Social dimensions of the nasal prefix in Jakarta Indonesian

Ferdinan Okki KURNIAWAN
Universitas Katolik Indonesia Atma Jaya

Standard Indonesian has a widely used active verbal prefix /məN-/ marking the active voice which alternates in its shape at the prefix-root boundary. Most previous studies were devoted to Standard Indonesian as spoken in formal contexts. Less attention, however, has been given specifically to the socio-phonological variation of this verbal prefix in a more colloquial variety of Indonesian, as the casual everyday language spoken in Jakarta, i.e., Jakarta Indonesian, with the cognate prefix /N-/ or /ŋə-/. Rather than relying on impressionistic observation, data in this study are drawn from a naturalistic speech corpus in informal settings (Gil et al. 2015). The use of a corpus could help us to verify impressionistic observations and allow us to understand more about the patterns of phonological variation of nasal assimilation. This study finds extra linguistic conditioning, namely gender and educational backgrounds of the speakers, determine patterns of variation of the nasal prefix.

1. Introduction[1]

The use of the nasal prefix in Jakarta Indonesian (JI) exhibits variation beyond the phonological conditioning environment. There are four variants that occur when the nasal prefix patterns with root-initial voiced obstruents [b-, d-, d͡ʒ-, and g-]. First, there is an active prefix that consists of the root alone (henceforth: bare verbs) as in (1a). Second, while /ŋə-/ (1b) vs. /N-/[2] (1c) is phonologically conditioned for voiceless obstruents, the realization is variable for voiced obstruents. Finally, there is a fourth variant [məN-] (1d) which is the Standard Indonesian (SI) form.

(1) Root: bəli 'to buy'

Active forms:
(a) Ø-bəli (JI)
(b) ŋə-bəli (JI)
(c) m-bəli (JI)
(d) məm-bəli (SI)[3]

When [məN-] is found in informal speech, it can be considered as a case of register shift. This is also the case in our corpus, as further discussed in Section 4. While the bare verb is more common than other variants, it is not that informative for our study. Since the variant that consists of bare verbs is chosen equally by all of the groups that we examine in this study (see Section 4)—that is, all age groups, both genders, and all subjects classified by educational attainment, show almost the same percentage of occurrences of

[1] This article and the research behind it would not have been possible without exceptional support from Abby Cohn, John Wolff, and Draga Zec. I am also indebted to two anonymous reviewers for insightful comments that greatly improved the manuscript. All errors are mine.

[2] The symbol N- is explained in Sections 2 and 3.

[3] All forms in the data are given in international phonetic alphabet (IPA).

KURNIAWAN, Okki Ferdinan. 2022. 'Social dimensions of the nasal prefix in Jakarta Indonesian'. NUSA 72: 79–100.
Permanent URL: http://repository.tufs.ac.jp/handle/10108/120256 doi:10.15026/120256

the variant consisting of the bare verbs alone, this variant does not provide insight into social characteristics of the JI community.[4]

This study supports previous studies (Hidajat 2010, Sneddon 2006, among others) that the speakers of JI variably produce the two variants in (1b) and (1c). In order to describe these patterns of variation, this study examines their patterns of use in Gil et al.'s (2015) corpus. The study shows variability when the nasal prefix is combined with roots that have initial voiced obstruents. This study investigates the sources of this variability, whether the variability is due to phonological conditionings or social factors. The organization of this paper is as follows. Description of the nasal prefix is presented in Sections 2 and 3. Then, the corpora and speakers are explained in Section 4. It is then followed by the presentation of the results from the corpora in Sections 5-7. Finally, Section 8 provides the discussion and conclusion of this study.

2. Nasal prefix in Standard Indonesian

In this section, the standard SI prefix /məN-/ is discussed. In SI, a nasal in the coda of the prefix has various shapes when it is added to root-initial consonants and vowels. It is assumed widely that the underlying nasal coda of the prefix is a placeless nasal /məN-/ or velar nasal /məŋ/ in SI (following previous major works as in Sneddon et al. 2010, Sneddon 2006, Pater 1999, Dardjowidjojo 1978, MacDonald & Dardjowidjojo 1967, among others).

The prefix patterns as follows when combined with root-initial sonorous consonants and vowels.[5] First, we can see that nasal assimilation does not apply to the root-initial sonorants and vowels. We can see that the nasal prefix is realized as [mə-] when combined with the root-initial sonorants, including liquids and glides and nasals in (2a-h).

[4] The variant consisting of the bare verbs alone is homophonous with other verb forms. They consist of the verb morpheme in its underlying form. Other verb forms consisting of bare verbs alone have different morphemic context (cannot be substituted for by a root plus nasal assimilation, or a root plus [ŋə-], or a base plus [məN-]). The most frequent are a form of the passive (called "Type 2"), as described by Dardjowidjojo (1978), Sneddon (1996), Cole et al. (2006), among others, and also "middle" verb described by Wolff (1986:B.33).

[5] Some variations in this pattern are seen with loanwords. A production task study on this variation is conducted by Kurniawan (2016). The variation in SI occurs when the nasal prefix occurs with root-initial voiceless stops [p, t, k] in English, Dutch, Portuguese, and Arabic loanwords. For example, underlying form /məN+target+kan/ 'to target' are possibly uttered by speakers as [mənargetkan] and [məntargetkan] 'to target'.

(2) Sonorant initial roots:

Root:	Prefixed forms [mə-]:	
(a) lamar	mə-lamar	‘to propose’
(b) rusak	mə-rusak	‘to destroy’
(c) jakin+i[6]	mə-jakini	‘to believe’
(d) wabah	mə-wabah	‘to be epidemic’
(e) makan	mə-makan	‘to eat’
(f) nilai	mə-nilai	‘to grade’
(g) ɲaɲi	mə-ɲaɲi	‘to sing’
(h) ŋaŋgur	mə-ŋaŋgur	‘to do nothing’

The nasal in the coda position is realized as velar [ŋ] when the roots begin with vowels (3a-e), and [-h] (3f).

(3) Vowel and [-h] initial roots:

Root:	Prefixed forms [məŋ-]:	
(a) aŋkat	məŋ-aŋkat	‘to lift’
(b) obat+i	məŋ-obati	‘to medicate’
(c) undaŋ	məŋ-undaŋ	‘to invite’
(d) ekor	məŋ-ekɔr	‘to follow’
(e) isi	məŋ-isi	‘to fill’
(f) hantam	məŋ-hantam	‘to hit’

The data in (4) show the pattern of nasal prefix when it is combined with roots that begin with voiced obstruents in SI.

(4) Voiced obstruent initial roots:

Root:	Prefixed forms (assimilation):	
(a) bəli	məm-bəli	‘to buy’
(b) dapat	men-dapat	‘to get’
(d) d͡ʑawab	məɲ-d͡ʑawap	‘to answer’
(c) guntiŋ	məŋ-guntiŋ	‘to cut with scissors’

The underlying nasal assimilates to the place of articulation of the root-initial voiced obstruents forming homorganic clusters. The next pattern is with root-initial voiceless obstruents.

[6] The suffix *-i* has a locative function.

