# Nasal aerodynamics and coarticulation in Bininj Kunwok: Smoothing Spline Analysis of Variance

*Hywel Stoakes*[1], *Janet Fletcher*[1], *Andrew Butcher*[2]

[1]The University of Melbourne, Australia
[2]Flinders University, Australia

hstoakes@unimelb.edu.au j.fletcher@unimelb.edu.au andy.butcher@flinders.edu.au

## Abstract

Nasal phonemes are well represented within the lexicon of Bininj Kunwok.[1] This study examines intervocalic, word medial nasals and reports patterns of coarticulation using a Smoothing Spline Analysis of Variance (SSANOVA). This allows for detailed comparisons of peak nasal airflow across six female speakers of the language. Results show that in a VNV sequence there is very little anticipatory vowel nasalisation and greater carryover into a following vowel. The maximum peak nasal flow is delayed for coronals when compared to the onset of oral closure in the nasal, indicating a delayed velum opening gesture. The velar place of articulation is the exception to this pattern with some limited anticipatory nasalisation. The SSANOVA has shown to be an appropriate technique for quantifying these patterns and dynamic speech data in general.

**Index Terms**: nasals, SSANOVA, aerodynamics, Australian languages

## 1. Background

### 1.1. Nasals in Australian languages

Australian languages have sonorant-rich phoneme inventories with nasals at many places of articulation, matched with oral plosives. Many languages contain five [1], six [2] and sometimes seven [3] contrastive nasals, with additional laterals matched with coronals. Across the areal phylum, vowels are not documented as contrasting phonemically based on nasalisation [1]. Previous research across languages generally, shows that coarticulation makes segments less phonemically distinct and in order to keep phonemes phonologically contrastive coarticulation needs to be limited [4], [5]. In Australian languages, due to the large number of place of articulation contrasts, there is a phonological imperative to keep nasals perceptually separated from each other, yet the mechanisms behind this are still understudied [1].

In Australian languages anticipatory nasalisation is thought to be very tightly controlled. The phoneme inventories of Australian languages commonly contain many places of articulation which need to be distinguished acoustically. Phonetic nasalisation in vowels makes it more difficult to discriminate the place of articulation of following nasals [6]. This is because place of articulation cues are often marginal in nasals with acoustic cues found within the transitions between nasal and vowels most salient (see [7] for an overview). The perceptual consequences of these spectral concentrations is that formant transitions at the margins of the nasal rather than the low frequency nasal murmur convey the majority of the place of articulation information (e.g. [6] for American English, and [7] for Catalan). In order to preserve the place of articulation cues in nasals and surrounding segments, a delay in velum lowering would limit the confounding effect of vowel nasalisation on cues that are due to movements in the oral articulators [8]. The delay provides the maximum opportunity for the perception of theses transitional cues found within the speech spectrum within sonorants and vowels ([1], [9]).

In a related vein, previous studies of Australian languages show that Warlpiri [10] and Iwaidja [11], both allow temporal coproduction of an apical nasal with a dorsal stop showing only limited spatial modifications, particularly in apical nasals [9]. In Burarra, Gupapuyŋu, and Warlpiri, anticipatory vowel-consonant coarticulatary resistance exceeds that of carry-over coarticulation [12]. These results suggest that a coarticulatory gesture can be anticipated and controlled by the speaker and may be consciously resisted in order to keep phonemic categories distinct. The current study asks whether, in order to control the extent of anticipatory coarticulation, advanced planning motivated by a need to preserve place of articulation cues is needed to mitigate the masking or loss of crucial spectral or articulatory cues.

### 1.2. Smoothing Spline ANOVA

There has been increased interest in the quantitative analyses of dynamic speech data in recent years. This has led to renewed focus on the acoustic analysis of dynamic formants and fundamental frequency measured over time. The Smoothing Spline Analysis of Variance (SSANOVA) method, introduced by Gu [13] is one method of averaging complex time-series data. There are now a number of phonetic studies utilising this technique for a variety of dynamic speech data. The most prevalent to date have been the analysis of static ultrasound tongue splines [14],[15], [16], [17] and acoustic formant data ([18],[19], [20], [21]) which allow dynamic formant trajectories to be compared across speakers and words. Levels of speaker variation prove challenging to analyse and this technique is promising for comparing articulatory results across speakers.

