ACOUSTIC STUDY OF THE NEPALI NASAL CONSONANTS

Krishna Prasad Chalise

Central Department of Linguistics, Tribhuvan University
krishnaprasadchalise@gmail.com

The Nepali nasals have been characterized based on duration, formant structure, Cog, and SD. /n/ has the longest duration, followed by /m/ and /ŋ/. They are longer with homorganic vowels. The duration of each place of articulation remains constant across different vowel environments. N1 value increases from /m/ to /ŋ/ and N3 and N4 nearly follow the same pattern. BW1 is the largest for /ŋ/ followed by /m/ and /n/. Cog and SD do not correlate with place of articulation; rather, they correlate with the vowel environment. Similarly, amplitude does not distinguish a nasal from the adjacent vowel in all vowel environments.

Keywords: formant, anti-formant, bandwidth, spectral moment, nasopharyngeal cavity

1. Introduction

1.1 Background

The early works that describe Nepali phonology, like Ayton (1820), state that there exist five nasal phonemes: /m/, /n/, /ɲ/, /ɳ/ and /ŋ/. It is because of the Devanagari writing system that Nepali uses. In the writing system, there are five nasal letters and the early descriptions of Nepali phonology described the letters as the phonemes based on the Sanskrit loans in Nepali. But the spoken Nepali has only three nasal consonants: bilabial [m], alveolar [n], and velar [ŋ] (Clark, 1963; Bandhu et al., 1971; Bandhu, 1979; Pokharel, 1989 and Khatiwada, 2009). Dahal (1974) includes /m^h/ and /n^h/ as the marginal nasal phonemes in addition to /m/, /n/ and /ŋ/. Pokharel (1989) has a long justification about how [ɳ] is not a phoneme in Nepali, but Genetti (1994) accepts /ɳ/ as a phoneme. She elaborates that it is mainly found in the Sanskrit loans as well as in a limited number of native words. However, the findings of Khatiwada (2009) support Pokharel (1989) and, now, it is a well-accepted fact that /m/, /n/, and /ŋ/ are the nasal consonant phonemes in Nepali.

There are no varied ideas about the point of articulation of /m/ and /ŋ/ since the first one is bilabial and the second one is velar (dorso-velar), but there are varied ideas regarding the place of articulation of /n/. Dahal (1974) regards /n/, basically, to be a dental sound, and he explains that it is produced alveolar before fricatives, affricates, velars, and glides, and it is produced retroflex before the retroflex sounds. Khatiwada (2009) supports Dahal (1974). But, Bandhu (1979) supposes it to be an apico-alveolar sound which is the same as the finding of Pokharel (1989). Pokharel further explains that [n] is apical thus it is closer to the Nepali retroflex sounds [ʈ] and [ɖ] rather than the dental sounds [t̪] and [d̪].

The distribution of [m] and [n] is in both onset and coda positions in a syllable whereas [ŋ] occurs only in the coda position except in a limited number of onomatopoeic words like [ŋjarrʌ] (sound produced by a cat when angry), [ŋjitstsʌ] (the state of appearance of a person when s/he can't get something expected).

1.2 Articulation and acoustics of nasals

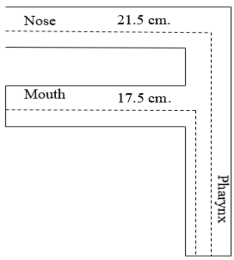
The nasals are produced with complete closure at a point in the oral cavity with lowered velum allowing the airflow through the nasal cavity. Therefore, there are two resonating cavities, i.e., the nasopharyngeal cavity (the combined cavity formed by the pharyngeal cavity and the nasal cavity) and the oral cavity, during the production of a nasal consonant. The approximate articulatory configuration of [m] is presented in Figure 1.

The cavities are combined in a complex way: the oral cavity is a side branch of the nasopharyngeal cavity open into it (the nasopharyngeal cavity) and closed at another end. The nasals bear some characteristics of vowels as well as stop consonants. Like vowels, they are produced by the airflow through the vocal folds and have well-defined patterns of formants that differ according to the place of articulation. Similarly, they are similar to the plosive consonants in articulation.