(5) Voiceless obstruent initial roots:

Root:	Prefixed forms (substitution):	
(a) pilih	məm-ilih	‘to choose, to vote’
(b) tulis	mən-ulis	‘to write’
(c) kasih	məŋ-asih	‘to give’
(d) sapu	məɲ-apu	‘to sweep’
Root:	Prefixed forms (assimilation):	
(e) t͡ɕari	məɲ-t͡ɕari	‘to seek’

The pattern in (5a-c) is termed as nasal substitution by many phonologists. From a derivational perspective, the nasal assimilates to the root-initial voiceless consonants [p, t, k] which then forms a homorganic cluster. This process is then followed by the deletion of the initial voiceless consonant. For example, the root [pilih] is first prefixed as [məm-pilih], and then [p] is deleted so that the final form is realized as [məm-ilih].

The fricative /s/ in (5d) becomes a palatal nasal [ɲ] rather than an alveolar nasal.[7] Lapoliwa (1981) proposed that the affricate /t͡ɕ/ in (5e) is realized as voiceless lamino alveolo-palatal affricate. No substitution is applied to /t͡ɕ/-initial root but rather the underlying nasal assimilates to the affricate and is realized as the palatal nasal /ɲ/ but the affricate /t͡ɕ/ is not deleted.

Let us now turn to the description of the nasal prefix in JI.

3. Nasal prefix in Jakarta Indonesian

The nasal prefix in JI is cognate with the nasal prefix in SI. The description below is based on previous studies by Ikranagara (1980), Muhadjir (1981), Sneddon (2006), and Kurniawan (2015).

As discussed in Section 1, the active verb has four different alternants that vary between bare verbs with no prefix, nasal assimilation, [ŋə-], and [məN-]. Further, the bare verbs are homonymous with other bare verbs which are of a different morpho-syntactic context.

Similar to the nasal coda of SI, it is assumed that the underlying nasal coda of the prefix is /N-/ in Betawi Malay (Muhadjir 1981) and JI (Sneddon 2006). However, other studies such as Ikranagara (1980) suggested /ŋ-/ as the underlying form for Betawi. The current study does not aim to solve these different hypotheses of the underlying forms, but rather to focus more on the patterns of variation that may contribute to our understanding of the development of JI.

Similar to SI, the nasal prefix in JI also has several alternants. The patterns of the alternants are presented as follows. First, the roots that begin with liquids, glides, and vowels in JI are presented in (6) and (7).

The alternants with [ŋə-] occur with the root-initial liquids, glides, and [-h] as in (6a-e), while with root initial vowels as in (7a-e), this prefix is realized as a velar nasal.

[7] Wolff (2010) shows that Malay/Indonesian /s/ derives from Proto-Austronesian phoneme *c, which is on alveolar or palatal affricate. This is likely the reason why /s/ is realized as [ɲ] rather than [n].

(6) Sonorant initial roots:

Root:	Prefixed forms [ŋə-]:	
(a) lamar	ŋə-lamar	'to propose'
(b) rusak	ŋə-rusak	'to destroy'
(c) jakin+in[8]	ŋə-jakinin	'to believe'
(d) wabah	ŋə-wabah	'to be epidemic'
(e) harus+in	ŋə-harusin	'to require'

(7) Vowel initial roots:

Root:	Prefixed forms [ŋ-]:	
(a) aŋkat	ŋ-aŋkat	'to lift'
(b) obat+in	ŋ-obatin	'to medicate'
(c) undaŋ	ŋ-undaŋ	'to invite'
(d) ekor	ŋ-ekɔr	'to follow'
(e) isi	ŋ-isi	'to fill'

The data in (8) illustrate the nasal prefix when conditioned by root-initial voiceless obstruents.

(8) Voiceless obstruent roots:

Root:	Prefixed forms (substitution):	
(a) pilih	m-ilih	'to choose, to vote'
(b) tulis	n-ulis	'to write'
(c) kasih	ŋ-asih	'to give'
(d) sapu	ɲ-apu	'to sweep'
(e) t͡ɕari	ɲ-ari	'to seek'

Except for [t͡ɕ-] initial (8e), the nasal prefix patterns in (8) are similar to SI, as presented in (5). In her study of Betawi, Ikranagara (1980) proposed that this phonological process involves two steps. First, the nasal assimilates to the root-initial voiceless consonants which then forms a homorganic cluster as in [m-pilih] 'to choose, to vote.' Second, the initial [p] is then deleted and it finally surfaces as [m-ilih]. Data in (9) show the patterns with root-initial nasals. When the nasal prefix is combined with root-initial nasal, they are realized as a single nasal, as shown in (9).

[8] The suffix -*in* has locative/applicative function.

(9) Nasal initial roots:

Root:	Prefixed forms Ø:	
(a) makan	makan	‘to eat’
(b) nilai	nilai	‘to grade’
(c) ɲaɲi	ɲaɲi	‘to sing’
(d) ŋaŋgur	ŋaŋgur	‘to do nothing’

The approach that is used to account for the SI facts could be extended to account for the phonological alternation in JI. However, root-initial voiced obstruents exhibit different patterns from the root initial sounds of the nasal prefix that we have discussed so far. For root-initial voiced obstruents there is a pattern of variation as displayed in (10a-d) below.

(10) Voiced obstruent initial roots:

Root:	Prefixed forms (assimilation ~ [ŋə]):	
(a) bəli	i. m-bəli	‘to buy’
	ii. ŋə-bəli	
(b) dapət	i. n-dapət	‘to get’
	ii. ŋə-dapət	
(c) d͡ʑawab	i. ɲ-d͡ʑawap	‘to answer’
	ii. ŋə-d͡ʑawap	
(d) guntiŋ	i. ŋ-guntiŋ	‘to cut with scissors’
	ii. ŋə-guntiŋ	

The forms in (10a-d. i) are used by some JI speakers. In addition, the forms in (10a-d. ii) with [ŋə-] variants–similar to the forms seen for liquids and glides–are also observed for some speakers. In his study of JI, Sneddon (2006) mentioned such variation but did not offer a full description of it. Kurniawan (2015) offered a fuller description of the variation based on a corpus study and speech production task. Using the same corpus used in Kurniawan’s (2015) study, the current study offers a more extensive investigation to provide a better understanding of what motivates the variation.

Based on previous studies of Betawi (Muhadjir 1981; Ikranagara 1980; among others), the variation in (10) is also found among Betawi speakers. As described specifically by Ikranagara (1980:135), the nasal assimilation [m-b, n-d, ɲ-d͡ʑ, ŋ-g] in Betawi is the same as the one used in Javanese and the variant with [ŋə-] is closely related to Sundanese [ŋa-]. Interestingly, Nothofer (1995) reported that [ŋə-] is also found in the southwestern dialects of Malay, particularly Bangka Malay. He proposed that Betawi speakers adopted [ŋə-] from Bangka Malay rather than Sundanese. While the solution to this disagreement is beyond the scope of this current study, this background will be relevant to our discussion in Section 8.

The key question to be addressed in this paper is what conditions the variation and why. One of the central goals of this study is to shed light on the pattern of variation of the nasal prefix and to examine how the results of this study may contribute to our understanding of the development of JI. Data in this study are drawn from a corpus of naturalistic conversations (Gil et al. 2015). It is important to see how this variation is

actually produced spontaneously by native speakers in naturalistic data. To the best of my knowledge, none of the prior studies of nasal assimilation in Indonesian have used a large data set from the naturalistic conversation and provided quantitative evaluations to gain a more insightful understanding of the social dimensions of the variation. The investigations with the JI corpus aim to seek evidence whether the two different variants shown in (10) are due to linguistic or social factors. The details of the corpus are provided in the next section.

