Bulletin of the *Transilvania* University of Braşov
Series IV: Philology and Cultural Studies • Vol. 17(66) No. 2 – 2024
https://doi.org/10.31926/but.pcs.2024.66.17.2.1

Partial voicing neutralization in unstressed stop-nasal sequences

Andrei MUNTEANU[1], Younggun KIM[2], Kuilin LI[1]
Banjamin LeBRUN[1], Simon LIVOLSI[1], Jackson MUNDIE[1]

We present the results of a production experiment that explore the realization of voicing in English stop-nasal sequences, a phonological environment with a low functional load. The results imply a separation between primary and secondary voicing cues. Primary cues – aspiration and vocal fold vibration during stop closure – robustly distinguish underlyingly voiced stops from and underlyingly voiceless ones. Meanwhile, secondary cues – vowel duration and stop closure duration – are limited in their use by phonological position or absent entirely. A principal component analysis of the data indicates that all speakers occasionally produce tokens that are ambiguous in voicing cues.

Keywords: *phonological contrast, phonetic cues, functional load, voicing*

1. Introduction

The English voicing contrast is typically conceptualized as a change in a single phonological feature, e.g. [voice]. However, acoustic studies show that the contrast is actually expressed as a combination of phonetic cues, most commonly aspiration, but also pre-voicing, duration of the preceding vowel, duration of the stop closure, and fundamental frequency (Cole et al., 2007). As such, the difference between, for example, the /p/ in /ɹæpɪd/ 'rapid' and the /b/ in /ɹæbɪd/ 'rabid' can be signaled with longer voice onset time (aspiration), less pre-voicing, shorter preceding vowel /æ/, longer closure duration, or a combination of the aforementioned. The realization of the contrast also depends on stress, word position, phonological environment, as well as the stop's place of articulation (Davidson, 2016).

More broadly, it has also been shown that lexical contrast can influence production. In historical linguistics, it is often assumed that segmental contrasts (e.g. /p/ vs /b/) that split fewer minimal pairs in the language are more likely to be

[1] McGill University, andrei.munteanu2@mcgill.ca (corresponding author)
[2] Université de Montréal

lost over time, also known as the *functional load hypothesis* (Martinet 1955). While not uncontroversial (King 1967), some have found evidence in its favour in the history of specific languages, such as Korean (Silverman 2009), or English and Japanese (Ogura and Wang 2018), but also in large cross-linguistic studies (Wedel et al. 2013; Round et al. 2022). Data from synchronic phonology likewise implies that lexical contrast can shape language evolution. Languages occasionally appear to avoid homophonous forms within a morphological paradigm, either through the introduction of sporadic sound changes (Gessner and Hansson 2004), the suppression of regular sound changes (Crosswhite 1999), or by omission of the homophonous forms entirely, i.e. leaving the paradigm as defective (Baerman 2010). Although homophony avoidance is typically treated as arising historically through errors in transmission (Winter and Wedel 2016), some arguments for an active anti-homophony restriction in the synchronic grammar have also been proposed (Munteanu 2021).

As such, it is not surprising that contrast plays a large role in phonetics as well, where it is often assumed that the realization of segments is dependent on the segmental contrasts required by the language, most famously in Dispersion Theory (Flemming 2017). For example, it is often assumed that inventories with a greater number of segmental contrasts allow for less variability in production than inventories with fewer segmental contrasts, though evidence of this claim has been limited (Hauser, 2022). It has, however, been shown that, within a language, speakers tend to hyperarticulate cues that separate phonemes if there is a possibility of lexical confusion, i.e. a minimal pair, confirmed both in an experimental setting with the English voicing distinction (Schertz 2013), and in a corpus study with a variety of cues (Wedel et al. 2018).

In this paper, we investigate the realization of the voicing distinction in English stops in stop-nasal sequences. Stop-nasal sequences in Modern English are only possible in hetero-syllabic position (across a syllable boundary), e.g. [tɛk.nik] 'technique' and [æd.mɪn] 'admin'. Historically, nasal-stop sequences were also found in (tautosyllabic) onset position, e.g. Middle English [kniçt] 'knight'; however, the initial stop was eventually deleted and, as a result, only hetero-syllabic stop-nasal sequences remain. In Modern English, words with this property tend to be loanwords of Greek and Latinate origin, e.g. [stæg.nənt] 'stagnant' and [hɪp.no.sɪs] 'hypnosis' and are relatively rare in the language; a quick search through an English corpus (Baayen et al. 1995) reveals only 219 example words.