Nepalese Linguistics, vol. 37(1), 2023, pp. 1-9. DOI: https://doi.org/10.3126/nl.v37i1.60006

The only difference is that the nasal port is open during the production of the nasals whereas it is closed during the production of the plosive consonants (Hayard, 2013).

Figure 1: The articulatory configuration of [m]



Fujimura (1962), Fant (1970), Stevens (2000), and Johnson (2003) have explained how the nasal consonants are produced. According to them, the nasals are produced with complete closure at a point in the oral cavity with lowered velum allowing airflow through the nasal cavity. So, there are two resonating cavities, i.e., the nasopharyngeal cavity (the combined cavity formed by the pharyngeal cavity and the nasal cavity) and the oral cavity, during the production of a nasal consonant.

The common acoustic characteristics investigated by the above-mentioned studies are that the nasals are characterized by the low first formant in the range of 200-300 Hz with wide bandwidth and the anti-formant produced by the coupling between the main resonating nasopharyngeal cavity, the side branch, and the oral cavity. The frequency of the anti-formant depends on the place of articulation of the nasal because different places of articulation determine the different lengths of resonating tube in the oral cavity.

Fujimura (1962) assumes the entire articulatory system of nasal production as the combination of the pharynx extending from the glottis to the velum, the oral cavity, and the nasal tract including the nasopharynx and nasal passages that are terminated by radiation impedance. In his calculation, the first nasal formant (N1) is in the range of 250-400 Hz. The second formant (N2) is in the range of 900-1150 Hz. The third formant (N3) is in the range of 1700-2200 Hz and the fourth nasal formant (N4) is in the range of 2300-3000 Hz. But, in the production of /m/ and /n/, the first anti-formant (Z1) falls within the range of 750-1250 Hz, for /n/, the Z1 falls within the range of 1450-2200 Hz, and for /ŋ/ the Z1 falls above 3000 Hz.

Stevens (2000) takes the production of the uvular nasal /ɴ/ as the special case for nasal modeling. In this mechanism, for a male adult speaker, the N1 is in the range of 250-300 Hz, N2 is around 1000 Hz, N3 is around 2200 Hz and N4 is around 2800-3000 Hz. For /m/, the Z1 is in the range of 1000-1200 Hz, for /n/, the Z1 is in the range of 1600-1900 Hz, and for /ŋ/, the Z1 is in the range of 2900-3200 Hz.

Johnson (2003) takes the production of /ɴ/ for the nasal modeling. In his calculation, N1 is around 407 Hz, N2 is around 1221 Hz, N3 is around 2035 Hz and N4 is around 2849 Hz. Similarly, the Z1 for /m/ is around 1094 Hz, and /n/ is around 1591 Hz.

Based on the study of nasals in English, Polish, Czech, German, Hungarian, Russian, and Swedish, Recasens (1983) shows that the frequency of N1 is around 200–300 Hz, increases gradually from the bilabial to velar places of articulations and the pattern is the same in the case of Z1. In his calculation for Catalan, (he did not study /m/) he found out the Z1 values for /n/ and /ŋ/ were 1780 Hz and 3700 Hz respectively.

Tabain et al. (2016) found out that N1 ranged from 300-350 Hz and it was the lowest for /m/ and highest for /ŋ/. /m/ had all the lowest nasal formants (N1, N2, N3, and N4), /ŋ/ had the highest N1 but low N2, N3, and N4), and /n/ had intermediate values.

The bandwidth (BW) of the nasal formants also have been regarded to characterize the nasals. Recasens (1983) points out that the bandwidth of the first nasal formant is much greater for the velar than for any of the other three nasal places of articulation. Johnson (2012, p. 185) says that the bandwidths of the nasal formants are wider than that of the vowels. Tabain et al. (2016) found out that /ŋ/ had the largest bandwidth values followed by /m/ and /n/ respectively, but the bandwidth of the first nasal formant (BW1). values for all the nasal consonants were found to be similar except for /ŋ/. Tabain (2019) supports the idea and states that BW1 is about 100 Hz, and for higher nasal formants, the bandwidths are around 300 Hz.

The production of the nasal consonants is exactly the same as the production of the plosive consonants except for the opening of the velopharyngeal port. So, the pattern of nasal-to-vowel and vowel-to-nasal formant transition is also similar to the formant transition in the plosive-to-vowel and vowel-to-plosive production (Tabain, 2019, p. 273).

