

Phonotactics and phonetic context in the perception of onset nasality in Taiwanese

Sheng-Fu Wang
New York University

1 Introduction

This paper reports an experiment that was designed to test the role of phonotactics and phonetics in Taiwanese speakers' perception of nasal contrast in voiced onsets. The empirical domain is the CV phonotactics in Taiwanese¹. In Taiwanese, a voiced onset and a vowel have to agree in nasality. In other words, there is no surface contrast between voiced stops (b, l, g)² and nasal stops (m, n, ŋ). The same restriction does not occur in French and Portuguese, both of which have phonemic nasal vowels just like Taiwanese. The patterns in Taiwanese and French/Portuguese are schematized in (1) and (2).

(1) Taiwanese	(2) French
p ^h a p ^h ã	pa pã
pa pã	ba bã
ba *bã	ma mã
*ma mã	

Studies have shown that Taiwanese nasal vowels have full nasality, whereas the counterparts in French and Portuguese have a slight delay in nasality (Chang et al., 2011; Hsieh, 2011; Delvaux et al., 2008; Parkinson, 1983). In light of these findings, I have proposed an analysis based on constraints on contrast (Wang, 2017), following the Dispersion-Theoretic framework (Flemming, 2006, 2008). The analysis is based on the hypothesis that whether a language has a surface contrast between voiced oral and nasal stops is crucially related to whether the language has full nasality in nasal vowels. For example, [b]-[m] is a phonemic contrast in French, and the analysis aims to connect the presence of such a contrast to the tendency for French nasal vowels to have delayed nasality. On the other hand, [b]-[m] is an allophonic distinction predictable from vowel nasality in Taiwanese, and the analysis tries to bridge this fact with the phonetic findings that Taiwanese nasal vowels tend to have full or almost full nasality.

The main point of the analysis is that since voiced oral stops have a weaker burst than voiceless stops (Zue, 1976), following from the fact that intraoral air pressures are lower for voiced stops (e.g., Slis and Cohen 1969; Lisker 1970; Warren 1976), they are already perceptually closer to nasal stops. If the following nasal vowel has full nasality, the discrimination between voiced oral and nasal stops would be more difficult, and thus the contrast is less likely to surface. Following this main point, in the analysis, full vowel nasality is the key correlate of the absence of nasal contrast in voiced stops in Taiwanese. The phonological grammars for Taiwanese and French are identical except for phonetic realization: Taiwanese requires full nasality for nasal vowels, while French does not. This phonetic difference derives different sets of attested syllable types.

This analysis makes some testable predictions. It predicts that Taiwanese speakers should improve their discrimination of voiced oral and nasal stops given a 'French-like' vocalic context that provides better

* I am grateful to Gillian Gallagher, Maria Gouskova, Lisa Davidson, Juliet Stanton, Scott Seyfarth, Ildi Szabó, Yining Nie, the audience at NYUs PEP Lab, and the audience at AMP 2016 & 2017 for discussions and comments on this project.

¹ 'Taiwanese' is the common term that is used to refer to the variety of the Southern Min (a Sino-Tibetan language also known as Hokkien and Minnan) language spoken in Taiwan. In the literature, the same variety is also referred to as 'Taiwanese Southern Min', 'Taiwanese Hokkien', 'Taiwanese Minnan', etc. The term 'Amoy' or 'Xiamen' refers to a very similar variety that is spoken around the Xiamen, Fujian area in South East China.

² The alveolar 'stop' in Taiwanese is essentially a lateral sound.

cues. Taiwanese speakers' struggle with discriminating voiced oral ([b,l,g]) and nasal stops ([m,n,ŋ]) in different contexts has been well documented (Pan, 2004; Wang, 2003). Following the Dispersion Theoretic Analysis, where the distinctiveness of these sounds before a nasal vowel is related to the nasality of the vowel, if the nasal vowel has a delay in nasality like in French, Taiwanese speakers should at least be better in telling voiced oral and nasal onsets apart. A competing prediction is based on the assumption that language-specific phonotactics dictates perception, and the role of phonetic detail is minimal. This is also the prediction that follows an analysis of the typology according to different rankings of AGREE[NASAL], *V̄, and IDENT[NASAL]. These would predict that Taiwanese speakers' perception of the distinction between voiced oral and nasal onsets do not change according to the amount of contextual phonetic cues.

This current paper reports an ABX perception experiment that was designed to test these predictions. The results of the experiment show that while some phonetic differences in the stimuli influenced participants' performance, the presence of delayed nasality did not. The results on the manipulated variable are not consistent with the prediction based on the Dispersion Theoretic analysis, although the observed effects of phonetic differences show that phonotactics does not predict all aspects of participants' performances in the experiment. The experiment will be described in detail in the remaining of this paper.

2 Literature review

This section reviews two experimental studies on the distinction between voiced oral and nasal onset in Taiwanese. I start with Wang (2003), who conducted AX experiments to test whether Taiwanese speakers could perceive the oral-nasal contrast in the onset and in the vowel. For testing the nasal contrast in the onset, he presented the participants with pairs of disyllabic real words (e.g., [mã⁵⁵ tso⁵¹] 'Mazu (place name)' vs. [bo³³ tsiŋ²⁴] 'loveless') and asked the participants whether the 'first sounds' were the same or different. By 'first sounds', the experimenter was trying to get the participants to compare the onsets of the first syllables for words in a pair. For testing nasal contrast in vowels, a similar method was used, with the target stimuli having an oral-nasal contrast in the vowel of the first syllable instead (e.g., [mĩ²¹ sũã²¹] 'misua (a type of salted noodle)' vs. [li⁵⁵ iu²⁴] 'reason/excuse'), and the question was whether the 'second sounds' were the same. Results are summarized in (3).

(3) Wang (2003) Results

segment	contrast	percentage of 'same' response
onset	b/m:	58%
	l/n:	32%
	g/ŋ:	32%
vowel	V/V̄	74%

For nasal distinctions in the onset position, different places of articulation had different results: a pair of labial voiced oral and nasal stops was more likely to be considered the same than the alveolar and velar pairs.