Crucially, the corpus of English (Bayeen et al. 1995) does not contain any voiced-voiceless minimal pairs for this phonological position. In other words, there is no lexical item in English (e.g. [pɪk.nɪk] 'picnic'), where articulating a pre-nasal voiced stop as voiceless or a pre-nasal voiceless stop as voiced would result in a different lexical item (e.g. *[pɪg.nɪk] 'pignic'). As such, speakers have little incentive to carefully articulate the voicing distinction in this environment, as there is little chance of misinterpretation. Note, however, that the contrast is evidenced elsewhere in the lexicon, e.g. [de.lɪ.kə.si] 'delicacy' vs [de.lɪ.gə.si] 'delegacy'.

Additionally, it could be argued that the expression of voicing in English word-medial coda stops is limited in general. Word-medial obstruent-obstruent sequences generally undergo assimilation, thus [əb̥steɪ̯n] 'abstain'. Meanwhile, word-medial obstruent-approximant sequences are assumed to syllabify as part of the onset of the following syllable, thus [ə.brʌpt] 'abrupt'. Therefore, it is rare for English to exhibit voiced-voiceless minimal pairs in word-medial coda position overall, though the contrast between voiced and voiceless stops is still maintained word-finally, c.f. [slæp] 'slap' and [slæb] 'slab'.

In summation, we expect voicing in English in general to be signalled using a number of cues (Lisker and Abramson 1967; Davidson 2016) with substantial variation, as is true of phonological contrasts in general (Schertz and Kang 2022). Furthermore, for the same contrast in stop-nasal sequences in particular, we expect hypoarticulation of the voicing cues resulting from the low functional load of voicing in this position. As such, productions of voiceless and voiced stops may exhibit partial or complete overlap in cues, more so than in other positions in the language.

2. Methodology

The experiment was a nonce word production experiment. Participants were asked to produce nonce words in the carrier phrase 'Say ____ again'. The experiment was conducted in a sound-attenuated room in the phonetics lab at McGill University (Montreal, Canada). Stimuli were presented orthographically using *PsychoPy* (Peirce et al. 2019).

2.1. Stimuli

Target stimuli consisted of English nonce words containing a stop-nasal sequence. The sequences were always word-medial and hetero-organic (articulated at different places of articulation), so as to ensure individual closures for the stop and nasal. Labial and velar stops were always followed by the coronal nasal: [pn], [bn], [kn], [gn]; coronal stops were flowed by the labial nasal: [tm], [dm]. In the orthographic prompts, the voiceless velar stop was represented by the English letter 'c'; the remaining letters matched their IPA equivalents.

Stimuli fell into one of two stress conditions, according to whether the stop was in a stressed syllable: **stressed** (where the stressed vowel immediately preceded the stop), and **unstressed** (where the stressed vowel immediately followed the nasal). Stress was controlled because it has been previously shown to interact with the realization of voicing (Lisker and Abramson 1967; Davidson 2016). Stress was not marked orthographically. Instead, to ensure that participants produced the intended stress pattern, stimuli were suffixed with phonologically charged suffixes. On the one hand, stimuli in the stressed condition always ended in *-{CN}ify*, where *{CN}* corresponds to the target sequence, e.g. *ste{bn}ify*, *ca{pn}ify*, which should ensure stress on the vowel preceding the target stop; compare *ámplify* and *clárify*. On the other hand, stimuli in the unstressed condition always ended in *-{CN}VCic*, e.g. *ste{bn}alic*, *ca{pn}adic*, which should ensure stress on the vowel following the nasal; compare *athlétic* and *históric* (Gussenhoven 1994). The same 'stems' were used in both the stressed and unstressed conditions (e.g. *stebn-* and *capn-*).

Items were further split into the underlyingly voiceless condition, where the pre-nasal stop was one of [p], [t], [k], and the underlyingly voiced condition, where the pre-nasal stop was one of [b], [d], [g]. Each sequence appeared in both conditions, meaning that the experiment consisted of voicing minimal pairs, where all strings appeared in each combination of stress and underlying voicing. Thus, the item *capnify* was in the voiceless and stressed conditions, the item *cabnify* in the voiced and stressed conditions, the item *capnadic* was in the voiceless and unstressed conditions, and the item *cabnadic* was in the voiced and unstressed conditions.