Tabain et al. (2016) observe that the first two spectral moments, Centre of Gravity (CoG) and Standard Deviation (SD)] are useful cues to characterize the nasals as the values are lower for /m/ and /ŋ/ but distinctly higher for /n/.

A study of the Nepali nasals was carried out by Pokharel (1989). His study was based on the five tokens for /m/ and /n/ each and two tokens for /ŋ/, and he found out N1 was found below 250 Hz for all of them and among them the value of nasal murmur was lowest for /m/ followed by /n/ and /ŋ/ respectively. In terms of length of transition, their relation was /ŋ/>/n/>/m/.

2. Methodology

Disyllabic words with target phonemes in word-initial and intervocalic positions as presented in Table 1 were produced by a male and a female speaker.

Table 1: Words for the analysis of nasals

Nasals	i-i		a-a		u-u	
/m/	*t̪imi*	'you'	*t̪ama*	'copper'	*kumu*	'proper name'
/n/	*t̪ini*	'they'	*t̪ana*	'melody'	*bunu*	'sister'
/ŋ/	*siŋe*	'with horn'	*ʈaŋa*	' horse cart'	*suŋu*	'let's smell'

The words were produced in the carrier sentence: X; mʌile X bʰʌnẽ [X; I said X (where X is the target word)]. The corner vowels /i/, /a/, and /u/ were taken as the vowels in the environments. The utterances were recorded in a quiet room using an EDIROL R09HR digital audio recorder and an EDIROL CS-15 by Roland Cardioid Condenser Microphone. There were all together [3 (number of sounds) × 3 (number of environments) × 4 (tokens) × 2 (produced as a word and an utterance) × 2 (number of speakers)] 144 tokens for analysis. There were 48 tokens for each of the nasal sounds. The audio files were in .WAV format with a sample rate of 44100 Hz, a bit rate of 1411, and a 24-bit resolution. Praat 6.1.40 was used for spectral and durational analysis, with standard settings unless otherwise stated. To measure the duration, the portion between the sound's onset and offset in the wideband spectrogram was manually selected. The frequency range was between 0-5000 Hz and the window length was 0.0043 seconds. The boundaries were determined by comparing the intensity with adjacent vowels and formant transition.

Spectral characterization is basically motivated by Tabain et al. (2016). I have characterized them in terms of the first four nasal formants (N1, N2, N3, and N4), the bandwidth of the first nasal formant, the center of gravity (CoG), and the standard deviation (SD). The nasals have more than 5 formants within the range of 0-5000 Hz so 6 formats were measured within the range of 0-5000 Hz for the male speakers and 0-5500 Hz for the female speakers. The nasal formants and formant bandwidth were measured manually selecting the 50 ms portion from the steady state. The LPC spectra were obtained extracting 50 ms portion from the steady state part of the vowel and the extracted sound was resampled into 12500 Hz. Then the spectra were extracted using the LPC (burg) method.

The center of gravity (CoG) and standard deviation (SD) were calculated within the frequency range of 1000-5000 Hz using the filter (pass Hann band) option in Praat. After filtering the frequency ranges, I extracted the LPC spectrum (LPC burg) and calculated the spectral moments. Tabain et al. (2016) point out that "this particular range was chosen because the range below 1000 Hz is dominated by voicing, and is also thought to be very similar across nasal place of articulation".

3. Findings

The nasal consonants were characterized based on the features of duration, the structure of the nasal formants (N1, N2, N3, and N4), the bandwidth of the first nasal formant (BW1), and the first two spectral moments center of gravity (CoG) and standard deviation (SD).

3.1 Duration

The average duration for /m/, /n/, and /ŋ/ is 84 ms (SD = 10.39), 72 ms (SD = 11.47) and 90 ms (SD = 13.17) respectively. In terms of duration, /n/ has the shortest /ŋ/ has the longest duration, and /m/ has an intermediate between them. Repeated measure one-way ANOVA was conducted to find out the significance of differences in duration of the nasals which shows the differences among them are significant [$F(2, 94) = 34.74$, $p < 0.0001$]. The pattern of duration for nasals follows the just opposite pattern of duration for plosives. In both of the classes, the velar is the longest followed by the bilabial and dental/alveolar respectively. Although the pattern remains the same, the nasal consonants are relatively longer in the environment of homorganic vowels. In this case, /m/ is bilabial and there is no vowel produced from that position but /n/ is relatively longer in the i-i environment, and /ŋ/ is relatively longer in the environment of u-u and a-a.