As for the results on vowels, surprisingly, vowels differing in nasality³ were judged to be the same 74% of the time. The author took it as a motivation to question whether the nasal contrast is phonemic at all in Taiwanese. Assuming that nasality may be non-contrastive is not a trivial matter: it would lead to great difficulty in accounting for the contrast between oral and nasal vowels in syllables that have voiceless onsets. It would take results from experiments that control for all potentially relevant factors to legitimately challenge the view that nasality is contrastive.

The paper also reported results for filler pairs such as [si]-[su] and [ki]-[ku]. As a reminder, in the onset discrimination task, the participants had to tell whether the 'first sound' in each syllable in a pair was the same. Results on these pairs showed that onsets within such pairs (e.g., [s] in [si] and [su]) were considered the same only 38% and 52% of the time respectively, despite the fact that the difference should only reflect a phonetic difference, not a phonemic one. In other words, as the [s] sound with different coarticulations seems to be perceived as 'different sounds' for the participants. The participants' performance on the filler items made the results on nasal contrast in the onset difficult to interpret.

³ The vowels were mostly different in tones as well, since it was not controlled for.

Pan (2004) conducted four experiments that aimed to show that the oral voiced stops [b, l, g] are indeed in the same phonemic categories as their nasal counterparts [m, n, ŋ]. With two articulatory experiments that looked at nasal airflow, she found that following a heterosyllabic nasal coda, voiced oral stops are nasalized, as evidenced by greater nasal airflow during the stop and at the onset of the following oral vowels. For example, in a sequence like [kĩm#bo], [b] had high volume of nasal airflow, as well as the onset of the oral vowel [o].

Pan also conducted two perception experiments. The first one was a ‘concept formation’ task, where the participant had to figure out the association between abstract categories and stimuli with different sets of onsets. Results showed that concept formation was the easiest when [b] and [m] belonged to the same category and voiceless onsets belonged to the other one. Placing [b] itself in a category and voiceless onsets in the other elicited worse performances, and participants performed the worst when one concept was associated with [b], and other concept was associated with [m] along with other voiceless stops. This was taken as evidence that Taiwanese speakers put [b] and [m] in the same category. The second perception experiment was a gating task, where [b] and [m], which were followed by oral and nasal vowels respectively, appeared in the second syllable of the stimuli. Results showed that the participants needed to hear the onset of the oral and nasal vowels to disambiguate the oral syllable from the nasal syllable. Again, this was taken as evidence that voiced oral stops and nasal stops are not in a phonemic contrast themselves.

To sum up, the Pan study used attested patterns in Taiwanese to show that the voiced oral stops and nasal stops are not contrastive, either indirectly based on articulatory properties or directly based on the difficulty of Taiwanese speakers in perceiving the differences of these sounds. It provides a baseline and set up stages for exploring whether Taiwanese speakers can perceive the difference in different phonetic contexts, which is the target of the present study.

In order to directly observe how Taiwanese speakers discriminate this non-phonemic contrast, this current experiment differed from these two studies by putting voiced oral stops and nasal stops in exactly the same vocalic context in some of the experimental conditions, including conditions where the voiced oral stops are illicit. Pan’s (2004) gating experiment fits the description that voiced oral stops and nasal stops were in the same context: up until the vowels following those stops were revealed, they could be described as being in the same context. However, it should be noted that oral and nasal onsets were always followed by oral and nasal vowels respectively in Pan’s study, thus the experiment only showed how and when speakers resolve ambiguity, but not how they dealt with illicit patterns.

3 Method

3.1 Overall Design & Predictions The experiment followed an ABX paradigm. In the experiment, the participants answered whether the third item in a sequence of stimuli was identical with the first or the second item. There were three types of AB pairs, which can be schematized as [ba]-[mã], [b^aã]-[mã], and [b^aã̃]-[mã]. [ba]-[mã] refers to an attested contrast in Taiwanese where the whole syllables contrast for nasality. [b^aã]-[mã] refers to a syllable with an voiced oral onset and a nasal vowel with delayed nasality, and [b^aã̃]-[mã] refers to a syllable with an voiced oral onset and a nasal vowel with delayed nasality. The contrast between these two types of syllables with the nasal syllable [m] is the critical target of this experiment.

There are two predictions on this critical target of comparison, which I call ‘phonetic context’ and ‘phonotactics’ predictions respectively. The former predicts that Taiwanese speakers should improve their performance on the [b]-[m] contrast in the nasal context if the nasal vowel following [b] has a delay of nasality. The latter predicts that phonetic detail of vowel nasality should not matter because this contrast is unattested for Taiwanese speakers. The predictions are summarized in (4).

(4) Two main predictions on main manipulation of the experiment

Dispersion-Theoretic analysis: [ba]-[mã] > *[b^aã]-[mã] > *[b^aã̃]-[mã]
 Phonotactics : [ba]-[mã] > *[b^aã]-[mã] = *[b^aã̃]-[mã]

As controls, I also included ABX sequences the oral onset is a voiceless aspirated or unaspirated stops, which are contrastive to nasal stops in both the nasal and the oral vocalic contexts. Including these controls allowed us to see a more general phonetic effect within the attested contrasts: Do additional differences between the oral and the nasal onsets in voicing and aspiration affect the discrimination performance between oral and nasal onsets?

There are again two predictions on the effect of phonetics. The first one, ‘the phonetic cue prediction’, is that among the attested contrasts, the distance between the oral and the nasal onsets in terms of phonetic cues (and the differences in phonological features) makes a difference. The specific prediction is that the contrast between a voiceless aspirated stop and a nasal stop should be the easiest, followed by the contrast between a voiceless unaspirated stop and a nasal stop, and the contrast between a voiced and a nasal stop. Also, since the alveolar voiced ‘stop’ in Taiwanese is a lateral sound, there is an additional prediction that the difference between [l] and [n] may be harder to distinguish than between [b] and [m]: [l] and [n] often participate in synchronic alternations or diachronic sound change, for example, n-lateralization and l-nasalization are observed in Korean (Davis and Shin, 1999). It is also not uncommon for laterals and nasals to be in the same phonologically active natural class (Mielke, 2005). Another prediction, the ‘phonotactics prediction’, is that the participants would be equally good at discriminating attested contrasts regardless of the types of segments that are in the stimuli. The predictions are schematized in (5).