There were 24 critical items, 6 in each combination of voicing and stress conditions, 2 for each place of articulation. A summary of the items can be seen in Table 1 in orthographic form, as in the experiment.

Table 1*: critical items*

stressed	*unstressed*
ara(t/d)mify	ara(t/d)manic
ca(p/b)nify	ca(p/b)nadic
mi(c/g)nify	mi(c/g)neric
peri(c/g)nify	peri(c/g)nilic
sa(t/d)mify	sa(t/d)minic
ste(p/b)nify	ste(p/b)nalic

Critical items were generated using a bigram learning model trained on a corpus of English (Baayen et al. 1995) and manually selected for realism by the first author. Additionally, 40 filler items were generated using the same methodology. The filler items did not exhibit stop-nasal sequences, but did occasionally exhibit the same affixes as target items (-*ify*, -*ic*), as well as other affixes, or no affixes at all.

Critical items were split into two blocks, such that no two items of the same string differing only in underlying voicing condition or stress condition appeared in the same block. For example, if block 1 contained the item *capnadic* (stressed, voiceless), block 2 contained *capnadic* (stressed, voiced) and *capnify* (unstressed, voiceless). Each block was repeated twice. Participants were not made aware of block transitions.

2.2. Participants

A total of five participants took part in the study. All participants were in their 20s and graduate students at McGill University at the time of the experiment. Four of the five participants were native speakers of English; the remaining participant was a native speaker of Mandarin. Three of the participants were born in Canada, one was born in the US, and one was born in China.

Participants were not made aware of the purpose of the experiment. Furthermore, in conversation after completion, none were able to guess the purposes of the experiment and seemingly were not aware that the experiment contained voiced-voiceless minimal pairs.

2.3. Analysis

Recordings were manually annotated using *Praat* (Boersma and Weenik 2019). The stop closure, nasal closure, and preceding vowel were all marked separately, as in Figure 1. Aspiration, if present, was also marked separately. Stop bursts were counted as part of the aspiration.

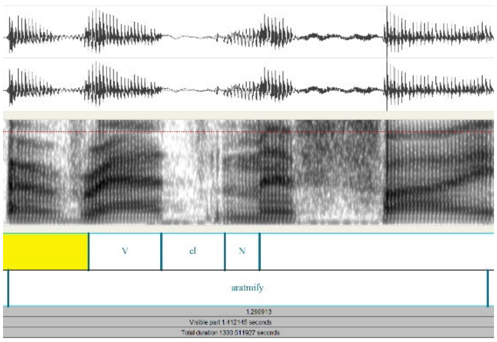

Figure 1. Example annotation of a production of the item *aradmify* in *Praat*. V = preceding vowel; cl = stop closure; N = nasal closure.

Duration for each interval was extracted automatically using a script. As such, preceding vowel duration, closure duration and aspiration duration (i.e. positive VOT) all featured in the analysis. To avoid speakers with slower speech rates disproportionately skewing the results, all duration measures were converted to by-participant z-sores. Additionally, f0 measurements were extracted at 3 evenly spaced points during the stop closure. The f0 measurements themselves did not feature in the analysis. However, the presence of pitch was used as a proxy for pre-voicing (since it is impossible to measure f0 if the vocal folds are not vibrating).

A linear mixed-effects model was run for each of the experimental measures. The model included the acoustic measure as the dependent variable and stress and underlying voicing as independent variables. Speaker and item 'stem', e.g. *ste(p/n)n* or *ca(p/b)n*, were included as random effects.

3. Results

3.1. Duration

3.1.1. Aspiration

Only 50.4% of stop productions exhibited aspiration. As expected, aspiration was more common in the underlyingly voiceless condition (71.7%) than in the underlyingly voiced condition (29.2%). Figure 2 shows the aspiration duration by stress and underlying voicing conditions, with productions not exhibiting aspiration treated as having aspiration with duration of 0.

