest portion of the data (e.g., a minor blip on the baseline of 0.0004 Pa) to reach the top of the view window, suggesting there might be relevant information when there is not. In our analysis, this is most problematic when there are no nasal segments in the portion under analysis. To avoid this, we suggest comparing completely oral items with those that have at least one nasal segment within the same recording. It was suggested by a reviewer that a purely nasal consonant (ideally [ŋ] as it most closely approximates a naso-pharyngeal tube with little interference from the side branching oral tube) is best since nasal vowels make use of both the oral and nasal tracts.

As an anonymous reviewer also pointed out, ideally it would be best to calibrate both earbuds by playing a predetermined artificial sound source with both earbuds at an equal distance. While this is not always possible in a field scenario, the earbuds could be held at an equal distance while the consultant/participant phonates a given sound to be used as a base line (e.g., a vowel). It is even possible to do a post-hoc calibration using these suggestions. In theory, this would also make amplitude normalization unnecessary.

3.3.1 Praat for quick visualization of wave forms and intensity curves This method is advantageous as it allows for quick and accurate general comparisons of the nasal and oral data. Through this method, one should be able to identify both regressive and anticipatory nasalization, nasal vowels and consonants, closure in stop consonants, and emphasis. This method should not be used if there is interest in statistically quantifying a given segment, word, or phrase (see §3.3.2 for information on obtaining an objective metric for quantitative analysis).

After the stereo sound file has been loaded into Praat and the channels have been extracted, the intensity information needs to be isolated. To do so, begin by selecting both the extracted channels in the Praat Objects window and clicking the To *Intensity*...button. The default settings (Minimum pitch Hz: 100, Time step: Auto) have been found to be quite acceptable; however, the intensity can be smoothed by lowering the frequency or enhanced by increasing the frequency. This method was used to create Figure 2, and Figure 6 to Figure 9 in §4.2. The following steps detail how to create these visualizations in the Praat Picture window:

1. The first step is to select a box in the Praat Picture window beginning in the top left corner and dragging the highlighted box horizontally to 7.5 and vertically to 1.5.

2. In the main menu of the Praat Picture window, click *Pen* → *Solid line* and *Pen* → *Red*.

3. To draw the wave form, choose the nasal sound file in the Praat Objects window and click the *Draw* button.
    (a) We chose the *Poles* drawing method, though the *Curves* works just as well.
    (b) The ranges were left as their defaults and *Garnish* was unchecked.
    (c) Click *OK* to populate the wave form to the Praat Picture window.
4. In the Praat Picture window, highlight from 1 at the far-left side of the window and drag the mouse horizontally to 7.5 and vertically to 2.5.
5. Change colors to blue by clicking *Pen → Blue* in the main menu.
6. Steps 1 and 3 (skipping step 2) are then repeated using the oral track.
7. In the Praat Picture window, reselect the nasal track using the previous dimensions (0 VH to 1.5 V to 7.5 H).
8. Back in the *Objects* window, choose the oral intensity token and click the *Draw* button.
    (a) Maintain the defaults and uncheck Garnish.
    (b) Click *OK* to transpose the oral intensity curve over the nasal wave form.
9. The same steps are then repeated with the oral wave form.
    (a) Reselect the oral wave form in the Praat Picture window with the previous dimensions (1 V to 2.5 V to 7.5 H).
    (b) Change the pen color to red: *main menu → Pen → Red*.
    (c) Select the nasal intensity token in the Praat Objects window and click the *Draw* button.
10. To add the TextGrid, in the Praat Picture window, change the *pen* color to black: *main menu → Pen → Black*.
11. In the Praat Picture window, highlight from the top right corner to 7.5 H then to 3 V.
12. Back in the Praat Objects window, select the corresponding TextGrid and click the Draw button followed by *Draw...* using the default settings.
13. At times, it is also helpful to superimpose the nasal and oral intensity tracks over one another. To do this, we followed the same steps above for drawing the intensity tracks but kept them in the same highlighted area in the Praat Picture window. For the images presented here, we started on the far left side, from vertical point 3 to 4.5 and dragged the highlighted area horizontally to 7.5. For black and white publication purposes, it helps to change the pen style to dashed for one of the intensity curves (*Picture window → main menu → Pen → Dashed line*).

14. Finally, the image can be saved by choosing *File* in the Praat Picture window main menu followed by *Save as*... Several save options will appear, 300 dpi is lower quality than 600 dpi, but creates a smaller sized PNG image. For older versions of Praat, the Windows meta file or Praat picture files were the only options. The former can be visualized in MS Paint, while the latter allows you to reopen a closed image saved in this format in the Praat Picture window. Figure 2 is an example of the output created by following these steps. Further examples using this method can be found in §4.1. This figure provides an example from Kirundi, a Bantu language spoken in Burundi:

(1) *Ndavyumva*. [ⁿdadʑiõmva] 'I understand.'

The resulting visualization indicates that the word-initial syllable onset of this phrase is a prenasalized alveolar [ⁿd] which can be identified by an increase in amplitude lasting 94 milliseconds. The oral portion of the prenasalized stop, on the other hand, shows only minor fluctuations in amplitude. Upon cessation of the nasal portion of the segment, there is a brief loss of energy during the closure phase of [d] lasting 21 milliseconds. This can be observed in the decrease in energy in the wave form and the dip in the intensity curve in both the nasal and oral tracks. At the moment of release, a rapid increase in energy is observed in the oral track as seen in both the oral wave form and the oral intensity curve (superimposed over the nasal track). There is no noticeable change in either the nasal wave form or nasal intensity curve during the release phase. The difference in amplitude and intensity between the nasal and oral section from the beginning of voicing onset of the [a] vowel until the closure phase of the subsequent affricate ([ʥ]) suggests this vowel (128 ms in duration) is inherently oral. The greater amplitude readings from the oral track in the [ʥ] and [i] indicate these segments are also oral. For the [õ] segment, oral amplitude begins high, and tapers off throughout the vowel. At the same time, however, amplitude of the [o] segment in the nasal track begins low and increases throughout the vowel suggesting allophonic nasalization as the velum lowers in anticipation of the following [m] segment. The remaining portion of the phrase appears inherently oral after the velum closes and airflow into the earbuds is cut off after the [m] segment.

3.3.2 Methods for quantification This section describes methods for quantifying the nasal and oral data by extracting the amplitude and intensity values for analysis. To do this, each channel should be saved as a *.Sound file using the: *Save → Save as text file*... option in Praat. The intensity files extracted from the sound files with the *To Intensity*... button should be saved as a *.Intensity file using the same *Save* command as above. To manually extract the values from these files the following regular expressions can be run in R, EditPad Lite v.7.3.8, or any text editor equipped with regular expressions:

To visualize this data, the extracted values can be graphed in R using the automated *amplification* and *intensity* scripts in the appendix, which include the regular expressions in Equation 2. Figure 3 presents juxtaposed graphs of the amplitude and intensity data of the Shiwiar phrases shown in (2) and (3). The amplitude graph

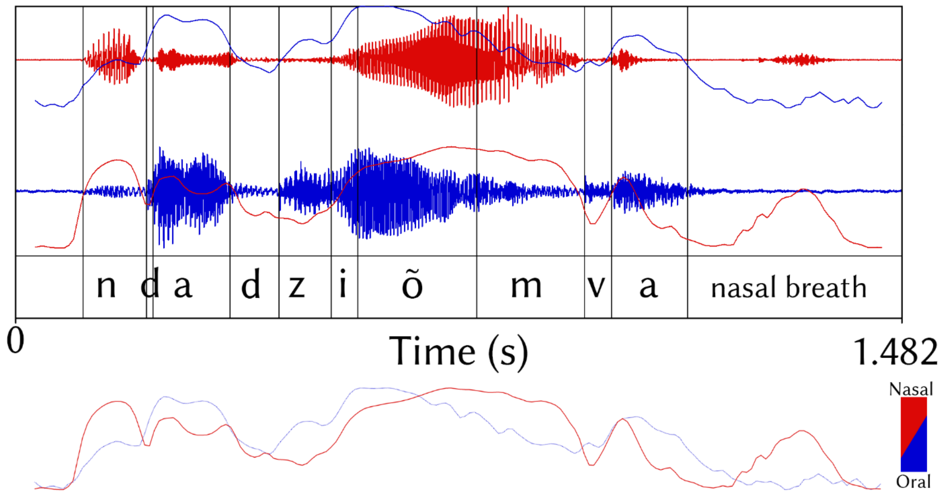

Figure 2. These graphs present a word-initial prenasalized stop and anticipatory nasalization in the [õ] vowel as the velum lowers to reach its target position for the [m] segment. The rest of the segments appear as oral. The nasal track appears as a red solid line while the oral track appears as a blue dashed line.

Step 1: Extract amplitude (Pa) /intensity (dB) values with mark-up.