In this study peak airflow is averaged using a similar spline smoothing algorithm, utilising functions contained within the *gss* package [22]. In order to apply an SSANOVA successfully to speech data, each segment must first be temporally normalised before subsequent statistical analyses are calculated (see Section 2.4). Each phoneme is considered individually with separate confidence intervals. The resulting plot, averaged across speakers, indicates the peak nasal airflow rate plotted across time separately for each of the nasal phonemes.

[1]There was a recent decision to standardise the orthography to use the Kunwinjku conventions, thus *Gun-wok* is now *Kunwok* (see http://bininjgunwok.org.au/information/orthography/)

# 2. Methods

## 2.1. Speakers and Materials

The recordings in were made with six female speakers of Bininj Kunwok (Kunwinjku variety) who repeated a list of disyllabic lexical items each containing intervocalic medial nasals. The lists were compiled by the first and third authors with reference to the Kuninjku Dictionary [23] and the Kunwinjku learners dictionary [24]. The list was then checked and revised by Murray Garde and by the first author in consultation with Bininj Kunwok speakers to ensure both semantic and phonological accuracy.

All words begin with a voiceless velar stop (except *bininj*) and each was uttered within the same carrier phrase (*yun yime X yimen Y*). Each speaker made three repetitions of the word list, although not all recordings were usable due to data capture errors, giving a total of 107 tokens (see Table 1).

Table 1: *Word list and number of tokens (n).*

| | | | | |
|---|---|---|---|---|
| **Word:** | *bininj* | *kamak* | *kangokme* | *kanjok* |
| **Phonetic:** | [ˈpɪnɪɲ] | [ˈkɐmɐk] | [kɐˈŋɔkmɛ] | [ˈ kɐɲɔk] |
| **Gloss:** | ‘male’ | ‘good’ | ‘carry away’ | kin-term |
| *n* | 22 | 11 | 9 | 15 |
| **Word:** | *karnubirr* | *kinga* | *kumoken* | *kunak* |
| **Phonetic:** | [ˈkɐɲˌʊbɪr] | [ˈkɪŋɐ] | [ˈ kʊmɔˌkɛn] | [ˈkʊnɐk] |
| **Gloss:** | f.w. mussel | ‘crocodile’ | ‘f.w. crocodile’ | ‘fire’ |
| *n* | 8 | 13 | 13 | 16 |
| **Total:** | *107* | | | |

## 2.2. Aerodynamic Recordings

This study reports the results from a single peak nasal airflow channel ($U_n$ measured in $\mathrm{cm}^3\,\mathrm{s}^{-1}$). Simultaneous peak oral airflow ($U_o$) was also recorded, although oral airflow data are not reported here. The multichannel articulatory recordings were gathered via Scicon R&D oral and nasal airflow masks with an in-built microphone connected to a Scicon R&D 916 capture device. The airflow acquisition hardware was controlled using the PCQuirer software (Version 7, Scicon R&D California, USA). Calibration was done before and after the equipment was moved to the field site.

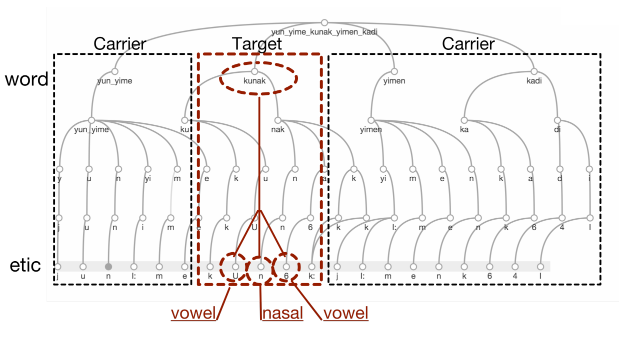


Figure 1: *A example of the hierarchy showing the target word within the carrier phrase*