(5) Predictions on other phonetic dimensions:

Phonetic cue prediction: [p^ha]-[mã] > [pa]-[mã] > [ba]-[mã]
 [p^hã]-[mã] > [pã]-[mã]
 [b]-[m] > [l]-[n]

Phonotactics prediction: [p^ha]-[mã] = [pa]-[mã] = [ba]-[mã]
 [p^hã]-[mã] = [pã]-[mã]
 [b]-[m] = [l]-[n]

3.2 Stimuli The stimuli that were used to form ABX sequences were disyllabic nonce words. The first syllable, where the target contrast occurred (e.g., /ba²⁴/ vs. /mã²⁴/), was always an attested combination of syllable and tone, and the second syllable was always an unattested combination.

To have a consistent quality of the vowel as much as possible throughout its duration, only monophthongs were used in the stimuli. Furthermore, only the pairs [a]-[ã] and [i]-[ĩ] appeared, as these pairs of vowels differ the least between dialects of Taiwanese.

The critical contrasts that occurred in the first syllable of stimuli were [b]-[m] and [l]-[n]. There were control sets, where voiceless onsets [p, p^h, t, t^h] appeared, so that the targeted segmental contrasts were [p]-[m], [p^h]-[m], etc. There were also fillers, where the targeted contrasts were either a difference in both the onset and the vowels (e.g., [pi⁵⁵]-[p^hĩ⁵⁵]), only in the onset (e.g., [p^ha⁵⁵]-[pa⁵⁵]), or only in the nasality of the vowel (e.g., [pi⁵⁵]-[pĩ⁵⁵]). In total, there were nine critical pairs, seven control pairs, and eight filler pairs.

For the critical and control stimuli, three types of syllable comparisons were constructed. One of the members in each comparison was always a nasal onset followed by a nasal vowel (e.g., [mĩ]), which is an attested form in Taiwanese. The other member in the comparison always had an oral onset, followed by three types of vowels: a fully oral vowel, a nasal vowel with an oral onset, and a fully nasalized vowel. Thus, when the nasal member in a pair was mĩ, the three types of syllables occurring as the ‘oral’ member of the comparison can also be described as [ba], [b^aã], and [b^ãã]. Among them, [ba] is the attested form in Taiwanese, while [b^aã], and [b^ãã] are not.

The sound files for the stimuli were created from the recording of a male native speaker. The two unattested types of syllables, namely the [b^aã]-type and the [b^ãã], were created from manipulating the recordings of attested [ba]- and [mã]-type syllables in the following way. The [b^ãã]-type syllables were created by splicing the vowel of a [mã]-type syllable to a corresponding voiced stop. That is, the vowels from [mã] and [lã] were spliced to [b] and [l] respectively. The vowels of the [b^aã]-type syllables were created with the following steps. First, the initial portion (six pitch cycles, which amounted to roughly 10% of the vowel duration) of corresponding oral and nasal vowels carrying a matching tone (e.g., /i/ from /bi²⁴/ and /ĩ/ from /mĩ²⁴/) were extracted. These two portions of vowels were manipulated so that they had matching amplitude and fundamental frequency. A continuum of vowel portions with varying degrees of nasality was created by mixing the oral and the nasal vowel with varying oral-nasal amplitude ratios. The 90% nasalized vowels were then created by splicing the first pitch cycle of a 0% nasal token, the second pitch cycle of a 20% nasal token... up to the fifth pitch cycle of a 80% nasal token. These five pitch cycles with nasality increasing from one cycle to another replaced the initial five pitch cycles of the original nasal vowel. The length of the spliced five pitch cycle was roughly 10% of the vowel duration.

The filler sound files were generated similarly. The only difference is that each filler contrast (e.g., [pi⁵⁵] vs. [p^hĩ⁵⁵]) only constituted exactly one pair of comparison instead of three in the critical and the control items.

For the nonce-word second syllable, four lexical gaps were chosen ([sa²⁴, ka²⁴, t^hu³³, k^hu⁵²]) and used throughout the whole stimulus set. These gap syllables were used in the following way: Any given comparison (e.g., [bi²⁴] vs. [mī²⁴]) appeared in four types of ABX sequences, and each of these four gaps would appear as the second syllable in one ABX sequence.

The combinations were arranged so that any given syllable gap did not occur only within a given ABX sequence type. For example, [sa²⁴] did not always occur within the ABB sequence. This arrangement added up the overall variability in the stimuli and aimed to make it more difficult for the participant to track a given type of nonce word to form a learning strategy. For example, if [bi²⁴] was always followed by [t^hu³³], they might be able to remember [bi²⁴-t^hu³³] as a word and then form hypotheses about how to answer it. The experiment ended up having 224 ABX sequences (9 critical sets × 3 comparisons × 4 ABX types + 7 pseudo filler sets × 3 comparisons × 4 ABX types + 8 filler pairs × 4 ABX types). The inter-stimulus interval in an ABX sequence was 250 ms.

3.3 Participants The participants were forty speakers of Taiwanese. They were self-reported to be both native and fluent speakers of this language. They were recruited in Taipei, where the experiment was conducted. The age range was 18-30 years, and the participants were evenly split by gender.

3.4 Procedure The experiment was run with Experigen (Becker and Levine, 2010). Instructions on the screen were written in Mandarin Chinese, but the experimenter explained the experiment in Taiwanese on the spot. They were told that this was an experiment for testing how well the listeners could tell the difference between synthesized speech sounds.

During the experiment, the participant heard an ABX sequence that was played automatically and was instructed to answer by clicking on the button showing ‘1’ or ‘2’, referring to whether the participant thought the third item was identical to the first or the second item. Eight practice trials were presented first, during which the participant could listen to the ABX sequence as many times as they wanted. The practice items were randomly drawn from the filler items. The main experiment was broken down into four blocks each with 56 items. The items were presented randomly but arranged in a way that different ABX sequences of a given pair of comparison never appeared in the same block. For example, target-syllable sequences like [bi]-[mī]-[bi] and [bi]-[mī]-[mī] never appeared in the same block.