R 3.2.1 syntax: `'z\\s\\[.*\\s=\(-|)\\d.'`
EditPad Light v7.3.8 syntax: `'z\s\[.*\]\s=\s(|-)(|0|\d+.\d+|\d+)(.*\S)` →
Cut matches button → Paste in new file.

Step 2: Remove non-numeric mark-up from the data.

R 3.2.1 syntax: `'z\\s\[.*\\s=\\s',''`
EditPad Lite v7.3.8 syntax: `z\s\[.*\]\s=\s` → *Cut matches*

Equation 2. This regular expression, presented in both R and EditPad Lite (Goyvaerts 2014) syntax, extracts amplitude values in pascals from *Sound* text files.

(top) contains the nasal wave form (red) in pascals superimposed with the oral wave form (blue). The intensity image (bottom) contains the nasal intensity track (red) in decibels superimposed with the oral track (blue). The segment annotations are time aligned based on the original Praat TextGrid while the boundaries of each segment are shown in gray lines.

(2) *Yumi yutawai.* [jumĩ jutawai̯] 'It's raining.'

(3) *Numin achikmahai.* [nũmĩn atʃikmahai̯] 'I grabbed a stick.'

Like Figure 2, instances of increased nasal amplitude sustained for a meaningful duration compared to instances of decreased oral amplitude in the same segment

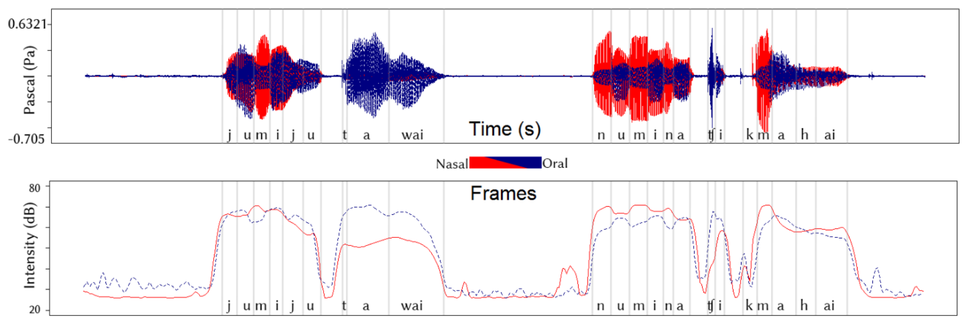

Figure 3. These graphs represent the wave form amplitude (Pa) and the intensity curves (dB) of the Shiwiar phrases presented in (2) and (3). The nasal track appears as a red line, while the oral track appears as a blue dashed line.

correlate to velum lowering in the speech signal (typically the case with nasal consonants). Instances where both tracks have similar values beyond the baseline might suggest both tracts are used in the production of a given speech sound (typically the case with nasal vowels). If a given sound is known to be completely oral (e.g., [p]) yet there is a substantial level of energy in the nasal track, it is recommended that additional tokens of the same segment, word, or phrase be tested for consistency. If a completely oral sound is showing a substantial degree of nasality, the nasal track should be independently played to see if output from the oral tract is being picked up by the nasal earbud; if so, consider reducing the gain or choosing a different pair of earbuds.

The basic calculation of nasalance can be used to create an objective metric in order to quantify the nasal and oral data presented in this method. Here, Equation 3 illustrates a simple interpretation of the intensity results by dividing a given nasal value (A) (in decibels (dB)) by the sum of itself (nasal A) and the oral (B) value found in the same frame. The calculation creates a nasalance ratio representing the proportion of nasal to oral energy within the signal. Multiplying this result by 100 provides the results as nasalance percentages. It should be noted that without further normalization, this calculation renders difficult to interpret results with Pascals (Pa) due to the large range of potential values.

$$\frac{A_{nasal}}{A_{nasal}+A_{oral}}(\text{Ratio})\frac{100*A_{nasal}}{A_{nasal}+A_{oral}}(\text{Percentage ratio})$$

Equation 3. This basic nasalance calculation (summarized in Rothenberg, unpublished MS) places the nasal and oral output into a ratio or percentage ratio to quantify the results for analysis.

Figure 4 provides three graphs representing the Siona phrase shown in (4). The top graph illustrates the superimposed nasal and oral wave forms, the center graph illustrates the superimposed nasal and oral intensity curves, and the lower graph illustrates the results from the ratio calculation of intensity from Equation 3. The ratio graph reveals that nearly the same portions which represent nasality in the wave forms and intensity curves have a nasal to oral ratio above 0.50 (50%).

(4) *J̰ao̰ baco*. [hã͡õ bak͡o̥] 'She has.'

Figure 4 reveals slight aspiration in the utterance-initial [h] followed by a clearly nasal diphthong ([ã͡õ]) which is attested by the higher amplitudes (represented in red) in the first two graphs. There also appears to be a degree of voicing taking place during the closure of [b] which can be observed in both the wave form and the fact that the intensity curves do not drop to the baseline as would be expected during complete voiceless closure (see [t] and [k] in Figure 3). In the rest of the audible utterance, the oral track dominates the nasal track as attested by the higher amplitudes represented in blue (the final segments ([k͡o̥]) are nearly inaudible). It is worth noting that while the ratio calculations may provide a reliable metric, it may report non-significant micro-perturbations as substantial differences between the nasal and oral tracks during moments of low energy (e.g., during closure, before or after a word). This can be seen during the [k͡o̥] segments at the end of this utterance which show the nasal track as having a higher level of energy compared to the oral graph. To avoid this, it is suggested to remove portions of the ratio data which appear during moments of silence (e.g., during stop closure, glottal stops, between words and phrases, and at the beginning and end of words). See Figure 10 and Figure 11 for examples of silence removed from the ratio data.

After quantifying the data, graphing the ratio values allows for an additional layer of comparison of a particular segment, word, or phrase. Typically, if the ratio line is above 0.5 (50%) the segment is considered to have more nasal energy than oral, while if it is below 0.5 it is considered to have more oral energy (see Figure 10 and Figure 11). This trend typically matches the data observed in the wave form and intensity curve graphs. Graphing the ratio values might also be beneficial for identifying segments which are visually inconclusive in the superimposed wave forms or intensity curves or for identifying trends across multiple tokens of similar segments to determine coarticulation, allophonic, or phonemic patterns.

4. Results and analyses

4.1 Wave form and intensity analysis Like the Kirundi example provided in Figure 2, Figure 6 to 11 allow for the identification of nasal and oral turbulence across segments, words, and phrases. Areas in the wave form and intensity curves with high energy and longer durations of sustained energy in the nasal track indicate that airflow from the nasal cavity made contact with the earbuds, causing an increase in pressure against the diaphragms which is represented as noise (areas of increased loudness) on the recording.

In purely nasal sounds (e.g., [m, ɱ, n, ɳ, ɲ, ŋ, ɴ]), it is expected that the amplitude of the nasal track will be greater than that of the oral. This can be clearly seen in the [n] and [m] segments in Figure 9 where the energy from the nasal track is apparent

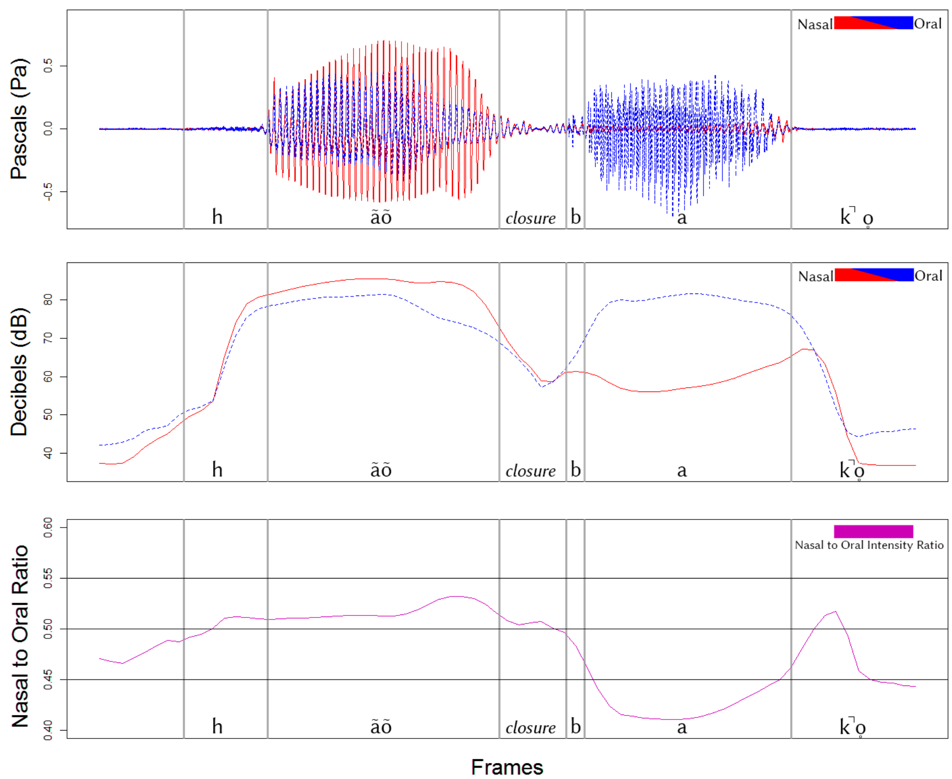

Figure 4. This figure represents the wave forms (superior), intensity curves (center), and nasal to oral ratio (inferior) of the Siona utterance *ja̱o̱ baco* [hã͜õ bak̚o̥] 'she has'. In the first two graphs the nasal track appears in red while the oral track appears in blue. The final graph represents the nasal to oral ratio of energy in the signal with segments of higher nasal to oral ratios extending to the top of the graph and those with higher oral to nasal ratios extending towards the bottom of the graph.