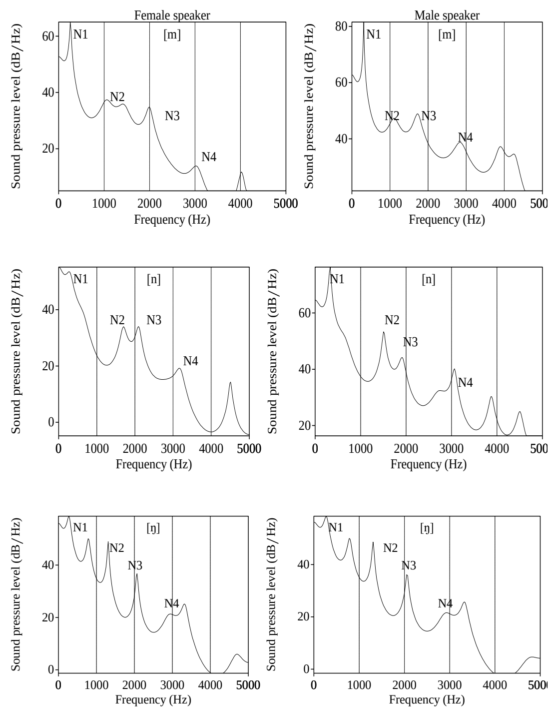
3.2 Formant structure

The formant structure of the nasal consonants is determined by the resonance of the nasopharyngeal tube. So, they have very similar formant structure as presented in Figure 5.1. The difference among them is the result of the nature of the oral cavity resonates which couples with the nasopharyngeal resonance.

In Figure 5.1, the LPC spectra for [m], [n] and [ŋ] produced by the female speaker and the male speaker are presented which shows that all of them have very low N1. N2 is around 1000 Hz, but in some cases, an extra formant is seen below it and, generally, it is not seen in the automatic formant tracking. This is the second nasal formant as explained by Chen (1997). N3 is around 2000 Hz and N4 is around 3000 Hz. But the formant location depends on the place of articulation of the nasal consonant and this variation in the formant location is useful to characterize the different places of articulation.

In this study, I measured the four nasal formants for each of the nasal consonants in three different vowel environments, viz. i-i, a-a and u-u. Certainly, the vowel environment influences the formant structure of the sounds. So, the researcher produced the consonants in isolation 15 times for each of the consonants. So, there were 15 tokens for each of the consonants which provided the formant structure without any influence of the environment. I compared the formant values obtained from two techniques for further justification.

Figure 2: LPC Spectra of /m/, /n/, and /ŋ/ by the Female and Male Speakers



Note. In the LPC spectrum, there is seen an additional formant between N1 and N2 in some cases. This formant is generally not identified by the automatic formant tracker in the spectrogram because it combines with the N1.

3.2.1 The first nasal formant (N1)

From the observation of the LPC spectrograms in Figure 5.1, I understand that the nasal consonants have very low N1 values. In this study, I found that /m/ has the lowest N1 value (M = 308 Hz, SD = 29.51) followed by /n/ (M = 325 Hz, SD =43.88) and /ŋ/ (M = 349 Hz, SD = 49.27). One-way ANOVA was conducted to find out the significance of the distinction which shows that the means are significantly distinct [$F(2, 94) = 23.33$, $p < 0.0001$]. Tukey HSD test shows that the three means are significantly distinct from each other ($p < 0.01$). This pattern is exactly found in the N1 values measured from the tokens produced in isolation. In isolated production, the N1 values for /m/, /n/, and /ŋ/ are 305 Hz (SD = 6.13), 310 Hz (SD = 7.62), and 323 Hz (SD = 6.81) respectively. The pattern that the N1 value increases from /m/ to /ŋ/ is justified in this case, too.