3.4.1 Statistical Analysis The results were filtered by excluding participants whose percentages of correct responses in the attested [ba]-[mā] and [la]-[nā] contrasts were below 68.75% (22 correct answers out of 32 items in this condition). According to the exact binomial test, success rate at and above this threshold has a true probability of success that is greater than 0.5. Under this criterion, six participants were excluded, leaving data from 34 participants for analysis. Filler items were also excluded from the analysis.

The data on the remaining items were analyzed with a generalized logistic linear mixed-effects model. Fixed effects included onset voicing/aspiration, place of articulation, and vowel nasality. The fixed effects were treatment coded following R’s default. The baseline level was ‘voiced onset’ for onset voicing, ‘labial’ for place of articulation, and ‘100% nasality’ for vowel nasality.

As for random effects, a group-level intercept and slopes for the main effects of onset voicing/aspiration, place of articulation, and vowel nasality and their interactions were fit for each participant, and group-level intercepts and slopes for vowel nasality for each item are were included.

Pair-wise comparisons between cell means were conducted with `lsmeans()` and `contrast()` from the `lsmeans` package in R (Lenth and Hervé, 2014), which adjusts the p-values with the Tukey method for a family of 18 estimates. Comparisons between cells that differ along each of the fixed effects will be reported and discussed.

3.5 Results

3.5.1 Distribution of individual performances Figure 1 presents the distribution of individual performances as a function of voicing/aspiration and vowel nasality of the items. The distribution of individual performance on the two unattested conditions, voiced onset with 100% nasal vowel and voiced onset with 90% nasal vowel, appear similar. The most striking difference is the difference between conditions with voiceless onsets and voiced onsets; the participants’ performance is better on items with voiceless onsets,

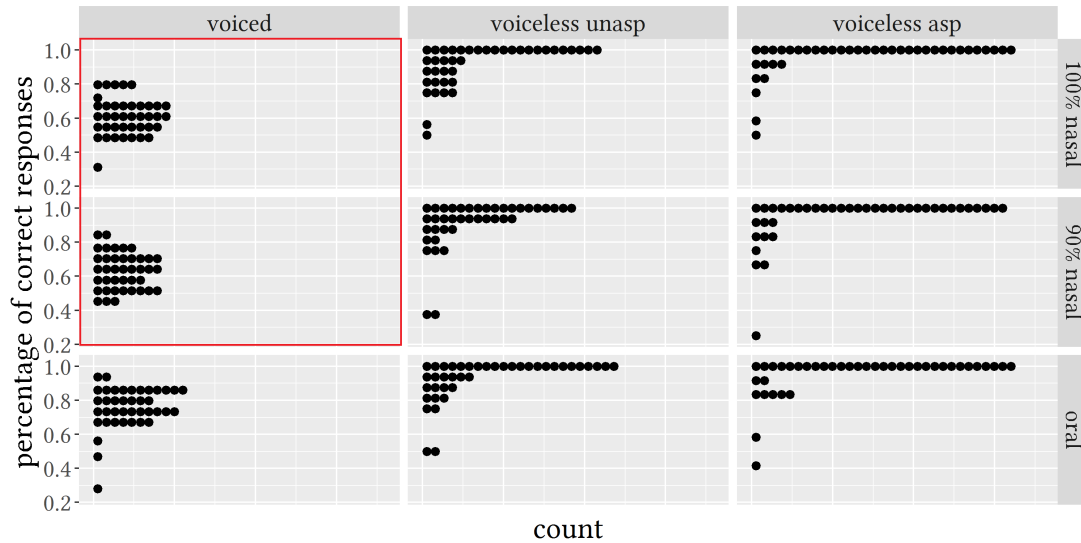


Figure 1: Distribution of individual performance as a function of voicing/aspiration and vowel nasality. In each panel, each dot presents a participant. Y axis refers to the overall percentage of the correct response of an individual participants and each step on the axis represent a 5% interval. Horizontal panel labels refer to conditions of voicing/aspiration. Vertical panel labels refer to conditions of vowel nasality following the oral onset. The two conditions that are highlighted are the distinctions that are not contrastive in Taiwanese, i.e., distinctions between voiced oral and nasal stops before a nasal vowel.

especially voiceless aspirated onsets. Another notable distribution can be found in the voiced onset with oral vowel condition ([ba]-[mã]): this is a phonotactically attested distinction in Taiwanese. However, the participants' performance is visibly worse in this condition.

3.5.2 Overall results The regression model yielded the main effects of vowel nasality and onset voicing/aspiration, as well as multiple two-way and three-way interaction effects. The estimates log odds for all conditions are plotted in Figure 2.

In what follows in this section, the presentation of results will focus on pair-wise comparisons of cell means based on the outputs of *lsmeans()* and *contrast()* functions from the *lsmeans* package in R (Lenth and Hervé, 2014). The examination of pair-wise comparisons between cells not only makes sense for inspecting whether the critical target contrast (e.g., [bã_{100%}]-[mã] vs. [bã_{90%}]-[mã]) is significantly different, but also can be motivated from the significant three-way interactions reported in the previous paragraph. The *p* values were adjusted with the Tukey method for comparing a family of 18 estimates. The estimates are given on the log odds ratio scale.

I will first focus on items with voiced stops, whose variation along vowel nasality included the the target contrast in this study (i.e., contrasts like [b^aã]-[mã] vs. [b^ãã]-[mã]) and also a distinction between attested vs. unattested items (i.e., contrasts like [ba]-[mã] vs. [b^aã]-[mã] & [b^ãã]-[mã]). After looking at this set of stimuli, I will zoom out and discuss the effect of voicing/aspiration and places of articulation.