As seen in Figure 2, voice onset time is greater for the underlyingly voiceless /p,t,k/ as opposed to the underlyingly voiced /b,d,g/, which is not surprising given that the voicing contrast in English is typically described as primarily a difference in VOT. A linear mixed-effects model on the results shows that underlying voicing has a statistically significant effect on aspiration duration ($p < 0.001$), while stress does not ($p = 0.750$). As such, aspiration as a voicing cue can be said to be robust even in the low functional load word-medial and pre-nasal environment.

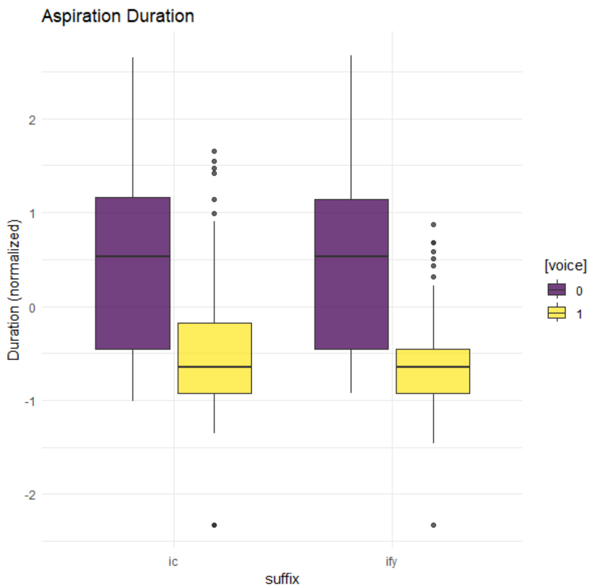

Figure 2. Boxplots of aspiration duration, measured from the stop burst to the onset of voicing in the following nasal, converted to by-speaker z-scores. Underlyingly voiceless condition /p,t,k/ is in purple (dark shade); underlyingly voiced condition /b,d,g/ is in yellow (light shade).

3.1.2. Preceding Vowel

Unlike the case with aspiration, all productions preserved the preceding vowel. Figure 3 shows the vowel duration by stress and underlying voicing conditions.

As seen in Figure 3, vowel duration is generally greater preceding underlyingly voiced stops /b,d,g/ as opposed to the underlyingly voiceless stop /p,t,k/, something that is well reported for English (Steffman 2019). However, the difference in vowel duration is more pronounced in the stressed condition, e.g. *cabnify* [ˈkæːbnɪfaɪ] vs *cabnadic* [kæ·bˈnædɪk]. In fact, the average vowel duration in unstressed vowels before a voiced stop is comparable to the average vowel duration in stressed vowels before a voiceless stop. Nevertheless, a linear mixed-effects model on the results does not reveal a statistically significant effect of underlying voicing on vowel duration ($p = 0.060$) or one of stress ($p = 0.316$). Given that the results lean in the same direction for both stress conditions, it is difficult to say if the lack of significance can perhaps be attributed to a lack of statistical power. In any case, we take the results to mean that vowel duration, unlike aspiration duration, follows the reported trends but exhibits substantial overlap between the voicing conditions, perhaps modulated by stress.

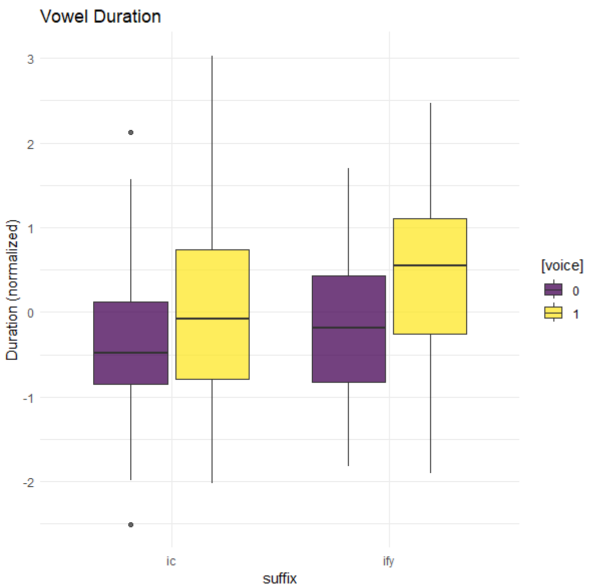

Figure 3. Boxplots of vowel duration, measured up to but not including the stop closure, converted to by-speaker z-scores. Underlyingly voiceless condition /p,t,k/ is in purple (dark shade); underlyingly voiced condition /b,d,g/ is in yellow (light shade).