The value of N1 for the nasal consonants is affected by the adjacent vowel. It is lower with the vowels that have low F1 and higher with the vowels that have high F1 values. In our experiment, the N1 value is lowest in the i-i environment followed by the values in the u-u and a-a environments. The Tukey HSD test shows that the values in i-i and u-u environments are not significantly distinct but the value in the a-a environment is significantly distinct ($p < 0.01$) from the other values. It shows that the N1 value is lower in the environment of high vowels and higher in the environment of low vowels.

3.2.2 The band width of the N1 (BW1)

The BW1 seems to be one acoustic cue to characterize the place of articulation of the nasal consonants. In our study, the velar nasal /ŋ/ has the highest BW1 value (M = 108 Hz, SD = 60.21) followed by /m/ (M = 108 Hz, SD = 60.21) and /n/ (M = 70 Hz, SD = 44.55). There is a slight difference in BW1 between /m/ and /ŋ/ but it is vastly different for /n/ and the difference is significant [$F (2, 94) = 12.71$, $p < 0.0001$]. The post hoc test, Tukey HSD test, shows that /n/ is significantly distinct ($p < 0.01$) from /m/ and /ŋ/ regarding the value of BW1.

This pattern is exactly found in the BW1 values measured from the tokens produced in isolation. In isolated production, the BW1 values for /m/, /n/, and /ŋ/ are 58 Hz (SD = 14.52), 48 Hz (SD = 15.60), and 72 Hz (SD = 21.74) respectively. The pattern that the BW1 value for /ŋ/ is the highest and it is followed by the value for /m/ and /n/ respectively is justified in this case, too.

The BW1 value is influenced by the height of the adjacent vowel. It is higher if the adjacent vowel is a low vowel and it is lower if the adjacent vowel is a high vowel. I found that it is the highest in the a-a environment (M = 160 Hz, SD = 44.55). It is relatively very low in the u-u environment (M = 75 Hz, SD = 43.85) and the i-i environment (M = 55 Hz, SD = 25.58). The distinction between the low vowel environment and the high vowel environment is significant [$F (2, 94) = 76.37$, $p < 0.0001$]. The post hoc test shows that the distinction between the BW1 values in the low vowel and high vowel environments is significant ($p < 0.01$).

3.2.3 The second nasal formant (N2)

From the observation of the LPC spectra in Figure 5.1, I understand that N2 lies in the region above 1000 Hz. I found that /m/ has the lowest N2 value (M = 1282 Hz, SD = 315.33) followed but /ŋ/ (M = 1408 Hz, SD =295.73) and /n/ (M = 1444 Hz, SD = 307.67). The second formant value is regarded to represent the frontness/backness of the tongue so in the sonorants produced from the alveolar region the second formant value is always higher and the same pattern was found in the case of nasal consonants, too. One-way ANOVA was conducted to find out the significance of the distinction which shows that the means are significantly distinct [$F (2, 94) = 9.64$, $p = 0.0001$]. Tukey HSD test shows that the N2 for /n/ is significantly distinct ($p < 0.01$). Similarly, in isolated production, the N2 value for /n/ is the highest (M = 1555 Hz, SD = 131.65) but it is lowest for /ŋ/ (M = 1125 Hz, SD = 120.34) and it is between them for /m/ (M = 1174 Hz, SD = 111.45).

The N2 value is also influenced by the height of the adjacent vowel. As in N1, it is higher if the adjacent vowel is a low vowel and it is lower if the adjacent vowel is a high vowel. I found that it is the highest in the a-a environment (M = 1481 Hz, SD = 187.59). It is relatively lower in the u-u environment (M = 1361 Hz, SD = 544.26) and lowest in the i-i environment (M = 1292 Hz, SD = 352.22). The distinction between the low vowel environment and the high vowel environment is significant [$F (2, 94) = 9.43$, $p = 0.0001$]. The post hoc test shows that the distinction between the N2 values in the low vowel and high vowel environments is significant ($p < 0.01$).