in the wave form amplitude, while the amplitude in the oral track hovers around the baseline. Similar results can be seen in Figure 3 in the [m] in *yumi*, the [n, m, n] in the word *numin*, and the [m] in word *achikmahai*. Conversely, sounds which have both nasal and oral output, such as nasal vowels (e.g., [ĩ, ẽ, æ̃, ã, ɔ̃, ũ]) are expected to show increased amplitudes in both tracks (see the [u, i] segments in the word *numin* in Figure 3 and the [ã͜õ] diphthong in Figure 4).

Because velum aperture can take place at varying degrees, we might expect to see observable differences in the nasal wave forms and intensity curves during instances of nasal leakage, allophonic nasalization, and phonemic nasalization. Nasal leakage may be observed as brief moments of low level amplification that stems beyond the baseline (see utterance initial [b] in Figure 5). Nearly every utterance-initial voiced stop in Media Lengua from a corpus of 960 utterances beginning with [b, d̪, g], from 10 speakers, revealed some degree of velum aperture. Figure 5 reveals a brief moment of nasalization in [b] as the negative VOT initiates. Nasalization continues throughout most of the segment but is eventually overtaken by oral prominence. This finding was expected as nasal leakage preceding an utterance-initial voiced stop has been reported in Spanish (the lexifier language of Media Lengua) as a mechanism to achieve a difference in transglottal airflow pressure allowing for voicing (Solé and Sprouse 2011).

Allophonic nasalization in the form of pre and post-nasalization (or anticipatory/carryover or regressive/progressive nasalization) is often observed as a gradient increase or decrease in nasal amplitude preceding or following a phonemic nasal segment (see [õ] in Figure 2 for an example of anticipatory nasalization). It is also common to see entirely inherently oral segments undergo nasalization between two inherently nasal segments (see the [u, i] segments in the word *numin* in Figure 3). Figure 5 also shows three clear instances of anticipatory nasalization and one clear instance of carryover nasalization in the vowels surrounding both dental nasals ([n̪]) and the bilabial nasal ([m]) – an allophone of [n̪] when preceding bilabial stops. In the first instance the air from the nasal tract begins to make contact with the earbud towards the beginning of the vowel, but does not take over the oral track in prominence until nearly three-fourths of the duration into the vowel. After the [n̪], nasal energy slowly decreases in the following [a] until the oral track takes over in prominence about half-way through the vowel. In the [on̪] cluster, nasalization begins at the beginning and steadily increases throughout the vowel ([o]) into [n̪]. Nasal prominence in the vowel clearly appears in the wave form, but the gradual transition is better seen in the intensity and ratio graphs. The last instance of anticipatory nasality ([ion̪]) is similar to the previous, though the entire cluster is nasalized as observed in all three graphs in Figure 5.

Phonemic nasalization in consonants is typically observed with increased levels of energy throughout an entire segment in the nasal track with the reverse trend in the oral track (see the word initial segment in Figure 2). Phonemic nasalization is often, though not always, more prominent than allophonic nasalization. For phonemic vowels, it is common to see similar levels of increased energy in both tracks as nasal vowels make use of both the oral and nasal cavities during production (see [ɛ̃] in the final word in Figure 11). However, it is typically the case that nasal amplitude outweighs that of the oral track in nasal vowels (see the diphthong in Figure 4 and the [u, i] segments in in the word *numin* in Figure 3). Figure 5 contains three phonemic nasal consonants (two [n̪] and [m]), which all show greater prominence in nasal energy compared to both the oral output during these segments and the anticipatory and carryover nasalization in the surrounding vowels. While the velum appears to lag in closure in the first instance of [n̪], therefore creating the carryover nasality in the following vowel, the last two instances abruptly stop at the end of the segment during the closure phases of the following affricate ([ʧ]) and stop ([p]) consonants. Finally, it is worth noting that there appears to be some degree of nasalization taking place in the oral cluster [e̞le], though the oral track maintains prominence. This may be feedback from the oral tract picked up by the nasal earbud. Figure 6 provides three graphs representing the Media Lengua phrase shown in (5).

(5) *Votanajunchi electionpi.* [botãn̪ãˈhõn̪tʃ e̞lexˈsĩõmpi] ‘We vote together in the election.’

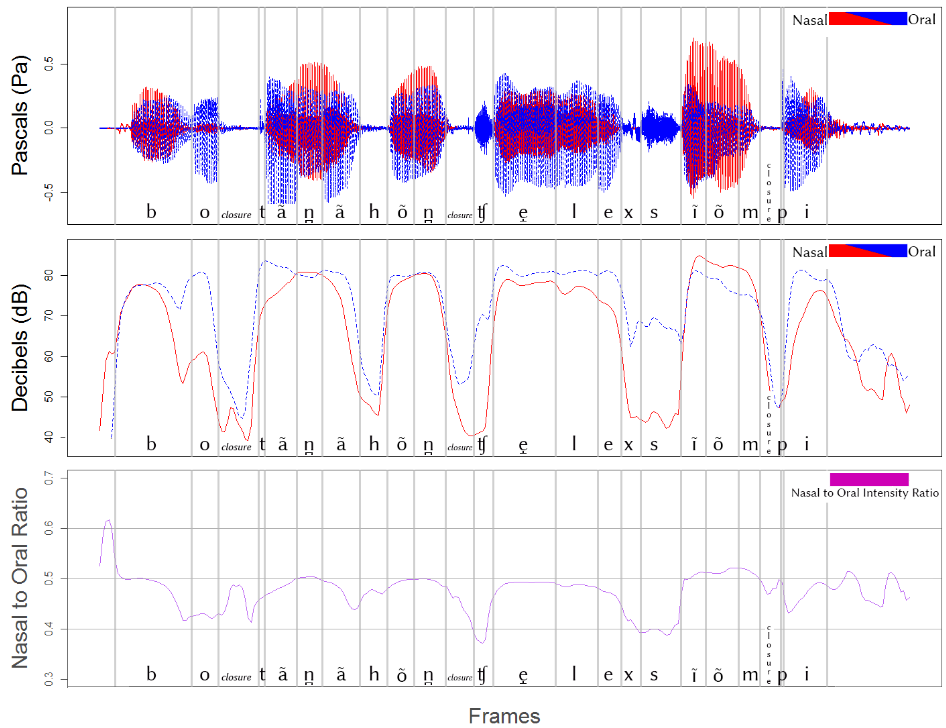

Figure 5. This figure represents the wave forms (superior), intensity curves (center), and nasal to oral ratio (inferior) of the Media Lengua utterance: *Votanajunchi electionpi.* [botãnãˈhõntʃ e̞lexˈsĩõmpi] ‘We vote together in the election.’ In the first two graphs the nasal track appears in red while the oral track appears in blue. The final graph represents the nasal to oral ratio of energy in the signal with segments of higher nasal to oral ratios extending to the top of the graph and those with higher oral to nasal ratios extending towards the bottom of the graph.

It is also worth mentioning once more that while we are conducting a comparative analysis, the automated normalization process may over- and under-compensate energy readings if there are no nasal segments in a given utterance under analysis. The best way to avoid this is by analyzing larger stretches of data at a time where at least one nasal segment is present (two or three words, an entire phrase, etc.).