3.5.3 Voiced oral vs. nasal onsets in different vowel conditions The estimates of cell means for the items that compare a voiced oral stop to a nasal stop are shown in Figure 3. The crucial comparisons in this study are between the black (100% nasal, e.g., [b^aã]-[mã]) and the dark gray bars (90% nasal, e.g., [b^aã]-[mã]). The result on these comparisons for both places of articulation shows that Taiwanese speakers did not perceive the distinction between a voiced oral onset and a nasal onset differently when the following nasal vowel had different phonetic properties: whether the nasal vowel was fully nasalized or not did not

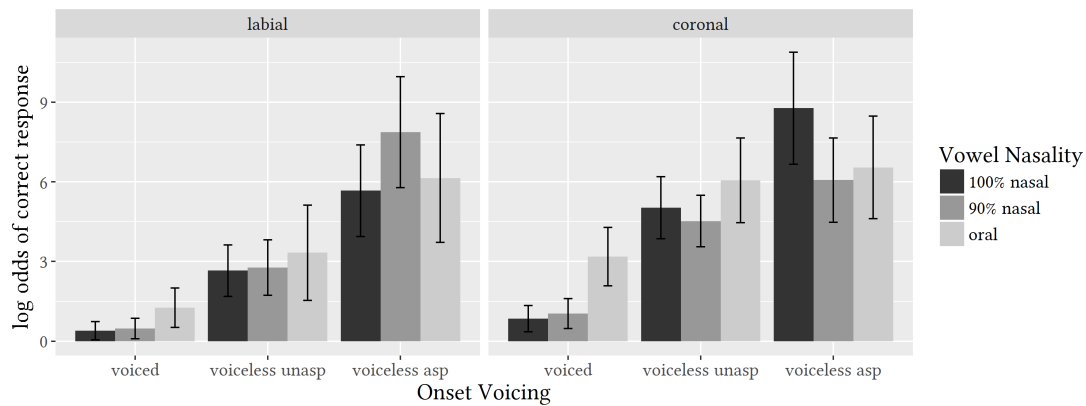


Figure 2: Overall result in estimated log odds of correct responses (the Y axis) as a function of onset voicing (the X axis), place of articulation (panel), and vowel nasality (color). The estimates and asymptotic confidence intervals are taken from a generalized linear mixed-effects logistic model.

make a difference in performance. The oral condition (e.g., ba-mã), represented by the light gray bars, serves as a control for items with voiced onsets, and some unexpected results can be seen in Figure 3: The attested contrast was only performed better in the coronal condition, but not in the labial condition.

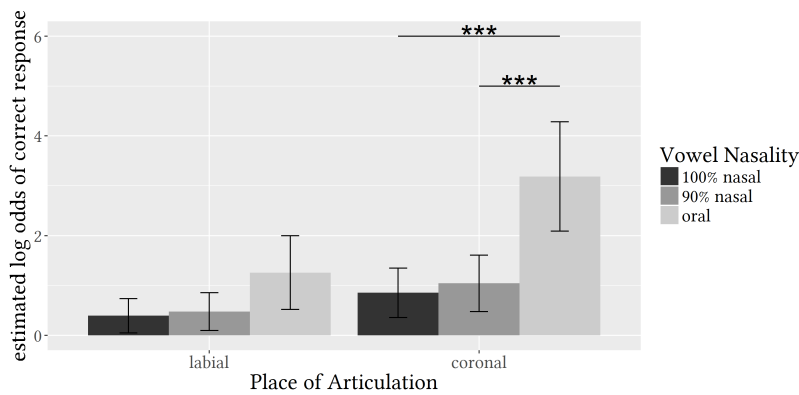


Figure 3: Results for items with a voiced onset. The Y axis shows estimated log odds of correct responses. The X axis shows place of articulation. Different colors show different conditions of vowel nasality. The error bars indicate asymptotic confidence intervals. Asterisks and lines indicate differences with statistical significance: $p < .05^*$, $p < .01^{**}$, $p < .001^{***}$

3.5.4 The effects of voicing/aspiration and place of articulation This section reports the results on the effect of voicing/aspiration and place of articulation. I start with the effect of onset voicing/aspiration, the overall results are plotted in Figure 4, with the neighboring bars showing difference conditions of voicing/aspiration.

The difference between voiceless aspirated and voiced onsets was found for all combinations of places of articulation and different conditions of vowel nasality, except for the condition where the onset was coronal and the vowel was oral (i.e., [t^ha]-[nã] ≈ [la]-[nã]). The difference between voiceless unaspirated and voiced onsets was significant when the vowel was nasal but not when the vowel was oral. Finally, the difference

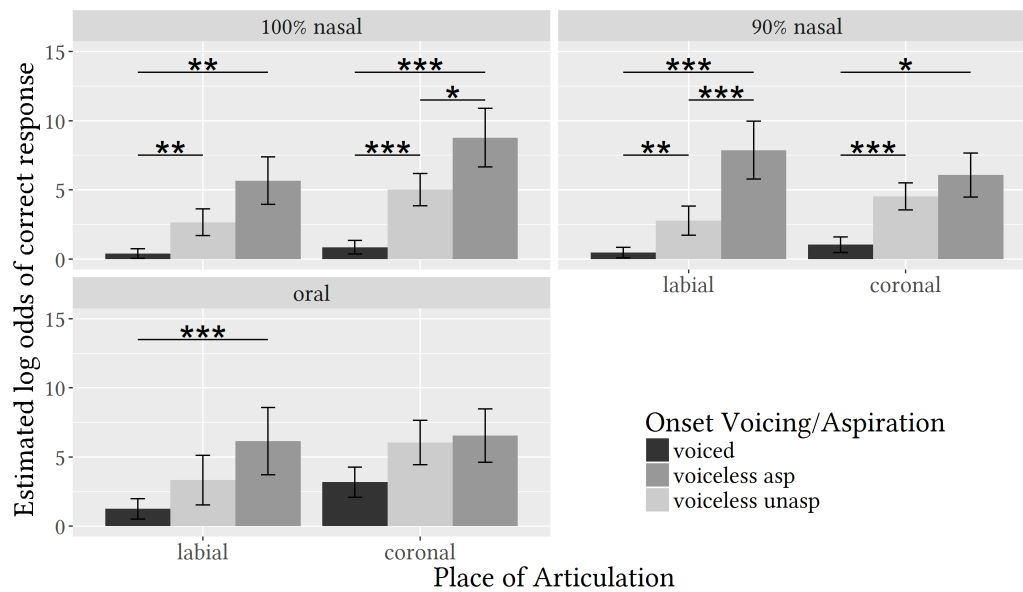


Figure 4: Overall results highlighting pair-wise comparisons for the effect of onset voicing/aspiration. Y axis shows estimated log odds of correct responses. The error bars indicate asymptotic confidence intervals. Different colors show different conditions of onset voicing/aspiration. X axis shows onset voicing/aspirations. Different panels show different vowel nasality conditions. Asterisks and horizontal lines indicate difference with statistical significance: $p < .05^*$, $p < .01^{**}$, $p < .001^{***}$