## 2.3. Labelling and Querying

Labelling and segmentation was done within the Emu WebApp [25] and further analyses used the *EmuR* package [26] within the R programming environment [27]. The acoustic signal was used as the basis of segmentation and the determination of the nasal vowel boundary. The hierarchical querying architecture of Emu is essential in order to restrict the measurements of the target nasal to the word medial intervocalic position ($V_1NV_2$). The following code queries a hierarchy that has both word and phonetic (etic) tiers temporally linked within Emu (see Figure 1).[2]

```
#queries for a VNV sequence
require(emuR)
# data base first loaded using:
BGW_AE_N_2006 <- load_emuDB(databasepath)

VNV.seq <- emuR::query(BGW_AE_N_2006,
   "[[etic = vowel ->
     [etic = nstop & Medial(word,etic) = 1
          ^ word =~ .*]] ->
      etic = vowel]",
timeRefSegmentLevel = "etic",
resultType = "emuRsegs"
                         )
```

This gives an *R* vector (an *Emu* segment list) containing the results for a word medial sequence of a medial nasal surrounded by two vowels. This can be refined to remove the carrier phrase tokens. We then return an individual segment list for each of the items in turn by specifying the target segment (see the documentation for the *Emu Query Language Version 2* for details). This gives four parallel vectors, one containing data for *V1*, one containing data for *N*, one containing data for *V2* and one that encompasses the entire word, used in subsequent analysis. The airflow channels are then extracted using the `emuR::get_trackdata` function providing an *R* data frame.

## 2.4. Normalisation

We use a similar aerodynamic measurement methodology to that reported by [28] for French nasal sequences in that the nasal flow is averaged for all speakers for each individual phoneme and then the sequence is then reconstructed in temporal order. The airflow ($U_n$) averaging is achieved by first, time normalising the signal and subsequently averaging the airflow for each segment separately which gives an average peak flow over time ($U_n$) [28, pp 594–5]. This method shows the absolute timing of dynamic changes in airflow. The flow magnitude information, however–as it is an average across speakers–is less valuable. A smoothing spline ANOVA is then calculated (see Section 2.5 below for the method). Each token has had the zero-offset, adjusted as over the course of a recording session the zero flow level drifted either upward or downward. The minimum value in $V_1$ was measured and used as the zero value for the entire sequence which was then then used to normalise the airflow values in $N$ and $V_2$.

## 2.5. SSANOVA

The process for calculating the SSANOVA closely follows the method introduced by Fruewald [19] who compared dynamic formant trajectories (F1 and F2). In the current study an SSANOVA (`gss::ssanova()`) is calculated using the *gss* package and subsequently the `stats::predict()` function which makes a prediction for each point based on the model (`fit`). The corresponding standard error is also calculated (`se.fit`). These

[2] Anonymised data can be accessed at http://hywel.github.io/data/df_VNV_V1.csv (1Mb)

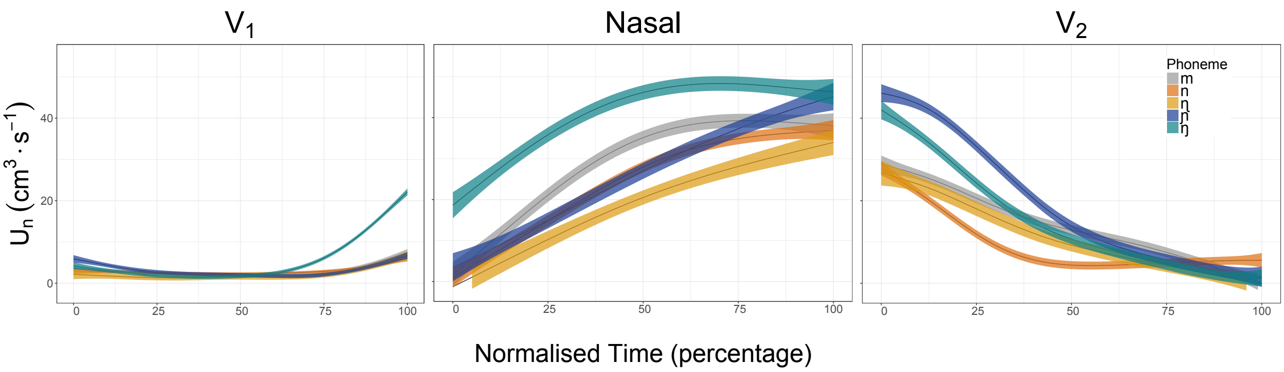


Figure 2: *An SSANOVA of average nasal airflow by normalised time in Vowel$_1$, Nasal, Vowel$_2$ sequences separated by phoneme*