3.2.4 The third nasal formant (N3)

As in the LPC spectra in Figure 5.1, I found that N3 lies in the frequency region around 2000 Hz. I found that the pattern of N3 is like the pattern of

N1 as it increases from the bilabial to the velar nasal consonants. /m/ has the lowest N3 value (M = 2057 Hz, SD = 185.88) followed by /n/ (M = 2166 Hz, SD = 110.74) and /ŋ/ (M = 2237 Hz, SD = 88.25) respectively. One-way ANOVA was conducted to find out the significance of the distinction which shows that the means are significantly distinct [F (2, 94) = 28.17, $p < 0.0001$]. Tukey HSD test shows that the N3 values for all of the nasal consonants are significantly distinct ($p < 0.01$) from each other. Similarly, in isolated production, the pattern is the same. The pattern is also the same in the isolated production.

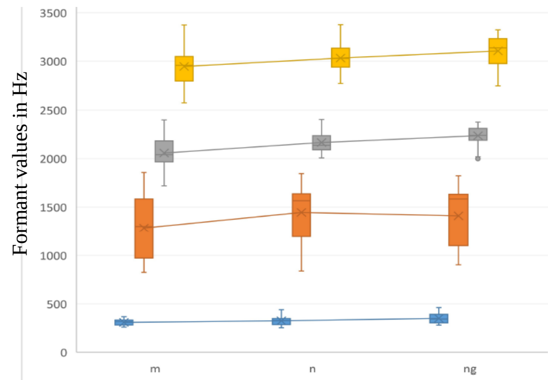
The N3 value, too, is influenced by the height of the adjacent vowel. It is higher if the adjacent vowel is a high vowel and lower if the adjacent vowel is a low vowel. I found that it is the highest in the u-u environment (M = 2198 Hz, SD = 121.08). It is slightly lower in the i-i environment (M = 2180 Hz, SD = 185.85) and lowest in the a-a environment (M = 2082 Hz, SD = 119.87). The distinction between the low vowel environment and the high vowel environment is significant [F (2, 94) = 9.43, p = 0.0001]. The post hoc test shows that the distinction is significant ($p < 0.01$) between the low and high vowel environments.

3.2.5 The fourth nasal formant (N4)

I found that N4 lies in the frequency region around 3000 Hz. The pattern of N4 is like the pattern of N1 and N3 as it increases from the bilabial to the velar nasal consonants. /m/ has the lowest N4 value (M = 2944 Hz, SD = 211.83) followed by /n/ (M = 3032 Hz, SD = 133.16) and /ŋ/ (M = 3104 Hz, SD = 151.26) respectively. One-way ANOVA was conducted to find out the significance of the distinction which shows that the means are significantly distinct [F (2, 94) = 28.17, $p < 0.0001$]. Tukey HSD test shows that the N4 values for all of the nasal consonants are significantly distinct ($p < 0.01$) from each other. Similarly, in isolated production, the pattern is the same. The N4 value does not seem to be affected by the feature of the adjacent vowel.

The overall formant structure for the nasal consonants in Nepali is presented in Figure 2, which shows that the nasal formant values, except N2 which is highest for the alveolar nasal, gradually increase from the bilabial nasal to the velar nasal. The average space between N1 and N2 is 1050 Hz, between N2 and N3 is 775 Hz; and between N3 and N4 is 873 Hz.

Figure 3: Overall formant structure of the nasals



Note. The N1, N2, N3 and N4 are from bottom to up

3.3 Spectral moments (CoG) and SD)

Tabain et al. (2016) found that the first two spectral moments, the Center of Gravity (CoG) and standard deviation (SD) were very prominent features to distinguish the places of articulation of the nasal consonants. They measured them within the frequency range of 1000-5000 Hz. I extracted 50 ms selection from the middle part of the nasal consonants and filtered 1000-5000 Hz frequency range using the pass Hann band filter using Praat and then I got the LPC (burg) spectra. Then, I measured the CoG and SD values at power 2 using Praat. However, in our findings, both of them were completely insignificant to characterize the places of articulation of the nasal consonants. The CoG and SD values for all of the nasal consonants are approximately 1900 Hz and 600 respectively. However, both of the values are significantly distinct for the vowel environments in which the nasal consonants occur. So, our finding is that the spectral moments are not useful cues to characterize the places of articulation of the nasal consonants.

4. Discussion

The pattern of duration is /n/ (dental) < /m/ (bilabial) < /ŋ/ (velar) which is just opposite to the duration pattern of the plosive sounds (velar < bilabial < dental). /n/ and /ŋ/ are relatively longer in the environment of the homorganic vowels and

/m/ is not affected by the vowel environment. However, Tabain et al. (2016) did not find out any definite pattern of duration for different places of articulation of the nasals.