between voiceless aspirated and unaspirated onsets was different in conditions where the onset was lateral and the vowel was 100% nasal ($[t^{h\tilde{a}}]-[l\tilde{a}] > [t^{\tilde{a}}]-[l\tilde{a}]$) and where the onset was labial and the vowel was 90% nasal ($[p^{\tilde{a}}]-[m\tilde{a}] > [p^{\tilde{a}}]-[m\tilde{a}]$).

As for the effect of place of articulation, the overall results are plotted in Figure 5, but with the neighboring bars showing difference places of articulation. The only comparison that showed the effect of place of articulation is that the $[t^{\tilde{a}}]-[n\tilde{a}]$ contrast was performed better than the $[b^{\tilde{a}}]-[m\tilde{a}]$ contrast.

3.6 Discussion

3.6.1 Overall Performance One of the most striking observations from the results is that participants did not reach ceiling level in terms of their performances on the attested contrasts. A substantial number of the participants' performances were below 90% for the contrast between attested voiceless unaspirated onsets and nasal onsets (e.g., $[pa]/[p^{\tilde{a}}]/[p^{\tilde{a}}]$ vs $[m\tilde{a}]$). For the attested contrast between voiced oral onset with an oral vowel versus nasal onset with a nasal vowel (e.g., $[la]$ vs. $[n\tilde{a}]$), all but two subjects performed below 90%. In other words, the task was difficult for the participants at least for a subset of the items.

Since the subjects performed well for the voiceless aspirated items (although about one fourth of the participants did not perform at 100%), it was possible that the inclusion of these easier items, along with fillers, affected subjects' performance by shifting them to a higher standard as to whether a given pair of sounds was different, which in turn made the discrimination of acoustically similar items more difficult for them. This interpretation also suggests that the acoustic similarity affected participants' performance.

The reader may wonder whether the subjects' expectation about balance amounts of answers affected their performance. Since the task was an ABX task, balancing the answers would mean balancing between A and B. The balance should be easier done with easier items and fillers, thus the expected amounts of answers on the harder items should still be balanced. An expectation like this might affect their performance in a subtle way: if they had answered some items wrong, expectation of balancing A and B responses might

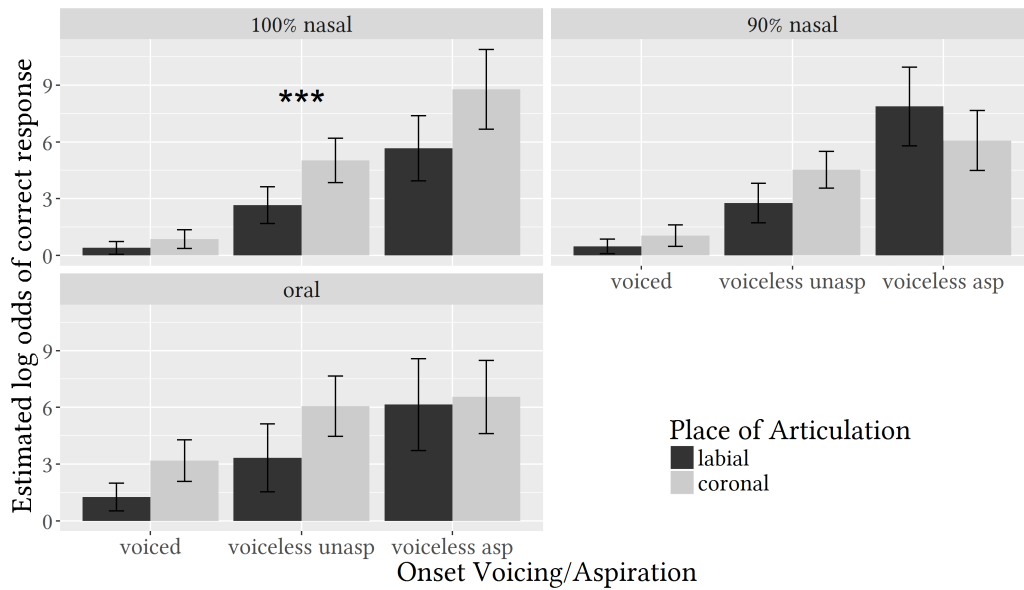


Figure 5: Overall results highlighting pair-wise comparisons for the effect of place of articulation. The Y axis shows estimated log odds of correct responses. The X axis shows onset voicing/aspirations. Different colors show different places of articulation. The error bars indicate asymptotic confidence intervals. Different panels show different vowel nasality conditions. The asterisks indicate a difference with statistical significance: $p < .05^*$, $p < .01^{**}$, $p < .001^{***}$

bias them towards the wrong answer again. However, since all items were interspersed and the same pair of comparisons were embedded in different non-word contexts, this is not a very likely scenario.

3.6.2 Voiced oral vs. nasal onsets in different vowel conditions The experiment aimed to see how well Taiwanese speakers could discriminate the non-phonemic contrast between voiced oral and nasal stops in different vocalic contexts. The critical context was the ‘French-like context’, where the oral stop was followed by a nasal vowel that had an oral onset. This is the context where two different hypotheses diverged in their predictions, which I repeat as follows: The phonotactics-based hypothesis predicted no improvement in this condition from the ‘Taiwanese context’ where an oral stop was followed by a fully nasalized vowel. On the other hand, the hypothesis based on a Dispersion-Theoretic analysis predicts that Taiwanese speakers should improve on the discrimination of this non-phonemic contrast in this condition.