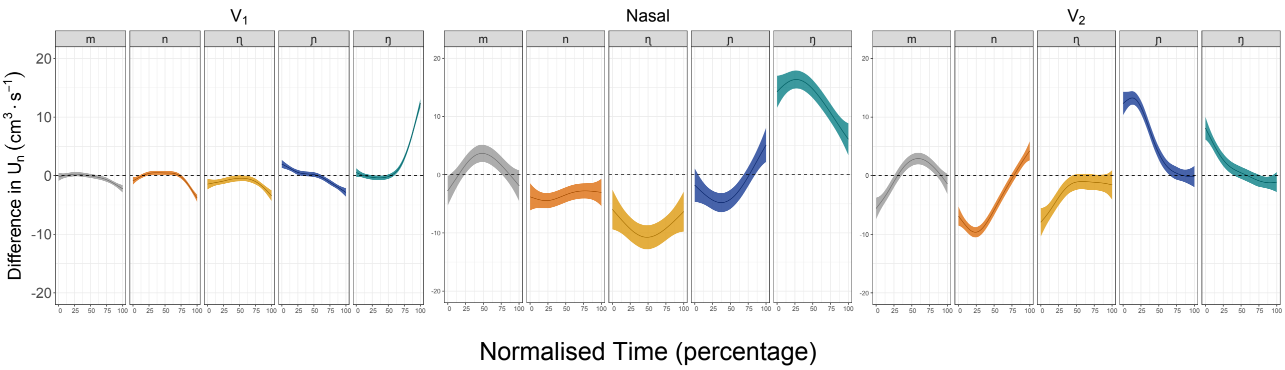


Figure 3: *The group plus interaction over normalised time between the phonemes (nasal:time) in Vowel$_1$, Nasal, Vowel$_2$ sequences*

are returned individually for each nasal phoneme (`nasal`) and results are then plotted using the *ggplot2* package. The *R* [27] code below produces a vector allowing plotting of the first panel (Vowel 1) shown in Figure 2.

```
df.VNV.V1 <-read.csv
          (file = "df.VNV.V1.csv")
require(gss)
Un.VNV.V1.model <-
  ssanova(data~nasal + time + nasal:time,
          data = df.VNV.V1)
grid.VNV.V1 <-
  expand.grid(time = seq(0,1,length = 100),
  nasal = c("m","n","ɳ ","ɲ","ŋ"))
grid.VNV.V1$Un.Fit <-
  predict(Un.VNV.V1.model,
           data_n = grid.VNV.V1,
           se = T)$fit
grid.VNV.V1$Un.SE <-
  predict(Un.VNV.V1.model,
          data_n = grid.VNV.V1,
          se = T)$se.fit
```

This generates the data frame for the first panel of the plot in Figure 2. The group and time interaction is then shown in Figure 3 indicating the difference in airflow over time for each phoneme.

# 3. Results

## 3.1. Smoothed nasal airflow over time

The figures 2 and 3 report the results from the SSANOVA. Figure 2 shows the SSANOVA for each phoneme plotted across 100 sample points as a percentage. The standard error (se) is shown as a ribbon in Figure 2) and is calculated using a 95% Bayesian confidence interval. The discontinuities at the edges of the panels are due to the averaging of the $U_n$ signal over time and minor perturbations in the flow signal. Figure 3 shows the group interaction between the phonemes over time (nasal:time). When the plots intersect the zero line it indicates that the phonemes are not significantly different at that timepoint.