In the case of the nasal prefix, the linguistic factors may involve morpho-phonological and lexical conditioning, while non-linguistic factors may include social factors that result in intra-speaker and inter-speaker variation. In terms of morpho-phonological conditioning, the place of articulation might determine patterns of variation of the nasal prefix. For example, speakers may produce more [m-b] than [ŋə-b] in bilabial-initial roots, but on the other hand, they produce more [ŋə-g] than [ŋ-g] in velar-initial roots. In Tagalog nasal substitution, Zuraw (2010) shows different rates of substitution based on place of articulation. The current study seeks to find out if the occurrence of nasal assimilation in JI may also differ across place of articulation.

In lexical conditioning, some words with the same root initial sounds may be codified [m-b], while other words may be codified [ŋə-b] in the speakers' lexicon. For example, native speakers may produce [m-bəli] 'to buy', but never [ŋə-bəli]. In other lexical items, they may produce [ŋə-bantu] 'to help', but never [m-bantu]. Equally important, we also need to investigate systematically if the variation is conditioned by non-linguistic factors such as gender and education. It is commonly discussed in sociolinguistic literature that females may use variants that are associated with higher social prestige (Angle & Hesse-Biber 1981, Trudgill 1972). In variationist sociolinguistics, educational background is often used as a social factor to determine speakers' position in a social structure (Mallinson 2007, Labov 2001). With that in mind, the current study intends to examine if the nasal assimilation and the [ŋə-] forms may interact with these two non-linguistic factors.[9] Interestingly, the current investigation finds that the variant [m-b] has a higher frequency among females and speakers of higher educational backgrounds.

4. Corpus, speakers, and statistical methods

As mentioned in the previous section, this study is based on corpus collected by Gil et al. (2015) between 2004 and 2012. This corpus is a part of larger language documentation projects conducted at the Max Planck Institute for Evolutionary Anthropology, Jakarta Field Station. This corpus consists of naturalistic conversations collected in relaxed and informal settings. The corpora also involve speakers from other ethnicities whose first language are not JI, such as Betawi (a minority dialect of Malay spoken in Jakarta), Javanese, Sundanese, Batak. In this study, I only include JI speakers who are considered by Wallace (1976) as those who were born and grew up in Jakarta and their parents are not of Betawi descents. Please see Wallace (1976), Grijns (1991), Sneddon (2006), Kurniawan (2018), among others, for the discussion about JI and Betawi distinction.

For this study, I analyze the speech of a total of five male speakers of lower educational

[9] Although occupation is coded in the corpus and may also be a potential factor that interacts with the two variants, the current study does not include it because in the metadata we have, sometimes the occupation does not correlate with educational attainment. For example, a speaker might have a lower educational background but own a highly profitable business.

background and seven male speakers of higher educational background. For female speakers, there is a total of five female speakers of lower educational background and four female speakers of higher educational background. Thus, there is a total of twenty-one speakers investigated in the corpus. The speakers of lower educational background are the speakers who have no college degree and the speakers of higher educational backgrounds are those who have completed undergraduate degree or are undertaking a degree at the time of data collection. Their ages range between twenty-one and forty-nine years old. The JI speakers from the corpus are chosen because they produce more tokens than the other speakers involved in the documentation.

Following Gries' (2017) methodology for analyzing corpus data quantitatively, the current study applies the chi-squared test for goodness of fit. The frequency distribution of the corpus data and all other variables (i.e., social categories) involved in this study are categorical. For these reasons, the chi-squared test for goodness of fit is applicable to analyze the data we have. We want to examine how good the fit of the observed occurrence of the variants is to some expected distribution. In the chi-squared test, we expect the distribution of the two variants to be equally frequent. Suppose the *p* value we find in the test is larger than the standard threshold value of 0.05. In that case, we conclude that the occurrence of the variants is not different from the expected distribution (i.e., no significant difference). If we find the *p* value lower than 0.05, we conclude that there is a significant difference between the occurrence of the variants. All the tests are done in R statistical software (R Core Team 2022).

5. Results of the corpus study - phonological and lexical conditioning

As we have seen in (10), JI has two variants of the active prefix in voiced obstruent initial stems. They are realized in the variant with assimilation as in [m-b, n-d, ɲ-d͡ʒ, ŋ-g,] or a variant with [ŋə-] as in [ŋə-b, ŋə-d, ŋə-d͡ʒ, and ŋə-g]. As discussed in Section 1, there are also variants with bare verbs as in [b-, d-, g-, d͡ʒ-] and variants with [məN-] as in [məm-b, mən-d, məŋ-g, məɲ-d͡ʒ] found in the corpus. The variant with assimilation, [ŋə-], and bare verbs are used in informal settings (JI), while the variant with [məN-] is used in formal settings (SI). This is a normative distinction and we can consider the use of [məN-] in informal settings as a case of style shift.

The root /bəli/ 'buy', as exemplified in (1), when combined with the nasal prefix may be realized in four different variants, depending on different settings. Table 1 below shows the four variants and their relationships with their settings and varieties.

Table 1. Examples of four variants of nasal prefix

Variants	Settings	Varieties
a) [m-bəli]	informal	Jakarta Indonesian
b) [ŋə-bəli]	informal	Jakarta Indonesian
c) [bəli]	informal	Jakarta Indonesian
d) [məm-bəli]	(in)formal	Standard Indonesian

This study searched the four variants in the corpus but focuses primarily on analyzing the variant with assimilation, and the variant with [ŋə-]. The search is limited only to active

transitive forms.[10]

The organization of the results is as follows. This section is focused on phonological and lexical conditioning, regardless of speakers' social categories. Following that, Section 6 presents the results based on speakers' social categories, specifically, gender and educational background.

First, this section aims to observe whether the patterns of variation are conditioned by phonological conditioning, specifically place of articulation. Additionally, the second factor that is examined is lexical conditioning. To observe the phonological conditioning, the results presented in this section are generated from the results of a total of twenty-one speakers, regardless of their social categories. This method allows us to see if the occurrence of the nasal prefix is conditioned by place of articulation and if this condition applies to all speakers, regardless of their social background.

To examine the lexical conditioning, some speech samples from five speakers are presented to see whether the corpus exhibits occurrence of nasal prefix variation within the same lexical items. This enables us to observe intra-speaker variation, in which one lexical item can be produced in two variants by the same individual speaker. Let us now first observe the overall results from the total of twenty-one JI speakers included in the corpus study in 5.1.

5.1. Overall results - all variants

In the overall results, there are four variants of the active transitive prefix presented in this section. They are variants with bare verbs, assimilation, [ŋə-], and [məN-]. It is necessary to examine them all since we need to look at whether or not the occurrence of each variant in the corpus is conditioned by place of articulation. Figure 1 displays the overall occurrence of the four variants of the nasal prefix.

[10] A reviewer of this paper mentioned the possibility of syntactic, semantic, and paradigmatic factors determining the realization of the prefix. This might be true for the case of bare verbs, as discussed in Conners and Brugman (2013), Conners, Bowden, and Gil (2013), and Hidajat (2010). However, I could not find any previous studies that specifically discuss the syntactic, semantic, and paradigmatic differences between the variant with assimilation, and the variant with [ŋə-], which will be the focus of this study. Additionally, I have consulted with several native speakers and do not find the two variants to be syntactically, semantically, and paradigmatically different in the active transitive forms.