4.2 Cross-linguistic analyses This section provides examples from Banjar, Japanese, Quichua, and English to show how the earbud method can be used to examine a number of phenomena related to nasalization, including nasal spreading, nasal leakage, and anticipatory and progressive nasalization. Each figure is accompanied by a description of several points of interest regarding the nasal and oral patterns in the speech stream. Because the figures in this section were created in Praat, which automatically normalizes objects in the Praat Picture Window, observations only show the general trends in the data. The figures are formatted to contain the nasal track (red) above the oral track (blue). The opposing intensity curves are superimposed over the wave forms for additional analysis. Figure 6 presents example (6) from Japanese as spoken in Sapporo, Japan:

(6) 新聞 - しんぶん *(shimbun)* [ɕĩmbũɴ] 'newspaper'

As indicated by the high energy in both the oral wave form and oral intensity curve during the production of the word-initial alveolo-palatal fricative ([ɕ]) and the low energy level present in the nasal track in the same segment, this clearly suggests [ɕ] is oral. On the other hand, the increase in energy in both the wave form and intensity curve in the nasal track suggests the following two segments ([i] and [m]) are produced with airflow exiting the nasal passage. As would be expected with oral stop consonants, the energy levels in both the nasal and oral tracks quickly drop off during closure and then the oral track spikes during the release phase of [b]. Upon voicing of [u], nasal energy remains low until the mid-point of the vowel when the velum begins to lower in anticipation of the final nasal segment [ɴ].

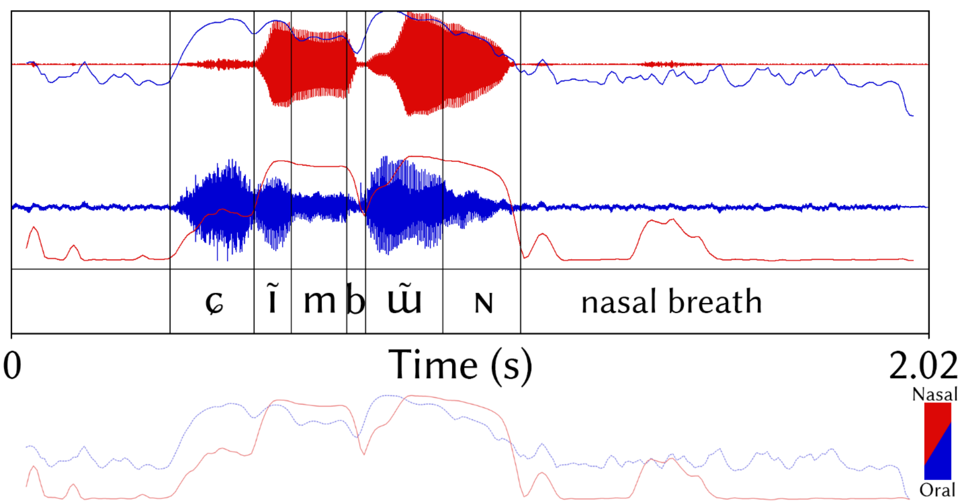

Figure 6. These graphs present the wave forms and intensity curves for the nasal and oral tracks of the Japanese word *shimbun* [ɕĩmbũɴ] 'newspaper.' The data suggests there is anticipatory nasalization in both vowels in anticipation of the nasal coda segments. The nasal track appears in red, and the oral track in blue.

Figure 7 presents example (7) from Banjar as spoken in Yogyakarta, Indonesia:

(7) *bang* [bãŋ] 'to rise'

The word in Figure 7 begins with a slight prominence of nasal pressure at the beginning of the [b] segment, which is interpreted as nasal leakage similar to the utterance-initial [b] in the Media Lengua example in Figure 6. There is some indication that [a] shows anticipatory nasalization just after the half-way point with a steady increase in energy in both the nasal wave form and nasal intensity curve. At the same time, as the nasal energy begins to increase, the oral energy decreases.

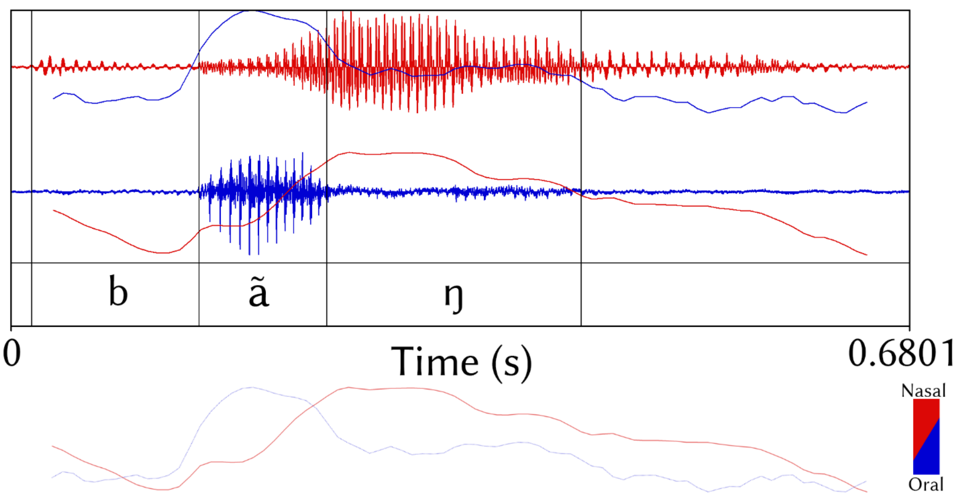

Figure 7. These graphs present the wave forms and intensity curves for the nasal and oral tracks of the Banjar word *bang* [bãŋ] 'to rise.' The data suggests utterance-initial prenasalization at the beginning of [b]. Allophonic nasality in the vowel [a] is also present in anticipation of the nasal coda segment. The nasal track appears in red and the oral track in blue.

Figure 8 presents example (8) from Karanki Quichua, spoken in Imbabura, Ecuador:

(8) *Cuyka mishki mishkimi kan.* [ˈku͜ika ˈmĩʃki mĩʃˈkimĩ kaŋ] 'This guinea pig is delicious.'

This analysis demonstrates that the first word in the phrase (*cuyka*) appears to be completely oral as indicated by the high energy in both the oral wave form and oral intensity curve, while the low levels of energy from the nasal track also support this interpretation. The analysis also reveals carryover nasalization in the front high vowels ([i]) in the third, fifth, and seventh syllables throughout this phrase which are all made up of the same cluster ([mĩ]).

Figure 9 presents an example from English as spoken in the Pacific Northwest of the United States:

(9) *I can't make it.* [a͜ɪ kʰæ̃·n me͜ĩ kʰɪt̚]

Vowels in English are known to be realized as nasal when they precede a nasal consonant. However, Cohn (1990; 1993) and Keating (1988) report that English vowels are not completely nasalized in this environment as there is no specified target English speakers must reach to ensure the vowel is fully nasalized, unlike languages with phonemic nasal vowels such as French, Portuguese, or Guaraní. This type of coarticulatory effect can be seen in Figure 9. However, unlike the findings reported by Cohn (1990; 1993) and Keating (1988) showing the vowel is only partially nasalized,

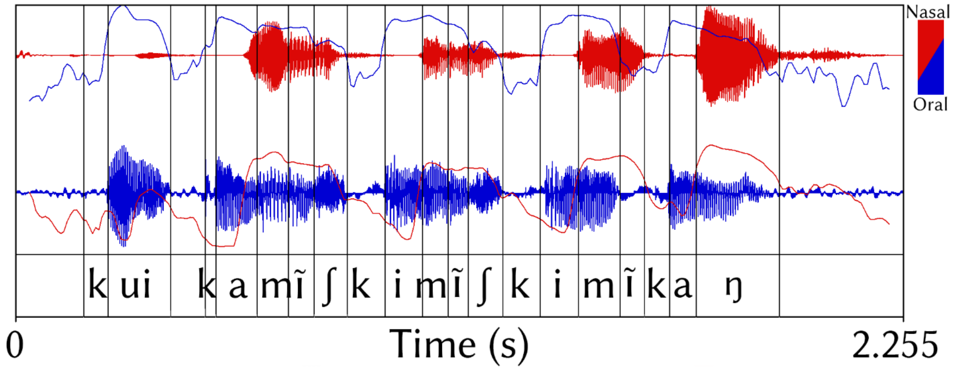

Figure 8. This graphs present the wave forms of the nasal and oral tracks superimposed with their opposing intensity curves. The phrase under analysis is *Cuyka mishki mishkimi kan* [ˈku̯ika ˈmĩʃki mĩʃˈkimĩ kaŋ] ‘The guinea pig is delicious’ from Quichua. The graph reveals carryover nasalization on vowels following nasal segments, a small degree of prenasality on the [a] preceding [m], and well defined closure and release phases for stop consonants in both tracks.

the energy output during the production of [æ̃] is greater in the nasal track (observed in both the wave forms and intensity curves). Moreover, carryover nasalization is also observed in [ẽɪ̃], suggesting [m] substantially influences the production of both its surrounding vowels in this utterance.