Results show very little anticipatory nasalisation in a vowel preceding a word medial nasal for all phonemes except the velar. The velar nasal shows an increased peak nasal airflow starting at 65% of the initial vowel ($V_1$). During the peripheral nasals /m/ and /ŋ/ have peak nasal airflow that occurs before that of the other phonemes with /ŋ/ just after 25% into the nasal and /m/ just prior to 50%. The coronal consonants all have their maximum peak of nasalisation at the acoustic offset of the nasal (centre panel of figure 2). Figure 3 shows that for velars (/ŋ/)the difference in flow is greater in $V_1$ and $N$ than the other phonemes. The palatal (/ɲ/) has a higher peak flow than the each of the other phonemes in the second vowel ($V_2$)indicating that it has the highest carryover nasalisation. This carryover effect may be due to the greater contact area of the laminal articulator meaning that coordination between oral closure and nasalisation is more difficult to maintain. In velars, velum lowering is less delayed because, unlike with coronal articulations, the velum needs to be lowered in order to make closure with the tongue dorsum during the articulation of the nasal itself.

## 4. Discussion

This study shows that anticipatory coarticulation of vowel nasalisation is tightly controlled in medial $VN$ sequences by Bininj Kunwok speakers. These patterns of nasal flow are interpreted as evidence of delayed velum lowering during the pre-nasal vowel. It is clear from the results that carryover nasalisation is not controlled in the same manner and that the peak of nasalisation is at the offset of oral closure for the coronal nasals. The variation in the location of the peak nasal flow suggests that there are physical differences between the articulation of these phonemes although this is not thought to be at the level of awareness. The tight control of velum lowering may be used as a strategy to ensure that place of articulation information is phonetically retrievable in an environment that can obscure place of articulation cues. This equates very well with qualitative examinations of acoustic signals in Bininj Kunwok, suggesting that the SSANOVA technique is appropriate for the analyses of complex time-course data. Further work will look at the duration and timing of the both the opening and closing phase in the language.

## 5. Acknowledgements

The authors would like to thank the speakers from the Mamardawerre community, Deborah Loakes and Murray Garde, as well as two anonymous reviewers. All errors remain our own. This research is supported by the ARC Centre of Excellence for the Dynamics of Language.

## 6. References

[1] Butcher, A. R., "Consonant-salient phonologies and the 'place-of-articulation imperative'," in Speech Production: Models, Phonetic Processes and Techniques, Harrington, J. and Tabain, M., Eds., New York: Psychology Press, 2006, 187–210.

[2] Tabain, M. and Butcher, A. R., "Stop consonants in Yanyuwa and Yindjibarndi: Locus equation data," Journal of Phonetics, 27 (1), 333–357, 1999.

[3] Breen, G., "The wonders of Arandic phonology," in Forty Years On: Ken Hale and Australian Languages, Simpson, J., Nash, D., and Laughren, M., Eds., Canberra: Pacific Linguistics, 2001, 45–69.

[4] Manuel, S. Y., "The role of contrast in limiting vowel-to-vowel coarticulation in different languages," The Journal of the Acoustical Society of America, 88 (3), 1286–1298, 1990.

[5] Scarborough, R., Zellou, G., Mirzayan, A., and Rood, D. S., "Phonetic and phonological patterns of nasality in Lakota vowels," Journal of the International Phonetic Association, 45 (03), 289–309, 2015.

[6] Miller, G. and Nicely, P., "An analysis of perceptual confusions among some English consonants," The Journal of the Acoustical Society of America, 27, 338, 1955.

[7] Recasens, D., "Place cues for nasal consonants with special reference to Catalan," Journal of the Acoustical Society of America, 73 (4), 1346–1353, 1983.

[8] Butcher, A. R., "What speakers of Australian Aboriginal languages do with their velums and why: The phonetics of the nasal/oral contrast," in XIVth International Conference of Phonetic Science, San Francisco, 1999.

[9] Fletcher, J., Butcher, A., Loakes, D., and Stoakes, H., "Aspects of nasal realization and the place of articulation imperative in Bininj Gun-wok," in Proceedings of the 13th Speech Science and Technology Conference, Melbourne, ASSTA, 2010.

