GARAI LUCA

ELTE BTK Nyelvtudományi Doktori Iskola
garailuc@gmail.com
https://orcid.org/0000-0002-9371-260X

Garai Luca: Acoustic characteristics of nasal and oral flaps
Alkalmazott Nyelvtudomány, Különszám, 2023/2. szám, 3–20.
doi:http://dx.doi.org/10.18460/ANY.K.2023.2.001

Acoustic characteristics of nasal and oral flaps

The study investigates the flapping of intervocalic /t/, /n/, and /nt/ in nonsense words uttered by six North American speakers. The main focus of the experiment is to explore the acoustic properties which distinguish oral [ɾ] and nasal [ɾ̃] flap sounds, and to investigate the effect flaps and preceding vowels have on each other in terms of duration and nasality. The results show that the flapping of /nt/ in nonsense words occurs infrequently, while /t/ is consistently flapped. In terms of consonant duration, oral flaps were found to be shorter than nasal flaps, whereas there was no significant difference in flap nasality based on H1*−H2* values. Preceding vowels were nasalized before nasal flaps but not before oral flaps. The results broaden our understanding of the acoustics of alveolar flap sounds, as well as the appearance of oral and nasal flaps in nonsense words.

Keywords: lenition, alveolar flapping, nasalization, spectral tilt, American English

1. Introduction

1.1. Alveolar flapping

Flapping is a lenition process which is present in various languages and results in the shortening of consonants into a brief, sonorous segment, a flap [ɾ] (Kenyon, 1994). In American English, flapping primarily affects the alveolar plosives /t d/ when they appear between sonorants (Kahn, 1976). The most frequent flapping environment is between a stressed and an unstressed vowel, possibly preceded by a rhotic liquid (V́(r)_V) but flapping may occur between two unstressed vowels and across word boundaries as well (Kahn, 1976; Patterson & Connine, 2001). In general terms, when an alveolar plosive is flapped, instead of a longer hold+release sequence, only a short, voiced closure (similar to that of liquids /r l/) is uttered. As a result, flapping neutralizes the difference in voicing between /t/ and /d/, and word pairs such as *matter–madder* are both realized as [mæɾɚ].

The prominent differentiating feature between flapped and unflapped realizations is segment duration, which is “extra short” in the case of flaps (Steriade, 2000: 323), usually between 10 and 40 ms, whereas the duration of unflapped alveolar plosives reportedly falls in the 100 − 140 ms range (Zue & Laferrière, 1979). Apart from duration, the only acoustic feature which is regularly mentioned in connection with flaps is a higher degree of voicing in flaps compared to unflapped realizations, but previous studies did not investigate the exact degree of this difference between realizations. The relationship between alveolar flap sounds and the sonorants preceding them has been examined from various aspects. Zue and Laferrière (1979) compared the realizations of /t/ and /d/ in utterances of various English words with different phonetic environments, including the aforementioned (V́(r)_V) environment, where they found the majority of /t d/ segment durations to fall with in the 10 to 40 ms range. However, they observed that "[f]lap duration after the vowels [i, e, ʊ, ɑy, oy, yu] was greater, for both /t/ and /d/ flaps, than duration after all other vowels [ɪ, ɛ, u, ɜ, ʌ, o, aw, ɑ, æ, ɔ]" (Zue & Laferrière, 1979: 1045). This means that, according to the findings of Zue & Laferrière (1979), speakers produce longer flaps after high palatal vowels compared to nonhigh and/or nonpalatal ones.

Another notable vowel-related phenomenon in connection with flapping is pre-fortis clipping. In English, vowels before fortis obstruents are articulated with a shorter duration than before lenis obstruents and sonorants (Scheer, 2017). As Braver (2014) notes, pre-fortis clipping occurs even if the obstruent in question is flapped, which means that even though the acoustic features of /t/ and /d/ are merged via flapping, the neutralization remains incomplete, and the stressed vowel in *betting* will be shorter than the one in *bedding*, for instance. However, despite the measurable difference in vowel duration, speakers cannot distinguish flapped utterances of such word pairs based on this difference alone, thus the neutralization is perceived as complete, and the difference is primarily detectable via acoustic analyses (Braver, 2014).