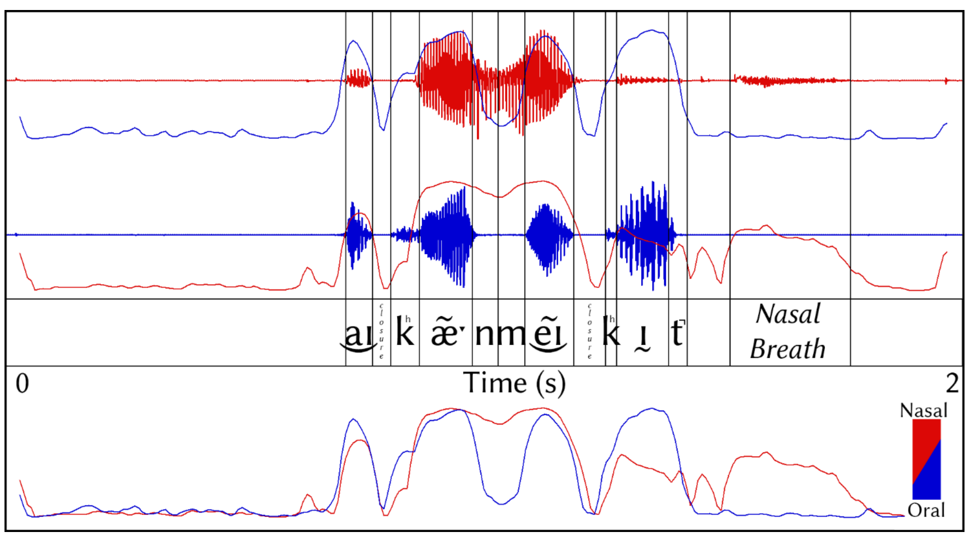

Figure 9. These graphs present juxtaposed nasal and oral tracks in the form of wave amplitudes with their opposing intensity curves superimposed on the wave forms for the English phrase *I can’t make it*.

4.3 Cross-technique comparisons This section provides examples from Portuguese (Figure 10) and French (Figure 11) – two languages with phonemic vowel nasality. Here, raw data used to generate images in Praat, which automatically normalizes the data in the Praat Picture Window, are compared to the same data, but where each track has undergone independent automatic external amplitude normalization to -3 dB of the peak amplitude. These latter images are generated using R. Figure 10 presents several isolated words from Brazilian Portuguese as spoken in Rio de Janeiro. All of the examples in (10) contain the open-syllable word-final nasal diphthong [ɐ̃u̯].

(10) *não* [nɐ̃u̯] 'no', *cartão* [kahtɐ̃u̯] 'cardboard', *tapetão* [tapetɐ̃u̯] 'soccer tribuna', *inflação* [ĩflasɐ̃u̯] 'inflation'

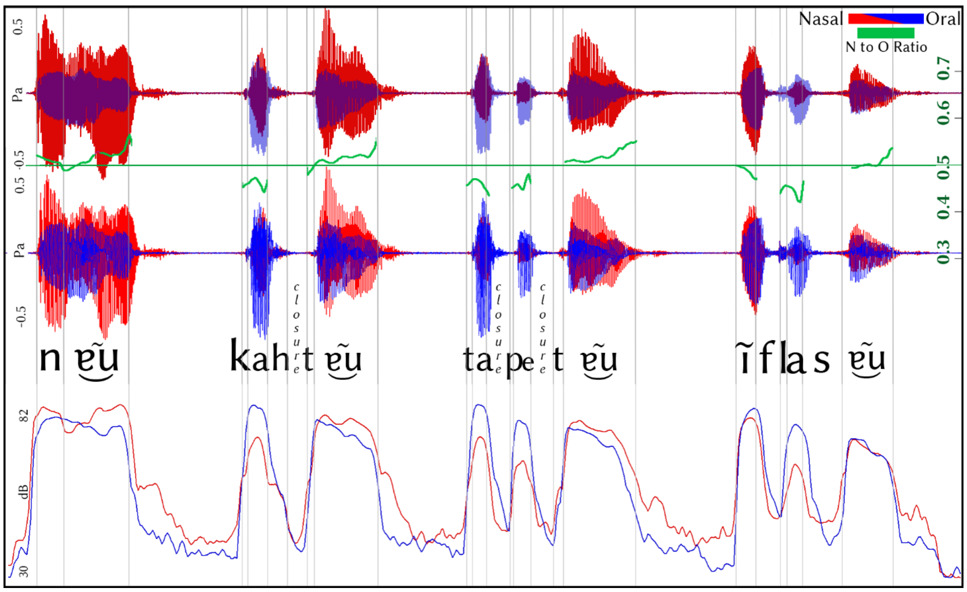

Figure 10. These graphs present wave forms generated on both Praat (upper) and R (lower), along with the nasal to oral ratio data presented in green. Below the wave forms, superimposed intensity curves generated in Praat are placed to use as an additional point of comparison. This figure represents the four Portuguese words found in example (10). All four words contain word-final nasal diphthongs which are reflected in the high-energy readings in the nasal tracks (in red) in word-final position.

Figure 10 provides a comparison between the superimposed wave forms generated in Praat (upper) and those generated in R (lower). In addition, the nasal to oral ratio based on the intensity track of these data is superimposed in the middle of the graph in green, acting as a metric to judge both wave form plots. At first glance, both wave forms appear to be quite similar. However, one important difference is the nasal to oral amplitude. In the Praat version, the oral track is more compact, which might lead to a misinterpretation if values are similar. Based on the nasal to oral ratio graph, data above the 0.5 line (the 50% point where nasal and oral energy output is equal), correlate to regions of high energy output in the nasal track, while data below the 0.5 line correlate to regions of higher oral output.

One discrepancy in this trend appears in the word-initial nasal high vowel ([ĩ]) in the final word *inflação* 'inflation.' In the Praat generated version (upper), [ĩ] shows greater nasal energy than oral, whereas in the R generated version (lower), the energy values appear to be more even. Comparing these representations to the ratio plot, it appears the R generated graph better reflects the data, since the ratio line begins at the 0.5 mark and slightly decreases throughout the segment, similar to the R-generated wave form. To insure this was not simply caused by normalizing the amplitude prior to analysis, the intensity curves of the raw data version are also present. The superimposed intensity curves reveal a similar pattern of energy distribution during [ĩ] with the segment beginning with nearly equal energy before the oral track gains prominence by the end of the segment.

However, to make brief impressionistic observations of the data, the Praat-generated version reflects the R version relatively well. It is clear that the first word, made entirely of nasal segments, shows elevated nasal amplitude across the word in the wave forms, intensity curves, and nasal to oral ratio line. The middle two words begin with oral segments until the nucleus of the last syllable. This is reflected in all four methods of observation as well.

Figure 11 presents several isolated words from European French as spoken in the south of France. The examples in (11) contain three of the four phonemic nasal vowels found in French (/ɑ̃, ɛ̃, ɔ̃/).

(11) *médicament* [medikamɑ̃] 'medicine', *imperméable* [ɛ̃mpɛʁmiabl] 'raincoat', *tomber* [tɔ̃be] 'to fall', *brun* [bʁɛ̃] 'brown'

Figure 11 provides a comparison between the superimposed intensity curves generated in Praat (upper) and those generated in R (lower). In addition, the nasal to oral ratio based on the intensity track of these data is superimposed in the middle of the graph in green acting as a metric to judge both intensity curve plots. For additional analysis, the superimposed nasal and oral wave form generated in R is also present at the bottom of the figure. Similar to the wave forms in Figure 10, on first glance, both intensity curves appear to be quite similar. However, one important difference is observed in the height of the curves. In the Praat version, the oral track is more elongated which might lead to a misinterpretation if the automatic normalized values are similar.

One potential discrepancy between the two intensity curves appears in the final segment in the last word *brun* 'brown' where the height of the oral intensity curve is higher in the Praat generated graph compared to that in R. This could lead to an interpretation that [ɛ̃] is more oral than it actually is. Based on the ratio line (green) the last half of the vowel reaches over the 0.5 (50%) marker, suggesting there is more nasal energy being produced than at the beginning of the vowel. This is slightly obscured in the Praat generated graph as compared to that of R, which suggests the energy values are closer to equal – a result supported by observations in the wave form and ratio graph. However, to make brief impressionistic observations of the