[10] Fletcher, J., Loakes, D., and Butcher, A., "Coarticulation in nasal and lateral clusters in Warlpiri," in Annual Conference- International Speech Communication Association, ser. International Speech Communication Association, 1, 2009, 86–89.

[11] Fletcher, J. M., Butcher, A. R., Loakes, D., and Stoakes, H., "Coarticulation and consonant cluster production in Iwaidja," The Journal of the Acoustical Society of America, 129 (4), 2452–2452, 2011.

[12] Graetzer, S., Fletcher, J., and Hajek, J., "Locus equations and coarticulation in three Australian languages," The Journal of the Acoustical Society of America, 137 (2), 806–821, 2015.

[13] Gu, C. and Wahba, G., "Smoothing spline ANOVA with component-wise Bayesian "confidence intervals"," Journal of Computational and Graphical Statistics, 2 (1), 97–117, 1993.

[14] Davidson, L., "Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance," The Journal of the Acoustical Society of America, 120 (1), 407–415, 2006.

[15] Billington, R., "'Advanced tongue root' in Lopit: Acoustic and ultrasound evidence," in Proceedings of the 15th Australasian International Speech Science and Technology Conference, Hay, J. and Parnell, E., Eds., ASSTA, 2014.

[16] Heyne, M. and Derrick, D., "Using a radial ultrasound probes virtual origin to compute midsagittal smoothing splines in polar coordinates," The Journal of the Acoustical Society of America, 138 (6), EL509–EL514, 2015. [Online]. Available: http://scitation.aip.org/content/asa/journal/jasa/138/6/10.1121/1.4937168.

[17] Mielke, J., "An ultrasound study of Canadian French rhotic vowels with polar smoothing spline comparisons," The Journal of the Acoustical Society of America, 137 (5), 2858–2869, 2015.

[18] De Decker, P. and Nycz, J., "A new way of analyzing vowels: Comparing formant contours using smoothing spline ANOVA," in Proceedings of thet the 35th NWAV Conference, 2006.

[19] Fruehwald, J., "SS ANOVA," Academia.edu, 2010. [Online]. Available: https://www.academia.edu/268789/SS_ANOVA.

[20] Haddican, B., Foulkes, P., Hughes, V., and Richards, H., "Interaction of social and linguistic constraints on two vowel changes in northern England," Language Variation and Change, 25 (03), 371–403, 2013.

[21] Docherty, G., Gonzalez, S., and Mitchell, N., "Static vs. dynamic perspectives on the realization of vowel nuclei in West Australian English," in 18th International Congress of Phonetic Sciences. Glasgow, Scotland, 2015.

[22] Gu, C., "Smoothing spline ANOVA models: R package gss," Journal of Statistical Software, 58 (5), 1–25, 2014. [Online]. Available: http://www.jstatsoft.org/v58/i05/.

[23] Garde, M., Bininj Kunwok Dictionary. ANU, forthcoming.

[24] Manakgu, A. and Etherington, S., Basic Kunwinjku Dictionary: A simplified English-Kunwinjku and Kunwinjku-English dictionary. The Kunwinjku Language Centre, 1996.

[25] Winkelmann, R. and Raess, G., "Introducing a web application for labeling, visualizing speech and correcting derived speech signals," in Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), Chair), N. C. (, Choukri, K., Declerck, T., Loftsson, H., Maegaard, B., Mariani, J., Moreno, A., Odijk, J., and Piperidis, S., Eds., Reykjavik, Iceland: European Language Resources Association (ELRA), May 2014, 26–31, ISBN: 978-2-9517408-8-4.

[26] Winkelmann, R., Jaensch, K., Cassidy, S., and Harrington, J., EmuR: Main package of the Emu Speech Database Management System, R package version 0.1.8.9001, IPS Munich. [Online]. Available: https://github.com/IPS-LMU/emuR.

[27] R Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2016. [Online]. Available: https://www.R-project.org/.

[28] Delvaux, V., Demolin, D., Harmegnies, B., and Soquet, A., "The aerodynamics of nasalization in French," Journal of Phonetics, 36 (4), 578–606, 2008.