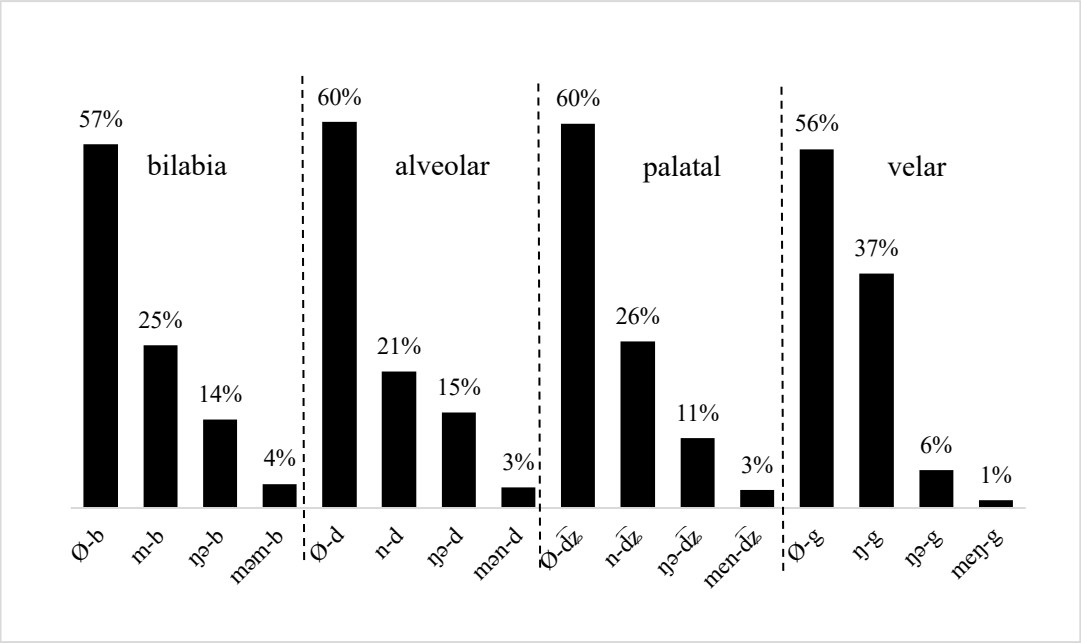

Figure 1. Four variants of nasal prefix

Table 2. Occurrence of nasal prefix – all variants

Place	Bare verbs	Assimilation	ŋə-	məN-	Total	Chi-squared test
Ø-b	226 (57%)					$\chi^2 = 253.15$;
m-b		101 (25%)				*df* = 3;
ŋə-b			55 (14%)			*p* <.001***
məm-b				15 (4%)	397	
Ø-d	113 (60%)					$\chi^2 = 137.9$; *df*
n-d		40 (21%)				= 3; *p*
ŋə-d			28 (15%)			<.001***
mən-d				6 (4%)	187	
Ø-d͡ʑ	127 (60%)					$\chi^2 = 162.82$;
n-d͡ʑ		55 (26%)				*df* = 3;
ŋə-d͡ʑ			23 (11%)			*p* <.001***
mən-d͡ʑ				6 (3%)	211	
Ø-g	323 (56%)					$X^2 = 468.86$;
ŋ-g		211 (37%)				*df* = 3;
ŋə-g			34 (6%)			*p* <.001***
məŋ-g				7 (1%)	575	
						$\chi^2 = 991.83$;
						df = 3;
Total	789 (58%)	407 30%)	140 (10%)	34 (2%)	1,370	*p* <.001***

Note: * indicates a significant difference; ** and *** indicate a higher significant difference (more asterisks indicate the more significant difference). This is applied to the rest of the paper.

The x-axis in Figure 1 exhibits the categories of each nasal prefix variant. They are structured based on the four categories of place of articulation, namely bilabials [Ø-b ~ m-b ~ ŋə-b ~ məm-b], alveolars [Ø-d ~ n-d ~ ŋə-d ~ mən-d], (alveo)-palatal [Ø-d͡ʑ ~ n-d͡ʑ ~ ŋə-d͡ʑ ~ mən-d͡ʑ], and velars [Ø-g ~ ŋ-g ~ ŋə-g ~ meŋ-g]. The vertical bars display

the percentages of nasal prefix variants. The percentage of each variant is calculated based on the actual tokens production in the corpus. For example, [Ø-b-] was produced 226 times (57%), [m-b-] was produced 101 times (25%), [ŋə-b-] was produced fifty-five times (14%), and [məm-b-] was produced fifteen times (4%). There are altogether 97 tokens produced (100%). The search of these 397 tokens was done on the twenty-one speakers. There is a total of 1,370 relevant tokens produced by the twenty-one speakers retrieved from the corpus. The results of the retrieval, however, do not come from all of these twenty-one speakers. For example, a speaker might produce [m-b] quite often but never produce [n-d]. The results retrieved are only from the speakers who produced relevant tokens. This calculation and search methods are applied to all results presented in the rest of this paper.

The overall occurrence of the nasal prefix in Figure 1 shows an interesting skewing occurrence across the place of articulation. We can see that bare verbs have robust occurrences between 56% and 60%, the occurrences of the variant with assimilation vary between 21% and 37%, the occurrences of the variant with [ŋə-] vary in the range of 6% and 15%, and the occurrences of the variant with [məN-] are between 1% and 4%. These patterns of occurrence are found similarly across the place of articulation, i.e., [b-, d-, d͡ʒ-, and g-] initials. The results of the chi-squared test for goodness of fit displayed in the last column of Table 2 show highly significant differences between the four variants in all places of articulations. The differences can be observed in *p* values, which are similar across bilabials, alveolars, palatals, and velars. The results suggest that the place of articulation does not determine the occurrence of the four variants of the nasal prefix in JI.

The similar percentages of the frequency of the bare-verb variant across all social classes, age groups, and both genders may indicate that this variant is the neutral form among other variants. The bare verbs and the standard variant [məN-] are not the primary focus of our analysis. The variation that yields substantial results are the variants with nasal assimilation and those with [ŋə-] before voiced stop initial roots. Therefore, this article is devoted to elaborating the results of these two variants only. When we are focusing on these two, results are presented as the percentage from the nasal assimilated variant and those with [ŋə-], since these two variants together make up 100%. The overall results for these two variants are presented in 5.2.

5.2. Overall results – nasal assimilation

After narrowing down the results to the assimilated and [ŋə-] variants, let us now see if the varied patterns of occurrence we have are conditioned by place of articulation or lexical conditioning.

Table 3. Occurrence of nasal prefix – two variants

Place	Assimilation	ŋə-	Total	Chi-squared test for goodness of fit
m-b	101 (65%)			$\chi^2 = 13.564$, $df = 1$, $p < .001$***
ŋə-b		55 (35%)	156	
n-d	40 (60%)			$\chi^2 = 2.1176$, $df = 1$, $p = .145$
ŋə-d		28 (40%)	68	
n-d͡ʑ	55 (71%)			$\chi^2 = 13.128$, $df = 1$, $p < .001$***
ŋə-d͡ʑ		23 (29%)	78	
ŋ-g	211 (86%)			$\chi^2 = 127.87$, $df = 1$, $p < .001$***
ŋə-g		34 (14%)	245	
Total	407 (74%)	140 (26%)	547	$\chi^2 = 130.33$, $df = 1$, $p < .001$***

In Table 3, we can see that across the place of articulation, nasal assimilation is generally produced higher than the [ŋə-] form. The chi-squared test for goodness of fit is conducted to observe whether the difference between the occurrence of the two variants in each place of articulation is statistically significant. The results show highly significant differences in bilabials, palatals, and velars. The only insignificant difference is found in alveolars. Similar to our results from the four variants in 5.1, the results of the two variants here suggest that, in general, place of articulation has little effect on the patterns of variation of the nasal prefix.