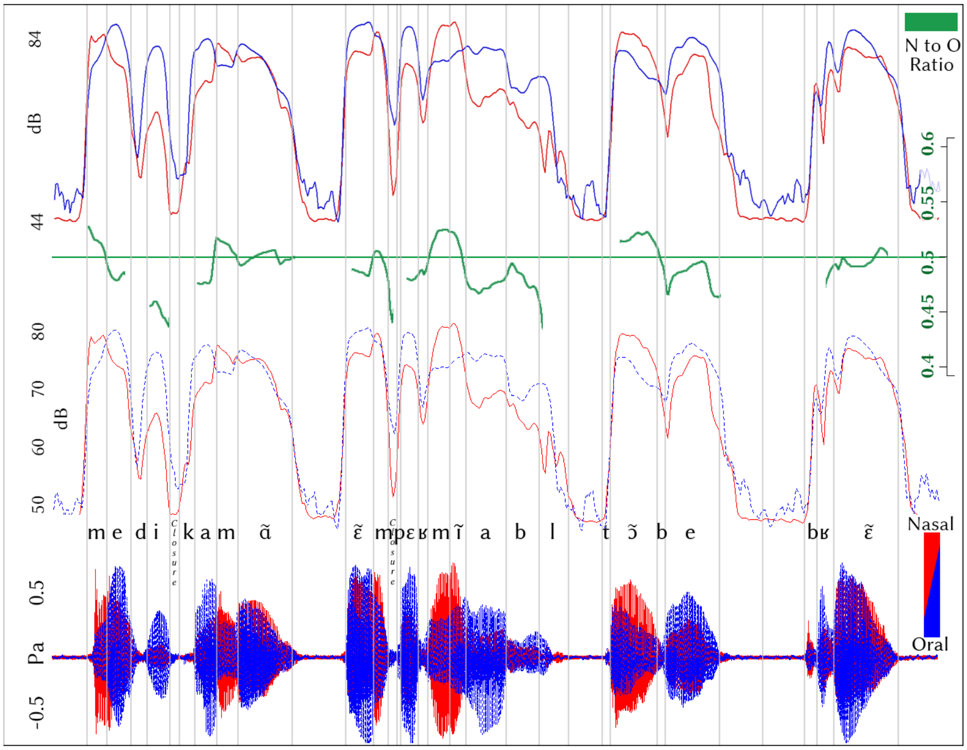

Figure 11. These graphs present intensity curves generated in both Praat (upper) and R (lower) along with the nasal to oral ratio data presented in green. Below the intensity curves, superimposed wave forms generated in R are placed to use as an additional point of comparison. This figure represents the four French words found in example (11).

data, the Praat generated version method also reflects the R version relatively well, and only causes issues when the values are similar in energy output.

This specific example also provides a basis for comparing differences in nasal and oral energy output in inherently nasal consonants and nasal vowels. Since the latter makes use of both the oral and nasal cavities for production, while the former only makes use of the nasal cavity, more nasal energy is expected from the consonants while the vowels should show increased energy in both tracks (though this is not always the case – see Figure 6 and Figure 9). The first two words contain four bilabial nasals ([m]) and two nasal vowels (word-final for the first ([ɑ̃]) and word-initial for the second ([ɛ̃])). The greater energy in the nasal track compared to that of the oral track during the production of the nasal consonants suggests that these are in fact purely nasal segments which do not make use of the oral cavity. The nasal vowels in these first two words, however, show similar energy levels, suggesting these segments are making use of both cavities during speech production. The final two words both contain nasal vowels but lack nasal consonants. The penultimate word contains the nasal vowel [ɔ̃], which, unlike the other nasal vowels, shows a great deal of prominence over the oral track. This may be a coarticulation effect caused by anticipatory closure for the following [b] segment which closes off the oral tract but might continue to allow air to escape through the nasal cavity, similar to an [m]. This hypothesis is supported by the overall decrease in oral energy from an initial position roughly equal to that of the nasal track, to a final position with little energy. Similar to the nasal vowels in the first two words, the word-final [ɛ̃] in *brun* 'brown' appears to be produced with similar energy output from both cavities.

5. Fieldwork and ethics The low cost and highly portable nature of this technique makes it a perfect tool for linguistic fieldwork. With it, descriptive and documentary linguists can provide nuanced and precise descriptions of nasality in the languages they work on without the need to take specialized equipment to their field site. However, as with any other data collection technique and especially in the context of fieldwork, the ethical pros and cons of using this tool must always be considered.

The main disadvantage of this technique involves the slight discomfort and awkwardness the consultants might experience by having to hold an earbud slightly below or inside their nostril during a recording session. Nevertheless, in our experience, consultants largely expressed amusement rather than discomfort, and any initial uneasiness was quickly resolved. This technique is only minimally invasive if the participants being recorded chose to insert the earbud into their nostril, but the linguist should explain that that is by no means a requirement. One important advantage of using earbuds is that consultants in many parts of the world are already familiar with them given the rapid spread of mobile phones (and their corresponding accessories) in recent years. This means that beyond explaining the peculiarities of where to place the earbuds for this method, consultants will not be intimidated by the equipment itself. To dispel any doubts that a consultant may still have about the method, it is also advisable for linguists to model the technique on themselves before asking the consultant to do it. Finally, it should be noted that the physical invasiveness and discomfort caused by this technique is virtually negligible compared to all other articulatory or aerodynamic methods to collect nasality data. In conclusion, given the methodology proposed in this paper, we do not believe that this method presents any particular ethical challenges for the linguist.

6. Discussion The technique described here provides an accurate and cost-effective method for analyzing nasality in the field.[5] The steps outlined here offer researchers a new method – based on previously existing nasalance and nasometry techniques – for identifying nasality in the field. Both phonemic and allophonic nasalization trends can be explored in the amplitude and intensity patterns generated through sound output from the nasal and oral tracts. The method also allows for a suprasegmental approach for visualizing oral and nasal interactions, and instances of microperturbations (such as momentary nasal leakage due to partial velum aperture) via

[5] This technique is also an excellent exploratory tool in a laboratory setting. However, detailing how this method can be used in conjunction with more advanced articulatory and aerodynamic tools in the lab is beyond the scope of this paper.

differences in energy values. Furthermore, coarticulation effects involving the velum, as in instances of anticipatory and progressive nasalization, can also be investigated with this technique. Future applications of this method could focus on segmental interactions (e.g., stop opacity in languages with nasal harmony).

One question that remains open for further investigation is how well this technique distinguishes voiceless nasal segments, such as those found in Burmese and Welsh. It may be that the nasal earbuds are picking up voicing cues rather than perturbations caused by nasal airflow hitting the nasal earbud diaphragm. It would also be important to compare this technique with more established methods (e.g., airflow masks, nasometers with an acoustic baffle, spectral measurements, nasoendoscopies, etc.), in order to understand further limitations of the earbud method. Finally, investigating how the formant patterns in oral vowels change when the nasal track is combined with the oral track in a mono recording may be of interest. While the formants from the nasal cavity do not change (as the shape of the cavity does not vary), combining these formants with the oral formants may shed some light on how oral formants and spectra react to the presence of nasal output.

References

Audibert, Nicolas & Angélique Amelot. 2011. Comparison of nasalance measurements from accelerometers and microphones and preliminary development of novel features. *Proceedings of INTERSPEECH 2011 – 12th Annual Conference of the International Speech Communication Association*, Aug 2011, Florence, Italy. 2825–2828.

Baken, Ronald. 1987. *Clinical measurement of speech and voice*. Boston: College-Hill Press.

Benguerel, A.P., H. Hirose & M. Sawashima. 1975. Velar height and its timing in French: A fiberscopic study. *Annual Bulletin Research Institute of Logopedics and Phoniatrics* 6. 67–78.

Berez, Andrea L. 2015. Reproducible research in descriptive linguistics: integrating archiving and citation into the postgraduate curriculum at the University of Hawai‘i at Mānoa. In Amanda Harris, Nick Thieberger & Linda Barwick (eds.), *Research, Records and Responsibility: Ten Years of PARADISEC*, 39–51. Sydney: Sydney University Press.

Boersma, Paul & David Weenink. 2013. *Praat* (version 5.3.47). Amsterdam: Institute of Phonetic Sciences, University of Amsterdam.

Brenner, Daniel, Andrea Davis, Natasha Warner, Andrew Carnie, Muriel Fisher, Jessamyn Schertz, Michael Hammond & Diana Archangeli. 2011. Can you say [v] or [x]? Aerodynamics of nasalized fricatives in Scots Gaelic. *Journal of the Acoustical Society of America* 130: 2550–2550.

Chen, Marilyn. 1996. *Acoustic correlates of nasality in speech*. Cambridge, MA: MIT dissertation.

Cohn, Abigail. 1990. *Phonetic and phonological rules of nasalization*. Los Angeles, CA: UCLA dissertation.

Cohn, Abigail. 1993. Nasalisation in English: Phonology or phonetics. *Phonology* 10. 43–82.

de Paula, Matías N. 2016. *A description of the sound system of Misiones Mbya*. Morgantown, WV: WVU MA thesis.

Demolin, Didier. 2011. Aerodynamic techniques for phonetic fieldwork. *Proceedings of the International Congress of Phonetic Sciences (ICPhS) XVII*, Aug 2011, Hong Kong. 84–87.

Fletcher, S., L. Adams & M. McCutcheon. 1989. Cleft palate speech assessment through oral nasal acoustic measures. In KR Bzoch (ed.), *Communicative disorders related to cleft lip and palate*, 246–257. Boston: Little Brown.

Fletcher, Samuel G., Iradj Sooudi & Sharon D. Frost. 1974. Quantitative and graphic analysis of prosthetic treatment for 'nasalance' in speech. *The Journal of Prosthetic Dentistry* 32(3). 284–291.