Several other linguistic and nonlinguistic factors have been considered as possible determiners of the likelihood of flapping. The gender of the speaker is claimed to influence flap frequency, as women are less likely to flap and their flaps tend to be longer in duration than those of male speakers (Garai, 2021; Zue & Laferrière, 1979). The lexical frequency of a word has also been cited as a determining factor: words which occur more frequently are more likely to be uttered with flaps (Garai, 2021; Patterson & Connine, 2001). In connection with this, lexical bias is considered heavily influential, since speakers store frequently heard forms, including flapped realizations, and associate them with the specific lexical item, making it more likely for some words to occur with flapping than others, regardless of the phonetic environment of the consonant (Connine, 2004). The flapping of /t/ and /d/ in American English is a well-established process due to its ubiquity, therefore it is expected to occur automatically in the appropriate environments for most speakers, even in previously unknown or nonsense words.

1.2. Acoustic properties of nasals

The current study is concerned with the flapped realization of alveolar nasals and nasal+plosive clusters in relation to /t/ flaps. Therefore, it is crucial to mention some key articulatory and acoustic features associated with nasal segments. During the articulation of nasals, the velum is lowered and air flows through the nasal tract, which is thus involved in the articulatory process, along with the pharyngeal and oral cavities (Fujimura, 1962). This phenomenon is called velopharyngeal coupling and it influences the acoustic properties of the sounds produced, since the different vibrations in the nasal and oral tracts interfere with each other, enhancing the intensity of the acoustic signal at some frequencies and reducing it at others. Fujimura (1962) observed various nasal formants and antiformants resulting from the process, the first one (P0) being the most prominent at 250−300 Hz and subsequent nasal formants (P1−P4) occurring based on the place of articulation of the nasal consonant in the oral cavity. Spectral tilt, which is measured by subtracting the second harmonic (H2*) from the first harmonic (H1*), is considered to be steeper in nasals until ∼ 2000 Hz compared to oral consonants (Fujimura, 1962). Spectral tilt is reported to be influenced by fundamental frequency (f0), with higher values of f0 being associated with steeper spectral tilt, or higher values of H1*−H2* (Garellek et al., 2016). It has also been observed that the bandwidth of nasal formants is wider than that of oral formants (Bunton et al., 2011).

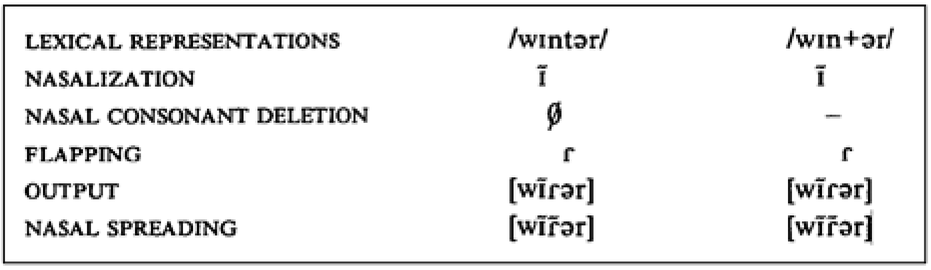
Nasality is a feature which can spread to neighboring segments, primarily vowels (Fujimura & Lindqvist, 1971). The reason for the spread is that the lowering and raising of the velum does not happen simultaneously with the oral closure, thus velopharyngeal coupling may overlap with some portion of the preceding and/or following vowel(s), changing its formant structure along the way. Vowel nasality can be a distinctive feature in languages where nasal deletion occurs, but in most cases, it occurs automatically around nasals. The nasality of vowels can be measured by A1*−P0 which is lower in the case of nasalized vowels compared to fully oral vowels (Styler, 2017). In nonhigh vowels, P0 corresponds to either H1* or H2* (whichever is higher in intensity); high vowels show formant interference in this frequency range and their nasality cannot be accurately measured using A1*−P0 (Styler, 2017).