The next aspect we should consider is lexicalization. If JI shows lexicalization, in which one lexical item may only be produced in one form, we would find no variation by lexical items, and we also would not be able to find intra-speaker variation. The current study shows that the corpus exhibits occurrence of the variant with assimilation and the variant with [ŋə-] forms the same lexical items and place of articulation. The same lexical item can be produced in two forms by the same individual speaker. Table 4 exemplifies intra-speaker variation within the same place of articulation, in this case, bilabials, as shown in the corpus.

Table 4. Examples of intra-speaker variation

Speaker	Root	Assimilation	ŋə-	English
M-H-S4*	bakar	m-bakar (1)	ŋə-bakar (1)	to burn
F-H-S3	bawa	m-bawa-in (1)	ŋə-bawa-in (2)	to carry something for someone
F-L-S1	bərantak	m-bərantak-in (1)	ŋə-bərantak-in (1)	to make a mess
M-H-S4	bəli	m-bəli (2)	ŋə-bəli (1)	to buy
M-H-S3	bajar	m-bajar (2)	ŋə-bajar (1)	to pay

*Note: M: male, F: female, L: lower educational background, H: higher educational background, S4: Speaker 4, S3: Speaker 3, and so on. These codes are used for the rest of the paper.

We can see from the results in Table 4 that indeed the same speaker produces both variants for the same lexical item. The number in the parentheses indicates the number of tokens. Here, intra-speaker variation occurs in the JI corpus where both variants were uttered in the spontaneous speech by the same speaker.

It should be noted that the intra-speaker variation in the results above does not seem to be

caused by different social settings. As can be observed in the corpus, all the speakers uttered them when interacting with the addressees in similar kinds of informal settings.

In summary, the choice of the variant with nasal assimilation or [ŋə-] is not generally conditioned by place of articulation nor is it conditioned by the particular lexical items (i.e., linguistic factors). Additionally, the results also show that the variant with nasal assimilation is of higher frequency than the variant with [ŋə-] and we need to know what factors may cause it. Since the linguistic factors discussed so far do not seem to be the source of the variation, we now need to examine a non-linguistic factor, namely gender, which is described in Section 6.

6. Results of the corpus study – gender

As described above, the total number of speakers investigated in this section is twenty-one. They consist of nine female speakers and twelve male speakers. First, we consider how gender categories may affect the occurrence of nasal assimilation. Trudgill (1972) discussed that the linguistic forms produced by female speakers are more often related to prestige standards than by male speakers. Although the variants with nasal assimilation and [ŋə-] do not belong to the standard variety, it is important to examine whether their patterns of use differ between female and male speakers.

To account for these variables, this section also uses a chi-square test for goodness of fit. We would like to examine if the occurrence of the variant with assimilation and the variant with [ŋə-] forms are conditioned by the speakers' genders. The results from the test are shown in the following table.

Table 5. Female and male speakers

Female speakers				
Place	Assimilation	ŋə-	Total	Chi-squared test for goodness of fit
m-b	76 (84%)			$\chi^2 = 40.89$, $df = 1$, $p < .001$***
ŋə-b		15 (16%)	91	
n-d	19 (76%)			$\chi^2 = 6.76$, $df = 1$, $p < .009$**
ŋə-d		6 (24%)	25	
n-d͡ʑ	41 (89%)			$\chi^2 = 28.174$, $df = 1$, $p < .001$***
ŋə-d͡ʑ		5 (11%)	46	
ŋ-g	163 (89%)			$\chi^2 = 111.74$, $df = 1$, $p < .001$***
ŋə-g		20 (11%)	183	
Total	299 (87%)	46 (13%)	345	$\chi^2 = 185.53$, $df = 1$, $p < .001$***
Male Speakers				
m-b	25 (40%)			$\chi^2 = 2.3226$, $df = 1$, $p = .127$
ŋə-b		37 (60%)	62	
n-d	21 (51%)			$\chi^2 = 0.02439$, $df = 1$, $p = .875$
ŋə-d		20 (49%)	41	
n-d͡ʑ	14 (47%)			$\chi^2 = 0.13333$, $df = 1$, $p = .715$
ŋə-d͡ʑ		16 (53%)	30	
ŋ-g	48 (80%)			$\chi^2 = 21.6$, $df = 1$, $p < .001$***
ŋə-g		12 (20%)	60	
Total	108 (56%)	85 (44%)	193	$\chi^2 = 2.7409$, $df = 1$, $p = .097$

In Table 5, we can see that the overall occurrence of the assimilation produced by both male and female speakers are different and as previously noted, not conditioned by place of articulation overall (with the possible exception of the velar place).

In terms of gender, the results in Table 5 show notably different patterns of occurrence between the male and female speakers. As shown in the last column of Table 5, the chi-squared test yielded different results. Across the place of articulation, the *p* values from the female speakers show highly significant differences between the assimilation and [ŋə-] forms. In contrast, the *p* values from male speakers show insignificant differences across place of articulation, except for velars.

Assimilation is more robustly produced by the female speakers than by the male speakers. It should be noted here that the percentage of variants with velar assimilation on the part of the male speakers (80%) is not very different from that of the female speakers (89%).

Among the female speakers, the assimilation ranges between 70% - 90% across places of articulation, while assimilation among male speakers ranges between 40% - 60% in bilabials [m-b], velars [n-d], and palatals [n-d͡ʑ]. In the velars, assimilation [ŋ-g-] is 80%. Aside from the case of the velars, gender affects the patterns of variation found thus far. The female speakers produced less variation: the variants with assimilation were produced robustly across places of articulation, whereas the male speakers produced more variation: the variants with assimilation were produced in the range between 40% and 60%.

The next key question to be addressed is, what conditions the more varied patterns of the nasal assimilation found among the male speakers. To account for this, we need to examine an additional non-linguistic factor, namely educational background, which is elaborated further in Section 7.

7. Results of the corpus study - education

This section is divided into two main parts. First, 7.1 compares the results between female speakers of higher educational background and the results from female speakers of lower educational background. Second, the results from male speakers of higher educational background are compared with the results from male speakers of lower educational background in 7.2. This section also performs the chi-squared for goodness of fit to observe if education affects the occurrence of the assimilation and [ŋə-] forms produced by the female speakers.

7.1. Female speakers and educational background

There are four female speakers of higher educational background and five female speakers of lower educational background involved in this study. The results are presented in Table 6.

Table 6. Female speakers and educational backgrounds

Female speakers of higher educational background				
Place	Assimilation	[ŋə-]	Total	Chi-squared test for goodness of fit
m-b	49 (82%)			$\chi^2 = 24.067$, $df = 1$, $p < .001$***
ŋə-b		11 (18%)	60	
n-d	6 (60%)			$\chi^2 = 0.4$, $df = 1$, $p = .527$
ŋə-d		4 (40%)	10	
n-dʑ	23 (88%)			$\chi^2 = 15.385$, $df = 1$, $p < .001$***
ŋə-dʑ		3 (12%)	26	
ŋ-g	123 (87%)			$\chi^2 = 76.169$, $df = 1$, $p < .001$***
ŋə-g		19 (13%)	142	
Total	201 (84%)	37 (16%)	238	$\chi^2 = 113.01$, $df = 1$, $p < .001$***
Female speakers of lower educational background				
m-b	27 (87%)			$\chi^2 = 17.065$, $df = 1$, $p < .001$***
ŋə-b		4 (13%)	31	
n-d	13 (87%)			$\chi^2 = 8.0667$, $df = 1$, $p = 0.004$**
ŋə-d		2 (13%)	15	
n-dʑ	23 (92%)			$\chi^2 = 17.64$, $df = 1$, $p < .001$***
ŋə-dʑ		2 (8%)	25	
ŋ-g	40 (98%)			$\chi^2 = 37.098$, $df = 1$, $p < .001$***
ŋə-g		1 (2%)	41	
Total	63 (89%	8 (11%)	71	$\chi^2 = 42.606$, $df = 1$, $p < .001$***

We can identify right away that the patterns of occurrence of assimilation and [ŋə-] form between the two groups are overall quite similar in Table 6. In the last column, the results of the chi-square test show significant differences between the two variants across the two groups, except for alveolars produced by the female speakers of higher education.