Goyvaerts, Jan. 2014. E*ditPad Lite* (version 7.3.3 x64).

Keating, Patricia. 1988. Underspecification in phonetics. *Phonology* 5. 275–292.

Ladefoged, Peter. 1999. Instrumental techniques for linguistic phonetic fieldwork. In William Hardcastle & John Laver (eds.), *The Handbook of Phonetic Sciences*, 137–165. Oxford: Blackwell Publishing.

Ladefoged, Peter. 2003. *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. Oxford: Blackwell Publishing.

Ohala, John J. 1971. Monitoring soft-palate movements in speech. *Journal of the Acoustical Society of America* 50. 140(A).

R Development Core Team. 2008. *R* (version 2.12.2). Vienna, Austria: R Foundation for Statistical Computing.

Riehl, Anastasia Kay. 2008. *The phonology and phonetics of nasal obstruent sequences*. Ithaca, NY: Cornell University dissertation.

Rothenberg, Martin. ms. Mask-based nasometry: A new method for the measurement of nasalance.

Solé, Maria-Josep & Ronald Sprouse. 2011. Voice-initiating gestures in Spanish: Prenasalization. *Proceedings of the International Congress of Phonetic Sciences (ICPhS) XVII*, Aug 2011, Hong Kong. 110–113.

Styler, Will. 2015. *On the acoustical and perceptual features of vowel nasality*. Boulder, CO: UC dissertation. http://savethevowels.org/will/dissertation.html.

Jesse Stewart
stewart.jesse@usask.ca

Martin Kohlberger
m.kohlberger@hum.leidenuniv.nl

Appendix A. Amplitude R script

The R script Amplitude.R is also available at https://scholarspace.manoa.hawaii.edu/bitstream/10125/24724/2/Amplitude.R.

```
#Amplitude visualisation in R v3.2.1

#To begin, open each tract in Praat (Boersma and Weenink 2013), select each 'Sound' file and save each one separately as a text file: Main menu --'Save' --'Save # as text file...'. Save these files in the same working directory you will use for R.

#Loading 'Sound' text files. Be sure to change address to the correct working directory.
nasal=read.table("%DIRECTORY%/SOUNDFILENAME_ch1.Sound", header=T, sep="\t", encoding="UTF-8")
oral=read.table("%DIRECTORY%/SOUNDFILENAME_ch1.Sound_ch2.Sound", header=T, sep="\t", encoding="UTF-8")

#For Windows users using characters other than standard ASCII you may have to change the encoding from UTF-16 to UTF-8.
#In EditPad Lite (Goyvaets, 2015) (free download) click 'Convert' in the main menu --> Text Encoding --> Under 'New Encoding' choose 'Unicode, UTF-8' --> OK. Then save.
textgrid=read.csv("%DIRECTORY%/TEXTGRIDNAME.TextGrid", header=F, encoding="UTF-8")

#Renaming the first column of the 'nasal' and 'oral' objects
colnames(nasal)[1]="Nasal"
colnames(oral)[1]="Oral"

#Subsetting the 'nasal' and 'oral' objects to get the amplitude (Pascal) data.
N=subset(nasal, grepl("z\\s\\[.*\\s=\\s(-|)\\d.",nasal$Nasal))
O=subset(oral, grepl("z\\s\\[.*\\s=\\s(-|)\\d.",oral$Oral))

#Removing non-numeric data from the 'N' and 'O' objects.
N=as.data.frame(apply(N,2,function(i)gsub('\\s+z\\s\\[.*\\s=\\s', '',i)))
N=as.data.frame(apply(N,2,function(i)gsub('\\s', '',i)))

O=as.data.frame(apply(O,2,function(i)gsub('\\s+z\\s\\[.*\\s=\\s', '',i)))
O=as.data.frame(apply(O,2,function(i)gsub('\\s', '',i)))

#Converting the N & O data into numeric values
```

```
N1=c()
N1$Nasal=as.numeric(levels(N$Nasal)[N$Nasal])
O1=c()
O1$Oral=as.numeric(levels(O$Oral)[O$Oral])

NO2=cbind(N1$Nasal, O1$Oral)

#Renaming columns of NO1 to 'Nasal' and 'Oral' objects
colnames(NO2)[2]="Oral"
colnames(NO2)[1]="Nasal"

NO2=as.data.frame(NO2)

#Creating a new column with the length of N & O listed numerically from 1...n
NO2$Frame=(1:nrow(N))

#Ratio:
NO2$Ratio=abs(NO2$Nasal/(NO2$Nasal+NO2$Oral))
#Percentage:
NO2$Percentage=NO2$Ratio*100


#TextGrid manipulation begins here.

#Subsetting the textGrid file to get the min time, max time, and text.
TGMin=subset(textgrid, grepl("            xmin = (|-)(\\d+\\.\\d+|\\d)", textgrid$V1))
TGMax=subset(textgrid, grepl("            xmax = (|-)(\\d+\\.\\d+|\\d)", textgrid$V1))
TGTxt=subset(textgrid, grepl('            text = ', textgrid$V1))

#Combining textGrid subsets
TG=cbind(TGMin, TGMax, TGTxt)

#Creating  an object with the time step between each frame.
dx=subset(oral, grepl("dx\\s=\\s\\d.", oral$Oral))
dx=as.data.frame(apply(dx,2,(function(i)gsub('dx = ', '', i))))
colnames(dx)[1]="dx"
dx=as.data.frame(apply(dx,2,function(i)gsub('\\s', '',i)))
colnames(dx)[1]="dx"
dx=as.vector(dx$dx)[1]
dx=as.numeric(dx)

#Removing the spaces 'xmin', 'xmax', 'text' and '=' signs
```

```
TG1=as.data.frame(apply(TG,2,function(i)gsub('(
    xmin = |              xmax = |              text = )', '',i)
    ))
TG1=as.data.frame(apply(TG1,2,function(i)gsub('\\s', '',i)))

#Renaming the columns
colnames(TG1)[1]="Min"
colnames(TG1)[2]="Max"
colnames(TG1)[3]="Segment"

#Converting the factor data to numeric
TG1$MinNum=as.numeric(levels(TG1$Min)[TG1$Min])
TG1$MaxNum=as.numeric(levels(TG1$Max)[TG1$Max])

#Calculating the average of the minimum and maximum values
TG1$Average=(TG1$MaxNum+TG1$MinNum)/2
TG1$MinFrame=round(TG1$MinNum/dx)
TG1$MaxFrame=round(TG1$MaxNum/dx)
TG1$AvgFrame=round((TG1$MaxFrame+TG1$MinFrame)/2)

#Adding in the annotation boundary markers with a "|".
v=as.data.frame(TG1$AvgFrame)
w=as.data.frame(TG1$MaxFrame)
x=as.data.frame(TG1$Segment)
y=data.frame(a=1:nrow(x),b="|\n|")
colnames(w)[1]="x"
colnames(v)[1]="y"
colnames(x)[1]="z"

#Gathering the frame data that matches the textGrid's
    boundaries with correct frames from the 'Sound' file.
SegmentFrame=as.data.frame(append(v$y,w$x))
colnames(SegmentFrame)[1]="Frame"
Segment=as.data.frame(append(as.vector(x$z),as.vector(y$b)))
colnames(Segment)[1]="Segment"

#Gathering the frame data that matches the textGrid's
    annotations with correct frames from the 'Sound' file.
Text=as.data.frame(cbind(as.vector(SegmentFrame$Frame), as.
    vector(Segment$Segment)))
colnames(Text)[1]="Frame"
colnames(Text)[2]="Segment"

#Dividing the frame values by 1000 to convert the data to
    milliseconds.
Text$FrameV=as.vector(Text$Frame)
Text$dec=(as.numeric(Text$FrameV)/1000)
Order=Text[order(as.numeric(Text$dec)),]
```

```
#Transforming N1 into a data frame.
NO2=as.data.frame(NO2)

#Populates the correct rows with the correct segments
b=0
z=c()
for (i in 1:NROW(NO2)){
b=b+1
i=NO2$Frame[b]
a=match(NO2$Frame==i, Order$FrameV==i)
num=max(a)
z[b]=as.vector(Order$Segment[num])}
NO2$SegmentsAligned=z


#Getting the maximum and minimum values from the textGrid (
    Order) object.
l=min(Order$dec*1000)
m=max(Order$dec*1000)

#Creating an object for the annotation data.
graphdata=subset(NO2, Frame<=m)
graphdata=subset(graphdata, Frame>=l)

#Graphing
#NOTE: If you are using Rstudio, the axes may appear
    distorted until you click the 'zoom' button. When saving
    , we suggest 1600w X 1000h.

#Diving the plot window into one row and one column.
par(mfrow=c(1,1))

#Graph
mar.default <- c(5,4,4,2) + 0.1
par(mar = mar.default + c(0, 4, 0, 0))
plot(NO2$Nasal,type="l", ylim=c(min(NO2$Nasal)+(min(NO2$
    Nasal/4)),max(NO2$Nasal)+(max(NO2$Nasal/4))), col="red",
     xlab="", ylab="", xaxt='n', frame.plot=TRUE)
lines(NO2$Oral, type="l", lty=2, col="blue")
mtext(NO2$SegmentsAligned,side=1, cex=1.4, at=1:nrow(
    graphdata), line=0)
mtext(side = 2, text = "Pascals (Pa)", line = 3, cex=2)
mtext(side = 1, text = "Frames", line = 2, cex=2)
```