In this study, N1 is in the range of 300-350 Hz, which is closer to the finding of Fujimura (1962) and Tabain (2019), but it is slightly higher than the findings of Stevens (2000) and Johnson (2003). The N1 value measured by Pokharel (1989), for the Nepali nasals, is comparatively fairly lower (below 250 Hz) than the findings in the literature. The N2 value, in this study, is higher than the value calculated by Fujimura (1962), Stevens (2000), Johnson (2003), and Tabain (2019), but it is lower than the value calculated by Tabain et al. (2016). The N3 and N4 values are almost similar to the calculations available in the literature, but the N4, calculated by Tabain et al. (2016), is much higher. The present study finds out that N1 is the lowest for /m/ followed by /n/ and /ŋ/ respectively, which is a universal tendency as stated by Recasens (1983). The same pattern is followed by N2, N3, and N4, except, N2 is the highest for /n/, but Tabain et al. (2016) find that N2, N3, and N4 are lower for /ŋ/ than for /n/.

It is well known that the formant values are relatively higher for female speakers. Similarly, The nasal formants are slightly influenced by the vowel environments that the nasals occur in. N1 is lower in the environment of high vowels, and it is higher in the environment of low vowels. Likewise, N2 follows the same trend in that it is lower in the environment of high vowels and higher in the environment of low vowels. It is because of the influence of the F1 and F2 of the adjacent vowels. But N3 shows the opposite pattern in that it is higher in the environment of high vowels and lower in the environment of low vowels. But N4 is not affected by the nature of the adjacent vowels. It is unknown what determines the variations in the N3 and N4.

In this study, only BW1 was measured. It is the largest for /ŋ/, followed by /m/, and it is the smallest for /n/. The finding of this study supports the finding of Johnson (2012) that /ŋ/ has the largest BW1 value. Tabain et al. (2016) discovered a similar pattern for the higher formant bandwidths, too. This study determines that the BW1 value is a cue to distinguish the nasal consonants in terms of the places of articulation. The BW1 value is also influenced by the nature of the adjacent vowel. In the environment of high vowels, it is far lower than in the environment of low vowels. So, it is suggested that the bandwidths of the nasal consonants are to be described in relation to the vowel environment, otherwise, the values are no more meaningful. For example, Tabain et al. (2016) mention that the BW1 value for nasal consonants is 100 Hz. On the contrary, in this study, it is 160 Hz in the a-a environment and only 55 Hz in the i-i environment with very high statistical significance.

As in Tabain et al. (2016), the first two spectral moments, the Center of Gravity (CoG) and Standard Deviation (SD) of the spectrum within the range of 1000-5000 Hz were measured. However, the spectral moments did not show any correlation with the place of articulation of the nasal consonants. Therefore, this study points out that the spectral moments do not characterize the place of articulation of the nasal consonant. Instead, they characterize the vowels that are adjacent to the nasal consonants.

Ladefoged and Maddieson (1996) claim that the amplitude of the nasal consonants is less than the amplitude of the adjacent vowels. However, this study did not support the relationship in all the environments except in the a-a environment. But in the i-i and u-u environments, in several cases, the relative amplitude as well the Root Mean Square amplitude were greater for the nasal than for the vowel preceding the nasal. Because of this inconsistency, I did not use the parameter of amplitude to characterize the nasals. Likewise, I measured the antiformants to characterize the nasal consonants as recommended by Fujimura (1962), Recasens (1983), Stevens (2000), Johnson (2003, 2012), and several others. But any objective evidence to clearly identify their exact positions could not be determined. Tabain et al. (2016) state that they had not been able to precisely measure the antiformants. Similarly, Johnson (2012) explains that antiformant may not be a reliable cue to characterize the nasal consonants in real speech, because the background noise may also cause the spectral minima.