The null result on the difference between the 90% nasal (‘French-like’) condition and the 100% nasal (‘Taiwanese-like’) condition was consistent with the prediction based on phonotactics. In light of this, the result can be interpreted by saying that Taiwanese speakers treated the 90% nasal vowels as they treated the 100% nasal vowels. Phonetic detail of this magnitude did not steer them away from phonemic processing, and the unattested difference between the [bã] type syllables and the [mã] syllables were difficult to the same extent for them.

One possibility that is at least partially consistent with the phonotactics hypothesis is that nasal contrast in a contrast like [ba]-[mã] in Taiwanese is syllable-based rather than phoneme-based. This suggests that during the ABX task, the participants were categorizing the whole syllable into oral or nasal when they decided on the matching between the stimuli. Thus, it was not important to them whether the onset of the vowel cued the voiced oral stop, as a large portion of the syllable was still nasal. This syllable-based categorical perception is consistent with the findings about ABX tasks: memories of the auditory traces may have faded away by the time the participant hear the X item, and they must rely on the categorical labels they have assigned to A and B (Gerrits and Schouten, 2004). Given that voiced oral stops and nasal stops are likely not two separate

categories (Pan, 2004) while oral and nasal vowels (and the oral [ba] and nasal syllables [mã]) are contrastive, it is possible that participants did the task based on categorical perception of vowel and thus syllable types.

Testing Taiwanese speakers with the same contrast in the context of an oral vowel (e.g., [ba] vs. [ma]) may help shed light on this possibility: If [ma] is conceived to be similar to [ba], it will strengthen the claim that the speakers use top-down processing based on phonotactics that are indifferent to phonetic details. On the other hand, If [ba] and [ma] can be distinguished better than [bã] and [mã], it suggests that nasal contrast in the onset position is distinguishable for Taiwanese speakers, but the acoustic difference is masked by syllable-based matching of oral and nasal items which follows from their experience with attested surface contrasts. It will also suggest that processing based on phonotactics and native sound categories is not absolute and is sensitive to the acoustic properties to a certain extent.

It should be noted that the result does not suggest a total dominant role of phonotactic knowledge in perception. It was possible that the vocalic cue for nasal contrast has to be augmented by cues from other dimensions, such as stop burst. Inspection of and comparison between French voiced stops and their Taiwanese counterparts revealed that the former have a much stronger stop burst than the latter, as shown in Figure 6. Thus it is possible that cues from the vocalic context was not enough to cue this nasal contrast in the onset position. This could be tested by manipulating bursts of onsets.

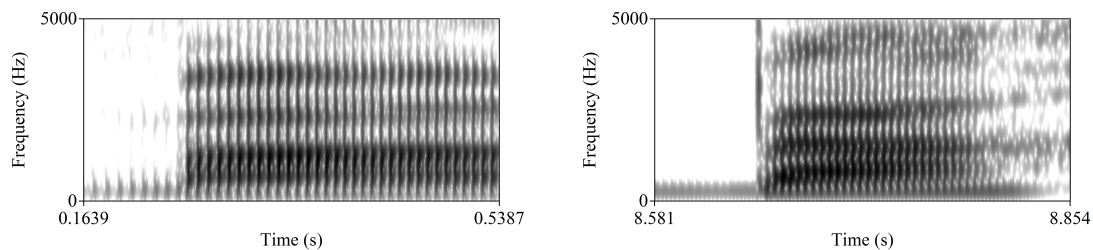


Figure 6: Spectrograms showing the difference in stop burst in Taiwanese (left: [ba⁵⁵] ‘father’) and French (right: [ba] ‘low’).

Another notable pattern from the results is the lack of expected phonotactic effect in the labial items: the attested contrast ([ba]-[mã]) was not distinguished significantly better than the unattested distinctions ([b^aã]-[mã] and [b^aã]-[mã]). On the other hand, the phonotactic effect was significant for coronal items ([la]-[nã] > [l^aã]-[nã] & [l^aã]-[nã]). Since voiced labial and coronal ‘stops’ also differ in manner of articulation, with the latter being laterals, this result indicates an asymmetry in both place and manner of articulation. One interpretation is that the difference between voiced oral and nasal stops is smaller than between laterals and nasal stops for Taiwanese speakers. This is surprising given the relationship between coronal laterals and nasals across languages. As mentioned earlier, [l] and [n] are predicted to be perceptually similar based on cross-linguistics phonological processes and distribution of natural classes. It remains to be seen whether this pattern can be replicated in other tasks. If so, we may need to put forward an interpretation based on cue saliency: because of the additional spectral energy during ‘closure’, lateral sounds may be more salient sounds themselves than non-lateral sounds such as voiced stops, which may make them more identifiable in a discrimination task.

3.6.3 The effect of voicing/aspiration and place of articulation In addition to the target distinctions, the inclusion of onset voicing/aspiration and place of articulation as factors made it possible to also examine effects of these factors and discuss possible implications concerning more general phonetic effects. As reported in the Results section, effects of onset voicing/aspiration and place of articulations were found. When the vocalic context was oral, where voiced and voiceless stops are attested, the participants performed better on items with a voiceless aspirated onset than items with a voiced onset. In other words, the participants distinguished pairs like [p^ha]-[mã] better than pairs like [ba]-[mã]. The difference between voiceless unaspirated and voiced stops was also significant: the participants distinguished pairs like [ta]-[nã] better than pairs like [la]-[nã]. It shows that within the attested contrasts, when the oral onset is more dissimilar from the nasal onset, the participants performed better, and it suggests effects that can be interpreted both in terms of phonetic detail and the difference of phonological feature values between sounds in a contrast.