3.1.3. Stop Closure

All productions exhibited a stop closure. Figure 4 shows the stop closure duration by stress and underlying voicing conditions.

As seen in Figure 4, stop closure duration is slightly greater for the underlyingly voiceless stops /p,t,k/ as opposed to the underlyingly voiced stops /b,d,g/, something that is reported for English (Cole et al., 2007). However, the difference is only seen in the stressed condition. A linear mixed-effects model on the results does not reveal any statistically significant effect of underlying voicing on closure duration in general ($p = 0.483$). It is difficult to say if the lack of significance can be attributed to a lack of statistical power. In any case, it appears that the difference in stop closure between voicing conditions, if reliable, is modulated by stress.

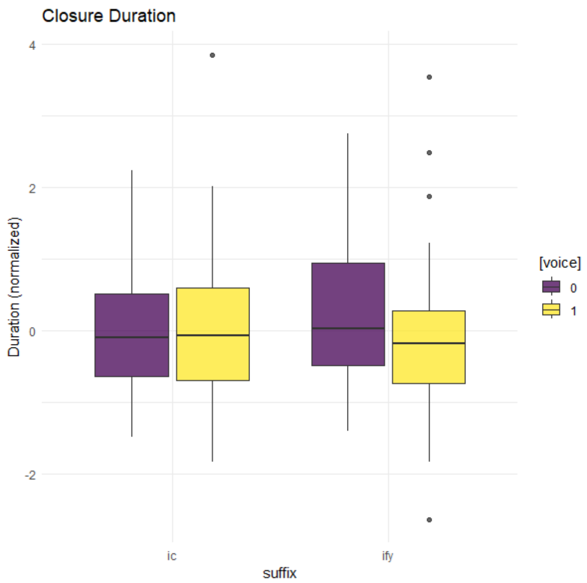

Figure 4. Boxplots of stop closure duration, measured from the preceding vowel offset to the burst, converted to by-speaker z-scores. Underlyingly voiceless condition /p,t,k/ is in purple (dark shade); underlyingly voiced condition /b,d,g/ is in yellow (light shade).

3.2. Voicing

Voicing during closure was approximated using the number of successful f0 measurements. Given that there were three f0 measurements throughout the closure, this resulted in a 4-point scale of voicing: 0/3, 1/3, 2/3, 3/3. To simplify the analysis, the position of voicing during closure (initial, final, medial) was disregarded, although it is known that this, too, is variable (Davidson 2016). Figure 5 shows the voicing proportion by stress and underlying voicing conditions.

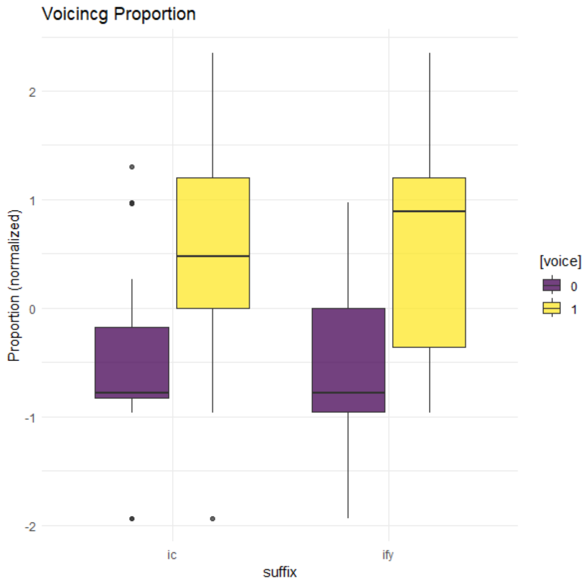

Figure 5. Boxplots of proportion of voicing, measured as the proportion of successful f0 measurements during closure, converted to by-speaker z-scores. Underlyingly voiceless condition /p,t,k/ is in purple (dark shade); underlyingly voiced condition /b,d,g/ is in yellow (light shade).