The case of alveolars might be caused by the lack of occurrence found in the corpus.

We can observe that the occurrence of nasal assimilation by both female speakers of higher and lower educational backgrounds are much more robust than [ŋə-] form. They range between 60% and 98%. Moreover, we can also see that place of articulation does not affect the patterns of occurrence, except in the case of the velars, where assimilation is much more robust than the other places of articulation and for alveolars for females with higher levels of education which is lower than expected. Interestingly, these robust patterns of assimilation among the female speakers of lower educational background are quite like the patterns of assimilation among female speakers of higher educational background.

This suggests that the educational factor among the female speakers has little effect on the patterns of variation of the two nasal prefix variants. Nasal assimilation is produced more frequently than [ŋə-] variant across places of articulation, as our linguistic factor, and across the educational categories as our non-linguistic factor.

We might wonder whether these patterns of occurrence are dominated by certain speakers due to individual speaker differences. To assess this, the results by speakers are examined and exhibited in Figure 2.

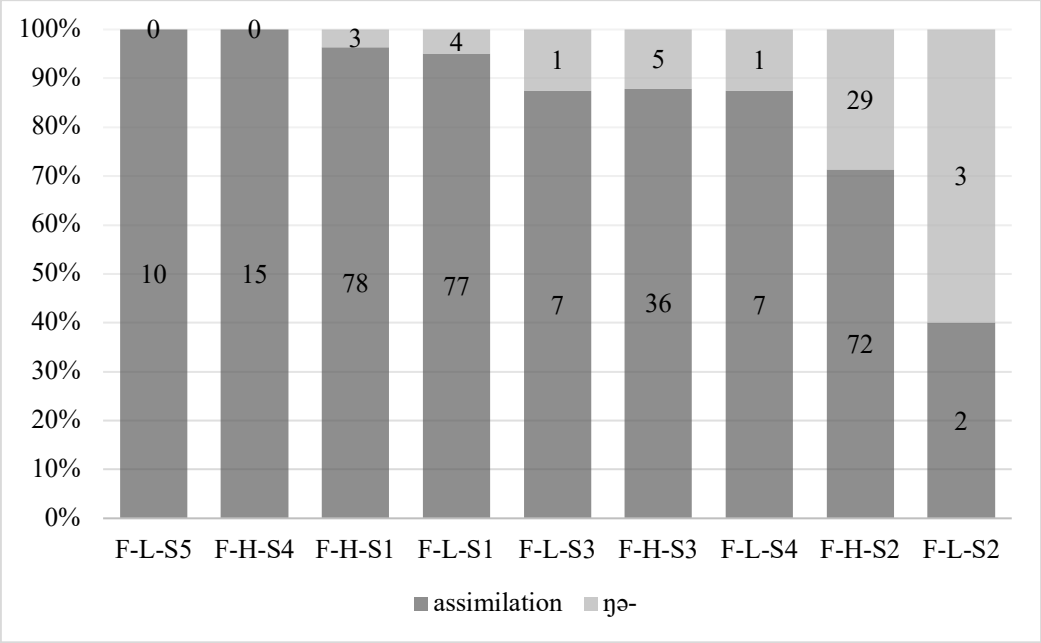

Figure 2. Female speakers

In Figure 2, we can see that in general, the female speakers of both educational backgrounds produced nasal assimilation robustly, with the frequency above 71%, except for speaker F-L-S2 who produced only 40% nasal assimilation. The speaker F-L-S2 lives in a neighborhood that is populated mostly by a Betawi population, and her results might be caused by her frequent interaction with her Betawi neighbors. We will come back to discuss the issue of Betawi influence in Section 8.

The overall results suggest that the patterns of variation presented in Figure 5 are not biased towards certain speakers, except for the case of speaker F-L-S2 whose patterns of usage are different.

Thus, based on Figures 4 and 5, we have supporting evidence to show that the place of articulation and educational category do not have a strong effect on the robust occurrence of the variant with nasal assimilation among female speakers. Let us now turn our attention to the results from male speakers, which are discussed in 7.2.

7.2. Male speakers and educational background

Seven male speakers of higher educational background and five speakers of lower educational background are included in this study. Table 7 here presents the results.

Table 7. Male speakers and educational backgrounds

Male speakers of higher education background				
Place	assimilation	ŋə-	Total	Chi-squared test for goodness of fit
m-b ŋə-b	19 (53%)	17 (47%)	36	$\chi^2 = 0.11111$, $df = 1$, $p = .738$
n-d ŋə-d	14 (64%)	8 (36%)	22	$\chi^2 = 1.6364$, $df = 1$, $p = .200$
n-dʑ ŋə-dʑ	13 (68%)	6 (32%)	19	$\chi^2 = 2.5789$, $df = 1$, $p = .108$
ŋ-g ŋə-g	46 (85%)	8 (15%)	54	$\chi^2 = 32.667$, $df = 1$, $p < .001$***
Total	92 (70%)	39 (30%)	131	$\chi^2 = 21.443$, $df = 1$, $p < .001$***
Male speakers of lower education background				
m-b ŋə-b	6 (23%)	20 (77%)	26	$\chi^2 = 7.5385$, $df = 1$, $p = 0.006$**
n-d ŋə-d	7 (37%)	12 (63%)	19	$\chi^2 = 1.3158$, $df = 1$, $p = .251$
n-dʑ ŋə-dʑ	1 (9%)	10 (91%)	11	$\chi^2 = 7.3636$, $df = 1$, $p = 0.006$**
ŋ-g ŋə-g	2 (33%)	4 (67%)	6	$\chi^2 = 0.66667$, $df = 1$, $p = .414$
Total	16 (26%)	46 (74%)	62	$\chi^2 = 14.516$, $df = 1$, $p < .001$***

Table 7 shows that the results from the male speakers of higher educational background differ from those of lower educational background. The results from the chi-squared tests in male speakers of higher educational background show insignificant differences between the nasal assimilation and [ŋə-] form in bilabials, alveolars, and palatals. The robust occurrence of nasal assimilation only occurs in velar place. For male speakers of lower educational background, there are significant differences between the two variants in bilabials and palatals (i.e., the [ŋə-] form is found significantly higher than the nasal assimilation) and insignificant differences between the two variants in the alveolars and velars. I should note here that the results in the velars might be caused by the lack of occurrence in the corpus.

Interestingly, the chi-squared tests in the 'total' rows exhibit the opposite results. Male speakers of higher educational background produced nasal assimilation significantly higher than the [ŋə-] form. On the other hand, male speakers of lower educational background produced [ŋə-] form significantly higher than nasal assimilation.