As for place of articulation, there was only a mild effect of coronal items over labial items, which stands out when the vowel was 100% nasal and the onset was a voiceless unaspirated stop. The effect of place of articulation has already been mentioned earlier in this paper while discussing the expected phonotactic effect when the vowel was fully nasalized and the onset was voiced: it only showed up when the onset was coronal ([la]-[nã] > [lã]-[nã]) but not when the onset was labial. Again, it is surprising that coronal voiced stops, which are laterals in Taiwanese, are less confusable with nasals.

To conclude this section, the results suggest a mix of phonotactic effects (null result on 90% vs. 100% nasal vowels) and phonetic effects (voiceless aspirated > voiceless unaspirated > voiced onsets). The main point of the Dispersion-Theoretic analysis on the vocalic cue-licensing of onset nasal contrast was not supported, although some phonetic effects appeared in the direction that are consistent with Dispersion-Theoretic terms.

4 Conclusion

Before concluding this paper, I would like to have a final discussion on the phonetic effects in a more general way. It is interesting that the difference on the dimensions of voicing and aspiration showed an effect, while differences only in terms of nasality did not. Voicing and aspiration are contrastive at the segmental level in most contexts in Taiwanese, while nasality is only contrastive at the segmental level for vowels following voiceless onsets. One interpretation is that listeners pay attention to acoustic dimension that matter. For example, it is important to know that [p] and [p^h] are different, so cues that are related to the difference in aspiration are learned as important cues. In the experiment, even though the difference between [pã] and [mã] and the difference between [p^hã] and [mã] do not differ in status (both are lexical contrasts), the participants reacted to the latter in a certain way that helped their discrimination because they are accustomed to paying attention to aspiration. The same thing is true for voicing, but is not true for onset nasality and a 10% change of nasalized duration.

It shows that listeners pay more attention to cues that are important for lexical contrasts, and even in cases where the cues do not matter as much (i.e., aspiration is only one of the differences between [p^h] and [m]), the presence of such cues still makes a positive effect in perceiving a contrast. On the other hand, since the difference in onset nasality is never that important for lexical contrast, enhancing the cues for onset nasality does not make a difference in perception.

Summing up this discussion, it should be noted that whether a phonetic dimension is used for contrast may correspond to perceptual saliency of acoustic difference in that dimension, which I have tried to model with the perceptibility scales. Examining the correlation between perceptual saliency of cues, whether those cues are used for lexical contrasts, and whether and how those cues make an impact in perception in general is potential direction of for future research beyond the issues that are immediately related to this paper.

To conclude, the experimental result failed to show that increasing vocalic cues to oral-nasal contrast can improve Taiwanese speaker's perception of this non-phonemic contrast, although other phonetic effects were found. The results suggest a need to model multi-dimensional phonetic cues and gradient phonotactic effects in phonological grammar. More specific future directions of this research mainly include teasing apart of potential task-related factors in the current experimental paradigm. Including contrasts such as [ba]-[ma] may help shed light on whether the participants were really using the vocalic information in the ABX task, and whether the [+NASAL] and [-NASAL] cues are used in the same way. Excluding fillers and control items may change the participants criterion as to whether items are different, so that some expected phonotactic effect that is absent in the current experiment (e.g., [ba]-[mã] vs [bã]-[mã]) and the putative Dispersion Theoretic effect of vowel nasality may show up. More fine-grained manipulation of vowel nasality and the additional manipulation of stop release are also potential directions that may help identify the exact cues that are crucial to the nasality contrast.

References

- Becker, M. and Levine, J. (2010). Experigen: an online experiment platform. Available (April 2013) at <https://github.com/tlozoot/experigen>.
- Chang, Y.-c., Hsieh, F.-f., and Hsieh, Y.-l. (2011). Phonetic Implementation of Nasality in Taiwanese (and French): Aerodynamic Case Studies. In *Proceedings of the 17th International Conference of Phonetic Sciences*, pages 436–439.

- Davis, S. and Shin, S.-H. (1999). The syllable contact constraint in Korean: An optimality-theoretic analysis. *Journal of East Asian Linguistics*, 8(4):285–312.
- Delvaux, V., Demolin, D., Harmegnies, B., and Soquet, A. (2008). The aerodynamics of nasalization in French. *Journal of Phonetics*, 36(4):578–606.
- Flemming, E. (2006). The role of distinctiveness constraints in phonology. Unpublished manuscript, MIT.
- Flemming, E. (2008). The realized input. ms. Unpublished manuscript, MIT.
- Gerrits, E. and Schouten, M. E. (2004). Categorical perception depends on the discrimination task. *Attention, Perception, & Psychophysics*, 66(3):363–376.
- Hsieh, Y.-I. (2011). Nasal and Nasalized Vowels in Taiwanese: An Aerodynamic and Acoustic Study. MA Thesis, National Tsing Hua University.
- Lenth, R. V. and Hervé, M. (2014). lsmeans: least-squares means. r package version 2.13. URL <http://CRAN.R-project.org/package=lsmeans>.
- Lisker, L. (1970). Supraglottal air pressure in the production of English stops. *Language and speech*, 13(4):215–230.
- Mielke, J. (2005). Ambivalence and ambiguity in laterals and nasals. *Phonology*, 22(02):169–203.
- Pan, H.-h. (2004). Nasality in Taiwanese. *Language and speech*, 47(3):267–296.
- Parkinson, S. (1983). Portuguese nasal vowels as phonological diphthongs. *Lingua*, 61(2-3):157–177.
- Slis, I. H. and Cohen, A. (1969). On the complex regulating the voiced-voiceless distinction ii. *Language and speech*, 12(3):137–155.
- Wang, H. S. (2003). The Nasality of Voiced Stops and Vowels in Taiwan Min Revisited. *Journal of Taiwanese Languages and Literature*, 1:213–225.
- Wang, S.-F. (2017). A Dispersion-Theoretic account of Taiwanese CV phonotactics. In *Supplemental Proceedings of the 2016 Meeting on Phonology*.
- Warren, D. W. (1976). Aerodynamics of speech production. *Contemporary issues in experimental phonetics*, 30:105–137.
- Zue, V. W. (1976). *Acoustic characteristics of stop consonants: a controlled study*. PhD thesis, Massachusetts Institute of Technology.