As seen in Figure 5, voicing proportion is substantially greater for the underlyingly voiced stops /b,d,g/ as opposed to the underlyingly voiced stops /p,t,k/, which is not surprising since voicing during closure is the canonical difference between voiced and voiceless stops cross-linguistically. The difference appears to be equally pronounced in both stress conditions. A linear mixed-effects model on the results reveals a statistically significant effect of underlying voicing on vowel duration ($p < 0.001$) but not stress ($p = 0.930$). As such, voicing proportion patterns with aspiration duration in that it appears to be a robust cue of underlying voicing, even in word-medial and pre-nasal position.

3.3. Principal component analysis

The previous sections explore the realization of the English stop voicing contrast in pre-nasal position for each of the acoustic cues separately. However, the question still remains whether the voicing contrast is maintained in this position along any combination of cues. In order to address the issue of contrast, the data was analyzed using principal component analysis (PCA). PCA is an unsupervised machine learning algorithm that projects the data onto a new set of uncorrelated dimensions. The algorithm begins by finding the dimension (i.e., principal component) of highest variance in the data. Thereafter, PCA iteratively identifies dimensions of highest variance orthogonal to the previous principal components. The dimensions may be 'diagonal' to the original measurements, that is, they may comprise multiple different measurements scaled by corresponding coefficients. PCA is particularly useful for analyzing the structure of multidimensional datasets comprising (potentially) correlated measurements (Greenacre et al. 2022).

For the current project, the PCA analysis was conducted on the three duration measures (aspiration, preceding vowel, stop closure), as well as voicing proportion. Additionally, the duration of the following nasal and vowel peripherality, measured as the Euclidean distance from the center of the vowel space, were also included. All measures were converted to by-participant z-scores. The PCA analysis effectively measures how different the productions are from each other along any combination of these acoustic measurements. The first principal component is the combination of acoustic features that separates productions the most, followed by the second principal component, etc.

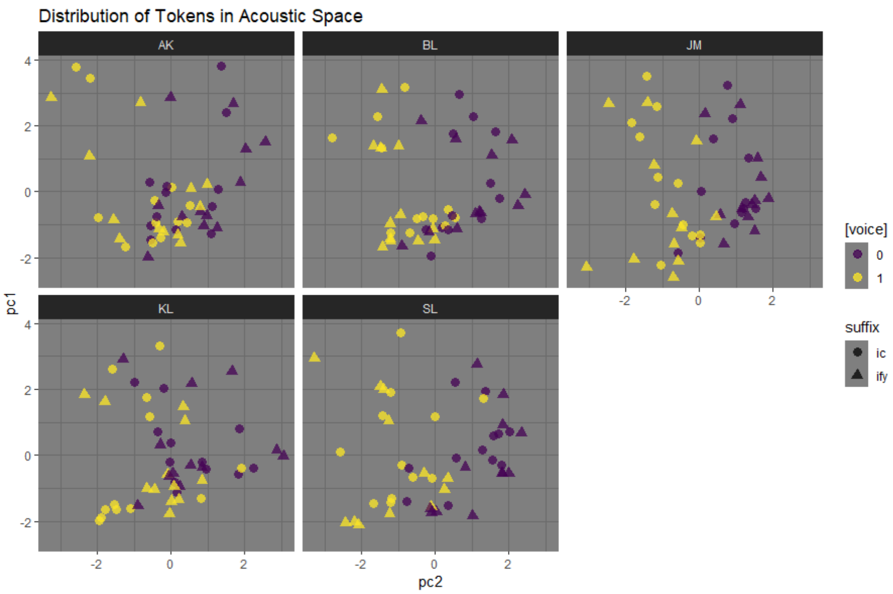

Figure 6. Principal components 1 and 2 for the critical items, shown separately by speaker. Underlyingly voiceless condition /p,t,k/ is in purple (dark shade); underlyingly voiced condition /b,d,g/ is in yellow (light shade).

Figure 6 displays the critical item productions from the experiment on the first two principal components (pc1 and pc2), which account for 49.9% of the variance in the data, separated by speaker. As can be seen in Figure 6, underlyingly voiceless items and underlyingly voiced ones differ primarily in the second principal component. The first principal component (29.5% of the total variance) is orthogonal to underlying voicing and, upon a closer inspection, appears to correspond to the stop place of articulation (negative = more coronal, positive = more dorsal).