The results of the male speakers are different from those of the female speakers discussed in previous sections. Across educational categories, female speakers generally produced nasal assimilation significantly higher than the [ŋə-] form. For the female speakers, educational background made little difference, whereas, for the males, educational background is clearly a factor in the choice of this variant. Overall, the occurrence of assimilation among male speakers of higher educational background is higher than the occurrence of assimilation among male speakers of lower educational background.

To assure that the results in Table 7 are not biased towards certain speakers, let us now observe the occurrence of the variant with nasal assimilation by speakers in Figure 3.

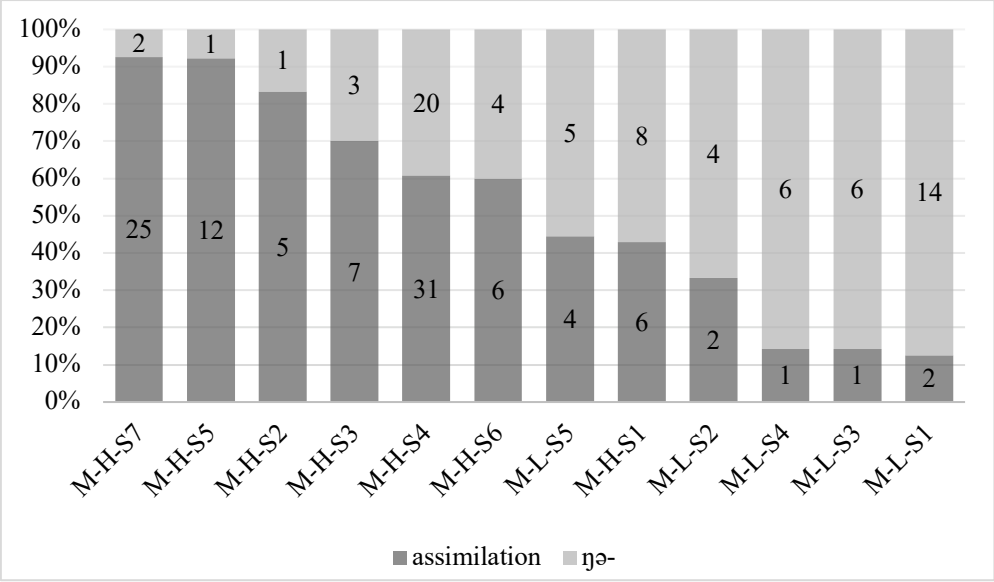

Figure 3. Male speakers

Figure 3 shows the occurrence of assimilation by twelve male speakers. We can see clearly in Figure 3 that the male speakers of higher educational background produced the variant with nasal assimilation for the most part at a rate greater than 50%, except for speaker M-H-S1 who chose the variant with nasal assimilation only 43% of the time, (i.e., 57% of the variants with [ŋə-]). On the other hand, the male speakers of lower educational background, in general, produced more limited occurrences of assimilation. It should be noted here that M-H-S1's parents are both Sundanese, and he also lived with his grandparents who are Sundanese. His higher occurrences of the variant with [ŋə-] might be also conditioned by his exposure to Sundanese that has [ŋa-] that is cognate with Betawi [ŋə-]. The Sundanese influence will be further discussed in Section 8.

The results presented in Figure 3 show that the results described are not biased towards some speakers but indeed shows that the majority of speakers of higher educational background produce a higher percentage of the variant with nasal assimilation, while the speakers of the lower educational background show a lower percentage of the variant with nasal assimilation. In fact, female speakers of all categories (except for F-L-S2) produced a substantially higher percentage of the variant with assimilation than the male speakers of all categories.

These indicate that the occurrence of nasal assimilation produced by the male speakers of

higher education background is closer to the occurrence of nasal assimilation produced by the female speakers, whereas the male speakers of lower educational background show results that are different in that they produce much fewer variants with nasal assimilation than the male speakers of higher educational background.

To sum up, the results from the four groups of speakers suggest that the linguistic factors, either lexical or dependent on the place of articulation do not determine the patterns of variation (though some effects of place are observed). There are differences based on gender. The results also show that among the female speakers, the factor of education does not condition the patterns of variation, while among male speakers, education appears to be a factor in choosing the variant with nasal assimilation.

8. Discussion and conclusion

Sneddon (2006), as cited in Hidajat (2010), suggested that the high percentage of the variant with nasal assimilation in JI mostly occurs among speakers with Javanese background. In addition, Conners and Brugman (2012) reported that the variant with nasal assimilation is ungrammatical and avoided in JI because native speakers recognized it as a Javanese variant. My study, however, shows that nasal assimilation is chosen not only by speakers with Javanese background but also by JI speakers with no ethnic or ancestry relation to Javanese.

Besides the ethnic background of the speakers, we should also consider the processes of language contact, which seem likely to be quite complex. Previous studies on Betawi, such as Muhadjir (1981) and Ikranagara (1980), mentioned that the variant with nasal assimilation (of Javanese origin) is used in variation with the variant with [ŋə-] (of Sundanese or Bangka Malay [Sd/BM] origin). We should recall from our discussion in Section 3 that nasal assimilation in JI shows the same morphophonological form as Javanese, while the variant with [ŋə-] might come from Sd/BM. The absorption of nasal assimilation to Betawi most probably took place around the seventeenth century when Javanese as an ethnic group was first recorded in *Dagh-Register* (1673), as discussed by Castle (1967). *Dagh-Register* reported that there were around 5,000 Javanese, i.e., 15% of the total population (32,068 people), in Jakarta at that time. In that case, the Javanese influence might contribute to the early absorption of nasal assimilation into the Betawi community.

However, the current Javanese variant used by present-day JI speakers is not only an inheritance of the same variant that characterized Betawi – it was furthered and also continues to be furthered by current Javanese influence. JI speakers, especially those of higher educational background, use the Javanese variant more frequently than Sd/BM variant. The choice of Javanese variant over Sd/BM variant by speakers of higher educational background is most probably caused by Javanese tradition (and language)[11] that is often associated with the image of *priyayi* (royalty or nobility) class (see Zents 2015, Errington 1988, 1985 for further discussions related to the Javanese language and tradition). During Dutch colonial, those who entered the civil service mainly were the

[11] The Javanese language and tradition discussed here specifically refers to the one in the Mataram's center (south central Java), which Sutherland (1975) considers the homeland of *priyayi* culture, and does not include all other Javanese traditions and dialects.

priyayi, i.e., Javanese elites, who were educated in European-style institutions (Hagen 1997). Furthermore, Nordholt (2011) mentioned that around 100.000 Dutch-trained administrators occupied higher echelons in the 1950s. This traditional *priyayi* practice was further supported by the new order regime under President Soeharto (Cribb & Kahin 2004).

Considering that, speakers of Javanese background often occupy a prominent place in Jakarta, the national capital. The Javanese language likely has prestige, and their version of Indonesian heavily favors the assimilated variant. This is an assumption based on my observation and needs to be tested in another study by examining on-going Indonesian speech on the part of speakers of Javanese ethnicity—especially, people born and raised in Jakarta but in families that had close connections to the Javanese homeland and who still used Javanese at home. The results in this study suggest that the variant with nasal assimilation is a prestige form that is associated with the Javanese, who predominate in the highest social echelons of Jakarta, and is increasing in usage at the expense of [ŋə-] associated with Betawi speech. Or possibly this is a matter of "negative prestige", where the variant with [ŋə-] has acquired a connotation of belonging to a masculine group. If this prediction is correct, it is interesting to see how the prestige use of a variant in the current investigation is influenced not by the standard variety, i.e., Standard Indonesian, but by another local variety, namely Javanese. This is different from the discussion in Trudgill (1972), where prestige is usually associated with standard varieties.