5. Conclusion

The closure duration of nasals (nasal stops) is thought to gradually increase from bilabial to velar places of articulation, while the closure duration of plosives (oral stops) is thought to gradually decrease. However, the presumed pattern for the plosive closure duration in Bengali (Mikuteit & Reetz, 2007), Hindi (Benguerel & Bhatia, 1980), and Nepali (Pokharel, 1989; Chalise, 2017) is not supported. In a similar vein, the supposed pattern of closure duration for nasals is not supported in this work, but the duration pattern for nasals is just opposite to the duration pattern for the plosives. In the environment of homorganic vowels, /n/ and /ŋ/ are longer, but /m/ is unaffected by the adjacent vowel.

N1 follows the universal tendency that it is the lowest for /m/ and gradually increases for /n/ and /ŋ/, respectively. N3 and N4 follow a similar pattern, but N2 is the highest for /n/. BW1 is the highest for /ŋ/, followed by /m/ and /n/, respectively. This pattern is true in the same vowel environment. Theoretically, the first antiformant is a good cue to characterize the nasal consonants, but in practice, it is not a useful cue because it is too difficult to figure out its exact location.

The spectral moments (CoG and SD) do not cue the places of articulation of nasals, but they characterize the vowel(s) adjacent to a nasal.

The amplitude of a nasal is lower than the amplitude of the adjacent vowel(s), but it may not always be true if the adjacent vowel is a high vowel.

References

Ayton, J. A. (1820). *A grammar of the Népalese language*. The Hindoostanee Press by Pereira.

Bandhu, C. M. (1979). *Nepāli Bhāṣāko Utpatti (The origin of the Nepali language)* (5th ed.). Sājhāprkāśan.

Bandhu, C. M., Dahal, B. M., Holzhausen A., & Hale A. (1971). *Nepali Segmental Phonology*. SIL & T. U. (reprinted in *Gipan, 3*(1) (2017).

Benguerel, A. P., & Bhatia, T. K. (1980). Hindi stop consonants: An acoustic and fiber-scopic study. *Phonetica, 37*, 134-148.

Chalise, K. P. (2017). Acoustic analysis of the plosives in Nepali. *Gipan, 3*(1), 41-64.

Chen, M. Y. (1997). Acoustic correlates of English and French nasalized vowels. *The Journal of the Acoustical Society of America*, *102*(4), 2350–2370. http://doi.org/10.1121/1.419620

Clark, T. W. (1963). *Introduction to Nepali*. W. Heffer and Sons.

Dahal, B. M. (1974). *A description of Nepali literary and colloquial*. [Unpublished doctoral dissertation]. Faculty of Arts, Deccan College, University of Poona.

Fant, G. (1970). *Acoustic Theory of Speech Production* (2nd ed.). Mouton.

Fujimura, O. (1962). Analysis of nasal consonants. *The Journal of the Acoustical Society of America, 34*, 1865-1875.

Genetti, C. (1994). Aspects of Nepali grammar. *Santa Barbara Papers in Linguistics, 6*. UCSB.

Hayard, K. (2013). *Experimental phonetics*. Routledge.

Johnson, K. (2003). *Acoustic and auditory phonetics* (2nd ed.). Blackwell Publishing.

Johnson, K. (2012). *Acoustic and auditory phonetics* (3rd ed.). Wiley-Blackwell.

Khatiwada, R. (2009). Nepali. *Journal of the International Phonetic Association, 39*(3).

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Blackwell.

Mikuteit, S. & Reetz, H. (2007). Caught in the ACT: The timing of aspiration and voicing in East Bengali. *Language and Speech, 50*(2), 249-279.

Pokharel, M. P. S. (1989). *Experimental analysis of Nepali sound system* [Unpublished doctoral dissertation]. Department of Linguistics, Post-Graduate & Research Institute, Deccan College, University of Poona.

Recasens, D. (1983). Place cues for nasal consonants with reference to Catalan. *The Journal of the Acoustical Society of America, 73* (4), 1346-1353.

Stevens, K. N. (2000). *Acoustic phonetics*. MIT Press.

Tabain, M. (2019). The phonetic properties of consonants. In W. F. Katz, & P. F. Assmann (Eds.), *The Routledge handbook of phonetics* (pp. 264-288). Routlege Taylor and Fransis Group.

Tabain, M., Butcher, A., Breen, G., & Beare, R. (2016). Acoustic study of nasal consonants in three Central Australian languages. *The Journal of the Acoustical Society of America, 139* (2), 890-903. http://doi.org/10.1121/1.4941659