Appendix B. Intensity R script

The R script intensity.R is also available at https://scholarspace.manoa.hawaii.edu/bitstream/10125/24724/3/Intensity.R.

```r
#Intensity visualisation in R v3.2.1

#To begin, open each tract in Praat (Boersma and Weenink 2013), select the 'Sound' files, then click the 'To Intensity...' button in the Object's window
#Save each Intensity file separately as a text file: Main menu --> 'Save' --> 'Save as text file...'. Save these files in the same working directory you will use for R.

#Loading intensity text files. Be sure to change address to the correct working directory.
nasal=read.table("%DIRECTORY%/INTENSITYFILENAME_ch1.Intensity", header=T, sep="\t", encoding="UTF-8")
oral=read.table("%DIRECTORY%/INTENSITYFILENAME_ch2.Intensity", header=T, sep="\t", encoding="UTF-8")

#For Windows users using characters other than standard ASCII you may have to change the encoding from UTF-16 to UTF-8.
#In EditPad Lite (Goyvaets, 2015) (free download) click 'Covert' in the main menu --> Text Encoding --> Under 'New Encoding' choose 'Unicode, UTF-8' --> OK. Then save.
textgrid=read.csv("%DIRECTORY%/TEXTGRIDNAME.TextGrid", header=F, encoding="UTF-8")

#Renaming the first column of the 'nasal' and 'oral' objects
colnames(nasal)[1]="Nasal"
colnames(oral)[1]="Oral"

#Subsetting the 'nasal' and 'oral' objects to get the intensity data.
N=subset(nasal, grepl("z\\s\\[.*\\s=\\s\\d.",nasal$Nasal))
O=subset(oral, grepl("z\\s\\[.*\\s=\\s\\d.",oral$Oral))

#Combining 'N' and 'O' objects
NO=cbind(N,O)

#Removing non-numeric data from the 'NO' object.
NO1=as.data.frame(apply(NO,2,function(i)gsub('\\s+z\\s\\[.*\\s=\\s', '',i)))
NO1=as.data.frame(apply(NO1,2,function(i)gsub('\\s', '',i)))

#Renaming columns of NO1 to 'Nasal' and 'Oral' objects
colnames(NO1)[1]="Nasal"
```

```
colnames(NO1)[2]="Oral"

#Converting the NO1 data into numeric values
NO2=c()
NO2$Nasal=as.numeric(levels(NO1$Nasal)[NO1$Nasal])
NO2$Oral=as.numeric(levels(NO1$Oral)[NO1$Oral])

#Creating a new column with the length of NO2 listed numerically from 1...n
NO2$Frame=(1:nrow(NO1))

#Ratio:
NO2$Ratio=NO2$Nasal/(NO2$Nasal+NO2$Oral)
#Percentage:
NO2$Percentage=NO2$Ratio*100

#TextGrid manipulation begins here.

##Subsetting the textGrid file to get the min time, max time, and text.
TGMin=subset(textgrid, grepl("            xmin = (|-)(\\d+\\.\\d+|\\d)", textgrid$V1))
TGMax=subset(textgrid, grepl("            xmax = (|-)(\\d+\\.\\d+|\\d)", textgrid$V1))
TGTxt=subset(textgrid, grepl('            text = ', textgrid$V1))

#Combining textGrid subsets
TG=cbind(TGMin, TGMax, TGTxt)

#Creating  an object with the time step between each frame.
dx=subset(oral, grepl("dx\\s=\\s\\d.", oral$Oral))
dx=as.data.frame(apply(dx,2,(function(i)gsub('dx = ', '', i))))
colnames(dx)[1]="dx"
dx=as.data.frame(apply(dx,2,function(i)gsub('\\s', '',i)))
colnames(dx)[1]="dx"
dx=as.vector(dx$dx)[1]
dx=as.numeric(dx)

#Removing the spaces 'xmin', 'xmax', 'text' and '=' signs
TG1=as.data.frame(apply(TG,2,function(i)gsub('(            xmin = |            xmax = |            text = )', '',i)))
TG1=as.data.frame(apply(TG1,2,function(i)gsub('\\s', '',i)))

#Renaming the columns
colnames(TG1)[1]="Min"
```

```
colnames(TG1)[2]="Max"
colnames(TG1)[3]="Segment"

#Converting the factor data to numeric
TG1$MinNum=as.numeric(levels(TG1$Min)[TG1$Min])
TG1$MaxNum=as.numeric(levels(TG1$Max)[TG1$Max])

#Calculating the average of the minimum and maximum values
TG1$Average=(TG1$MaxNum+TG1$MinNum)/2
TG1$MinFrame=round(TG1$MinNum/dx)
TG1$MaxFrame=round(TG1$MaxNum/dx)
TG1$AvgFrame=round((TG1$MaxFrame+TG1$MinFrame)/2)

#Adding in the annotation boundary markers with a "|".
v=as.data.frame(TG1$AvgFrame)
w=as.data.frame(TG1$MaxFrame)
x=as.data.frame(TG1$Segment)
y=data.frame(a=1:nrow(x),b="|\n|")
colnames(w)[1]="x"
colnames(v)[1]="y"
colnames(x)[1]="z"

#Gathering the frame data that matches the textGrid's
    boundaries with correct frames from the 'Sound' file.
SegmentFrame=as.data.frame(append(v$y,w$x))
colnames(SegmentFrame)[1]="Frame"
Segment=as.data.frame(append(as.vector(x$z),as.vector(y$b)))
colnames(Segment)[1]="Segment"

#Gathering the frame data that matches the textGrid's
    annotations with correct frames from the 'Sound' file.
Text=as.data.frame(cbind(as.vector(SegmentFrame$Frame), as.
    vector(Segment$Segment)))
colnames(Text)[1]="Frame"
colnames(Text)[2]="Segment"

#Dividing the frame values by 1000 to convert the data to
    milliseconds.
Text$FrameV=as.vector(Text$Frame)
Text$dec=(as.numeric(Text$FrameV)/1000)
Order=Text[order(as.numeric(Text$dec)),]

#Transforming N1 into a data frame.
NO2=as.data.frame(NO2)

#Populates the correct rows with the correct segments
b=0
z=c()
```

```
for (i in 1:NROW(NO2)){
b=b+1
i=NO2$Frame[b]
a=match(NO2$Frame==i, Order$FrameV==i)
num=max(a)
z[b]=as.vector(Order$Segment[num])}
NO2$SegmentsAligned=z

#Getting the maximum and minimum values from the textGrid (
    Order) object.
l=min(Order$dec*1000)
m=max(Order$dec*1000)

#Creating an object for the annotation data.
graphdata=subset(NO2, Frame<=m)
graphdata=subset(graphdata, Frame>=l)

#Graphing
#NOTE: If you are using Rstudio, the axes may appear
    distorted until you click the 'zoom' button. When saving
    , we suggest 1600w X 1000h.

#Diving the plot window into one row and one column.
par(mfrow=c(1,1))

#Graph
mar.default <- c(5,4,4,2) + 0.1
par(mar = mar.default + c(0, 4, 0, 0))
plot(NO2$Nasal,type="l", ylim=c(min(NO2$Nasal)-2,max(NO2$
    Nasal)+2), col="red", xlab="", ylab="", xaxt='n', frame.
    plot=TRUE)
lines(NO2$Oral, type="l", lty=2, col="blue")
mtext(NO2$SegmentsAligned,side=1, cex=1.4, at=1:nrow(
    graphdata), line=0)
mtext(side = 2, text = "Decibels (dB)", line = 3, cex=2)
mtext(side = 1, text = "Frames", line = 2, cex=2)

#Ratio
mar.default <- c(5,4,4,2) + 0.1
par(mar = mar.default + c(0, 4, 0, 0))
plot(NO2$Ratio,type="l", ylim=c(0.4, 0.6), col="purple",
    xlab="", ylab="", xaxt='n', frame.plot=TRUE)
mtext(NO2$SegmentsAligned,side=1, cex=1.4, at=1:nrow(
    graphdata), line=0)
mtext(side = 2, text = "Nasal to Oral Ratio", line = 3, cex
    =2)
mtext(side = 1, text = "Frames", line = 2, cex=2)
```