Notably, the underlyingly voiceless and underlyingly voiced productions are not completely separable in the pc1 x pc2 space. Participant KL in particular, who is a native speaker of Mandarin, exhibits substantial (but not complete) overlap in the two categories, which is perhaps not surprising, given the differences in realization of voicing between English and Mandarin (Hui and Oh 2015). However, a certain degree of overlap can also be observed for the remaining four participants, all of whom are native speakers of English and raised in English speaking environments.

Impressionistically, the overlap is greater in productions that are low in pc1, those that generally correspond to coronal stops. Thus, when taken together, the acoustic cues separate *aratmanic* and *aradmanic* less than *pericnilic* and *perignilic*. However, it is difficult to draw strong conclusions from this observation, as there was a limited number of experimental items tested. Stress does not seem to play a role in separability of the two voicing categories.

4. Discussion

The phonetic cues traditionally associated with the English voicing contrast exhibit different behaviours in our results. On the one hand, cues like aspiration and the presence of voicing during closure are robust indicators of phonological voicelessness and phonological voicing respectively, even in pre-nasal word-medial position. On the other hand, cues like preceding vowel duration and closure duration are suppressed either in their entirety or are realized only in stressed syllables. As such, our results imply that acoustic properties related to vocal fold activity (aspiration and voicing during closure) are the primary cues for the voicing contrast in English, present even in positions with a low functional load.

The remaining correlates of phonological [voice] appear to be secondary. Although reported in the literature (Cole et al. 2007), our study did not find a strong correlation between underlying voicing and closure duration. Likewise, the effect underlying voicing on preceding vowel length was observed but not significantly so and only in stressed syllables. The secondary cues may serve the role of enhancement features, working to ensure that the contrast between voiceless and voiced stops is clearly heard. As such, in word-medial pre-nasal position, where functional load is low and no minimal pairs are found, the enhancement features serve little purpose and are at least partially omitted, something that is in line with the literature on contrastive hyper-articulation (Schertz 2013; Wedel et al. 2018).

Nevertheless, we interpret the results with caution. Although aspiration and voicing during closure were reliable cues even in the limited dataset presented in this paper, it could be the case that the lack of statistical significance for the secondary cues can also be reasonably attributed to a lack of statistical power. In other words, a dataset with a greater number of participants or productions could very likely reveal a statistically significant effect of both closure duration and vowel duration, possibly modulated by stress. However, we still believe that the lack of some cues in our dataset but the presence of others indicates a separation between cue in terms of relative importance.

Finally, the principal component analysis conducted on the results shows that even when all available cues are taken together, some stop productions are ambiguous in voicing. For all speakers, therefore, the separation between voiceless and voiced stops was not complete. It remains to be seen if the lack of separation can be attributed solely to the low functional load of voicing in pre-nasal position. There exist languages where the voicing contrast is limited to particular phonological positions within the word, such as non-final position in most Slavic languages (Kavitskaya 2017). Although English does distinguish voicing in word-final position, the results imply that it is similar in that it suppressed contrast in phonological positions with a low functional load, such as pre-nasally.

5. Conclusion

This paper discusses the realization of voicing in English stop-nasal sequences, c.f. in *a*[gn]*ostic* vs *a*[kn]*owledge*. The results of an exploratory productions study suggest that aspiration and voicing during closure (i.e. voice onset time) reliably distinguish voiceless and voiced stops, although no voiceless-voiced minimal pairs in this environment exist in the language. Contrariwise, preceding vowel duration and stop closure duration between voiceless and voiced stops in this position are largely overlapping, somewhat more so in unstressed syllables than stressed ones.

When analyzed together using a principal component analysis, we see that all speakers occasionally produce stop-nasal sequences that are ambiguous for voicing.

This study contributes to the fields on English phonetics and phonetics of voicing in general. Moreover, the results contribute to the growing literature on the role of lexical contrast in language. Although further work is needed, we suggest that the low functional load of voicing in stop-nasal sequences is one of the reasons for the suppression of some but not all voicing cues in this position.

References

Baerman, Matthew. 2010. "Defectiveness and homophony avoidance". *Journal of Linguistics* 47: 1-29.

Baayen, Harald, Richard Piepenbrock, and Lennart Gulikers. 1995. *CELEX2 LDC96L14. Web Download*. Philadelphia: Linguistic Data Consortium.