To summarize, the two observed variants have different sources and result from different influences. The preference for nasal assimilation over [ŋə-] seems to reflect social prestige and influence from Javanese language. We have shown that the variation between the choice of nasal assimilation as opposed to [ŋə-] prefixation is at least partly conditioned by educational attainment and gender: in general, the variant with nasal assimilation has a higher frequency among all groups. Further, females and people of higher education lead—i.e., the choice of the variant with nasal assimilation has higher frequency among females than males and similarly, the choice of the variant with nasal assimilation has a higher frequency among those of higher educational attainment than those of lower attainment.

References

Angle, John & Sharlene Hesse-Biber. 1981. Gender and prestige preference in language. *Sex Roles 7*, 449–461. https://doi.org/10.1007/BF00288072

Castle, Lance. 1967. The Ethnic Profile of Djakarta. *Indonesia, 3,* 153-204.

Cole, Peter, Gabriella Hermon, & Yassir N. Tjung. 2006. Is there pasif semu in Jakarta Indonesian? *Oceanic Linguistics,* 45, 65-90.

Conners, Thomas J., John Bowden, & David Gil. 2013. Valency Classes in Jakartan Indonesian. In A. Malchukov & B. Comrie, (eds.), *Valency Classes in the World's Languages*. 941-986. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110429343-003

Conners, Thomas J. & Claudia Brugman. 2013. The event semantic role of the nasal prefix in Jakarta Indonesian. *Proceedings of the Twelfth International Conference on Austronesian Linguistics*. Canberra: Pacific Linguistics.

Cribb, Robert & Audrey Kahin. 2004. *Historical Dictionary of Indonesia.* Lanham, MD: Scarecrow Press.

Dagh-Register. 1673. *Batavia* (27-30).

Dardjowijojo, Soenjono. 1978. *Sentence Patterns of Indonesian.* Honolulu: University Press of Hawaii.

Errington, J. Joseph. 1988. *Structure and Style in Javanese : a Semiotic View of Linguistic Etiquette.* Philadelphia: University of Pennsylvania Press.

Errington, J. Joseph. 1985. *Language and Social Change in Java: Linguistic Reflexes of Modernization in a Traditional Royal Polity.* Athens, Ohio: Ohio University Center for International Studies.

Gil, David, Uri Tadmor, John Bowden & Brad Taylor. 2015. Data from the Jakarta Field Station, Department of Linguistics, Max Planck Institute for Evolutionary Anthropology, 1999-2015.

Grijns, Cornelis D. 1991. *Jakarta Malay I: A multidemensional approach to spatial variation*. Leiden: KITLV Press.

Hagen, James. 1997. Read All about It: The press and the rise of national consciousness in early twentieth-century Dutch East Indies society. *Anthropological Quarterly*, 70(3), 107-126.

Hidajat, Lanny. 2010. *The acquisition of verb argument structure in basilectal Jakarta Indonesian.* Doctoral dissertation. University of Delaware.

Ikranagara, Kay. 1980. Melayu Betawi Grammar. *NUSA: Linguistics Studies in Indonesian and Languages in Indonesia,* 9.

Kurniawan, Ferdinan. 2018. *Phonological Variation in Jakarta Indonesian: an emerging variety of Indonesian.* Doctoral dissertation. Cornell University.

Kurniawan, Ferdinan. 2016. Nasal assimilation and substitution in Standard Indonesian: evidence from loanwords speech production task. *Linguistik Indonesia*, 34(1), 31-56.

Kurniawan, Ferdinan. 2015. Nasal Assimilation in Jakarta Indonesian. *Proceeding of Austronesian Formal Linguistics Association (AFLA) 21.* Canberra: Australia National University. 149-165.

Labov, William. 2001. *Principles of Linguistic Change: Social Factors*. Oxford: Blackwell.

Lapoliwa, Hans. 1981. A Generative Approach to the Phonology of Bahasa Indonesia. *Pacific Linguistics,* D(34).

MacDonald, Ross R. & Soenjono Dardjowidjojo. 1967. *Indonesian Reference Grammar*. Washington, D.C.: Georgetown University Press.

Mallinson, Christine. 2007. Social Class, Social Status and Stratification: Revisiting Familiar Concepts in Sociolinguistics. *University of Pennsylvania Working Papers in Linguistics*, 13(2). https://repository.upenn.edu/pwpl/vol13/iss2/12

Muhadjir. 1981. Morphology of Jakarta Dialect, Affixation and Reduplication. *NUSA: Linguistics Studies in Indonesian and Languages in Indonesia,* 11.

Nordholt, Henk Schulte. 2011. Indonesia in the 1950s: Nation, Modernity, and the Post-Colonial State. *Bijdragen tot de Taal-, Land- en Volkenkunde*, 167(4), 386-404.

Nothofer, Bernd. 1995. The History of Jakarta Malay. *Oceanic Linguistics,* 34(1), 86-97.

Pater, Joe. 2001. Austronesian nasal substitution revisited. In L. Lombardi (ed.), *Segmental phonology in Optimality Theory: Constraints and Representations*. 159-182. Cambridge: Cambridge University Press.

Pater, Joe. 1999. Austronesian Nasal Substitution and Other NC Effects. In Rene Kager, Harry van der Hulst, and Wim Zonneveld (eds.), *The Prosody Morphology Interface*. 310-343. Cambridge: Cambridge University Press.

Prince, Alan & Paul Smolensky. 2004. *Optimality Theory: constraint interaction in generative grammar*. Malden: Blackwell.

R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Sutherland, Heather. 1975. The Priyayi. *Indonesia*, 19(19), 57–77.

Sneddon, James, N., Michael C. Ewing, K. Alexander Adelaar & Dwi Noverini Djenar. 2010. *Indonesian: A comprehensive grammar (2nd ed.)*. London: Routledge.

Sneddon, James N. 2006. *Colloquial Jakartan Indonesian*. Canberra, ACT: Pacific Linguistics.

Sumukti, Rukmantoro H. 1958. *Sundanese Phonemes*. Master thesis. Indiana University.

Trudgill, Peter. 1972. Sex, Covert Prestige and Linguistic Change in the Urban British English of Norwich. *Language in Society,* 1(02), 179-195.

Wallace, Stephen. 1976. *Linguistic and Social Dimension of Phonological Variation in Jakarta Malay.* Doctoral dissertation. Cornell University.

Wolff, John U. & Soepomo Poedjosoedarmo. 1982. Javanese Communicative Codes. *Data Papers Series 116*, Ithaca, NY: Southeast Asia Program, Department of Asian Studies, Cornell University.

Wolff, John U. 1986. *Formal Indonesian*. Ithaca: Southeast Asia Program Publications, Cornell University.

Wolff, John U. 2010. *Proto-Austronesian Phonology with Glossary*. Ithaca: Southeast Asia Program Publications, Cornell University.

Zentz, Lauren. 2015. 'Love' the local, 'use' the national, 'study' the foreign: Shifting Javanese Language Ecologies in (Post-)Modernity, Postcoloniality, and Globalization. *Journal of Linguistic Anthropology*, 24(3), 339-359.

Zuraw, Kie. 2010. A model of lexical variation and the grammar with application to Tagalog nasal substitution. *Natural Language and Linguistic Theory*, 28(2), 417-472.