Boersma, Paul and David Weenink. 2019. *Praat: Doing phonetics by computer*. Retrieved from https://www.fon.hum.uva.nl/praat/ (Computer program)

Cole, Jennifer, Heejin Kim, Hansook Choi, and Mark Hasegawa-Johnson. 2007. "Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech". *Journal of Phonetics* 35: 180-209.

Crosswhite, Katherine. 1999. "Intra-paradigmatic homophony avoidance in two dialects of Slavic". *UCLA Working Papers in Linguistics* 1: 48-67.

Davidson, Lisa. 2014. "Variability in the implementation of voicing in American English obstruents". *Journal of Phonetics* 54: 35-50.

Flemming, Edward. 2017. "Dispersion theory and phonology". In *Oxford Research Encyclopedia of Linguistics*, ed. by Mark Aranoff. Oxford: Oxford University Press.

Gessner, Suzanne and Gunnar Ólafur Hansson. 2004. "Anti-homophony effects in Dakelh (Carrier) valence morphology". Proceedings from the *Berkeley Linguistic Society of America* 30: 91-103.

Greenacre, Michael, Patrick Groenen, Trevor Hastie, Alfonso Iodice D'enza, Angelos Markos, and Elena Tuzhilina. 2022. "Principal component analysis". *Nature Review Methods Primers* 2.

Gussenhoven, Carlos. 1994. "English stress in lexical phonology". In *Phonologica 1992: Proceedings of the 7th International Phonology Meeting*, ed. by Wolfgang U. Dressler, Martin Prinzhorn and John J. Rennison, 87-96. Turin: Rosenberg & Sellier.

Hauser, Ivy. 2022. “Speech sounds in larger inventories are not (necessarily) less variable”. *The Journal of the Acoustical Society of America* 252: 2664.

Hui, Yang and Mira Oh. 2015. “Adaptation of English stops into Mandarin Chinese”. *Linguistic Research* 32: 403-417.

Kavitskaya, Darya. 2017. “Some recent developments in Slavic phonology”. *Journal of Slavic Linguistics* 25: 387-413.

King, Robert. 1967. “A measure for functional load”. *Studia Linguistica* 21: 1-14.

Lisker, Leigh and Arthur Abramson. 1967. “Some effects of context on voice onset time in English stops”. *Language and Speech* 10.

Martinet, André. 1955. *Économie des changements phonétiques: traité de phonologie diachronique*. Berne: Francke.

Munteanu, Andrei. 2021. “Homophony avoidance in the grammar: Russian nominal allomorphy”. *Phonology* 38: 401-35.

Ogura, Mieko and William Wang. 2018. “Evolution of homophones and syntactic categories noun and verb”. In *Proceedings of the 12th International Conference on the Evolution of Language (Evolang12)*, ed. by Christine Cluskey, Molly Flaherty, Hannah Little, Luke McCrohon, Andrea Ravignani and Tessa Verhoed, 401-435.

Peirce, Jonothan, Jeremy Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Lindeløv. 2019. “PsychoPy2: experiments in behavior made easy”. *Behavior Research Methods* 51: 195-203.

Round, Erich, Rikker Dockum, and Robin Ryder. 2022. “Evolution and trade-off dynamics of functional load”. *Entropy* 24: 507.

Schertz, Jessamyn. 2013. “Exaggeration of featural contrasts in clarifications of misheard speech in English”. *Journal of Phonetics* 41: 249-263.

Schertz, Jessamyn and Yoonjung Kang. 2022. “Phonetic cue competition within multiple phonological contrasts: Perception of Seoul Korean sibilants”. *Korean Linguistics* 18: 1-17.

Silverman, Daniel. 2009. “Neutralization and anti-homophony in Korean”. *Journal of Linguistics* 46: 453-482.

Steffman, Jeremy. 2019. “Phrase-final lengthening modulates listeners’ perception of vowel duration as a cue to coda stop voicing”. *Journal of the Acoustical Society of America* 145: EL560-EL568.

Wedel, Andrew, Abby Kaplan, and Scott Jackson. 2013. “High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128: 179-186.

Winter, Bodo and Andrew Wedel. 2016. “The co-evolution of speech and the lexicon: the interaction of functional pressures, redundancy, and category variation”. *Topics in Cognitive Science* 8: 503-513.

***. 2019. “R: A language and environment for statistical computing.” Details available at https://www.r-project.org/