1.3. Nasal flapping

The flapping of alveolar nasals and /nt/ clusters has been investigated primarily from a phonological point of view in the previous decades. Kahn (1976) observes that alveolar plosives can be flapped in a postnasal environment, but only if the nasal consonant is deleted after nasalizing the preceding vowel, thus the pronunciation of *winter* in casual speech can be realized as [wɪ̃ɾɚ]. The subsequent analysis of Jensen (1993) includes the nasalization of the flap sound itself, although the nasalization does not come directly from the nasal consonant. He claims that the preceding vowel is nasalized first, followed by the deletion of the nasal, then flapping occurs, and it is, in fact, the nasality of the vowel which finally spreads to the flap (see Figure 1). Contrary to the nasal deletion approach, Picard (1997) proposes a more straightforward explanation to the homophonous realization of word pairs such as *winter* and *winner*, [wɪɾ̃ɚ]. In his derivation, the plosive is deleted from the /nt/ cluster and flapping occurs directly on the nasal, resulting in a distinct nasal flap [ɾ̃] (see Figure 2).

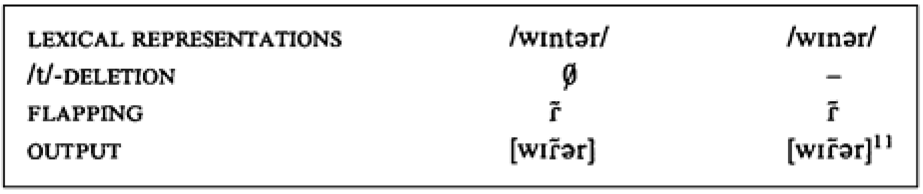
According to Picard (1997), flapping affects alveolar nasals in the same environments as /t/ and /d/. This claim is supported by Vaux (2000), who asserts that intervocalic /n/ after stressed vowels can be considered a flap sound based on its phonetic characteristics, and that this flap sound differs from flapped plosives. These distinctive features are not specified, however; and, as mentioned above, previous acoustic analyses of flapping were concerned with segment duration only. The findings of Garai (2019) show that the duration of /n/ and /nt/ in a V́_V environment was significantly shorter than in a V_V́ environment (where the flapping of alveolars is not expected), which suggests that /n/ and /nt/ can be realized as flap sounds; the duration of these nasal flaps was found to be longer than that of /t/ flaps.

Figure 1. Jensen’s analysis of flapping in winter and winner (from Picard, 1997: 291)



LEXICAL REPRESENTATIONS	/wɪntər/	/wɪn+ər/
NASALIZATION	ɪ̃	ɪ̃
NASAL CONSONANT DELETION	0	–
FLAPPING	ɾ	ɾ
OUTPUT	[wɪ̃ɾər]	[wɪ̃ɾər]
NASAL SPREADING	[wɪ̃ɾ̃ər]	[wɪ̃ɾ̃ər]

Figure 2. Picard’s analysis of flapping in *winter* and *winner* (from Picard, 1997: 291)



LEXICAL REPRESENTATIONS	/wɪntər/	/wɪnər/
/t/-DELETION	0	–
FLAPPING	ɾ̃	ɾ̃
OUTPUT	[wɪɾ̃ər]	[wɪɾ̃ər][11]

Spectral differences between unflapped and flapped surface realizations of /t/ and /nt/ can be seen in Figure 3 below. Before flapping, the nasal and the plosive in *mounting* are clearly distinguishable, with a silent closure following the formants of the nasal; whereas when flapping does occur, the two segments coalesce into a brief, sonorous consonant between the two vowels. A zoomed in image of flapped /t/ and /nt/ are compared in Figure 4. There are also notable differences in the spectral structure of oral [ɾ] and nasal [ɾ̃]: while the former only appears to be voiced without a definite formant structure, the latter exhibits visible formants (and it remains distinct from its neighboring vowels due to the lack of intensity in certain frequency ranges, that is, antiformants).

The primary aim of the current study is to investigate the measurable acoustic characteristics of oral and nasal flap sounds in English nonsense words uttered by North American speakers, as well as the interaction between flaps and preceding vowels regarding duration and nasality features. The purposes of the experiment described below can be summed up with the following two research questions:

(1) Do nasal flaps [ɾ̃] produced from underlying /n/ and /nt/ differ from oral flaps [ɾ] produced from underlying /t/ in terms of segment duration and spectral tilt, the latter signaling consonant nasality?
(2) How do flapped segments and the vowels preceding them affect each other in terms of duration and nasality?

Figure 3. Spectrograms of the words *shouting* (left) and *mounting* (right) produced by the same speaker, with unflapped [t] and [nt] (top) and flapped [ɾ] and [ɾ̃] (bottom); arrows indicate the boundaries of the target consonant(s) of flapping

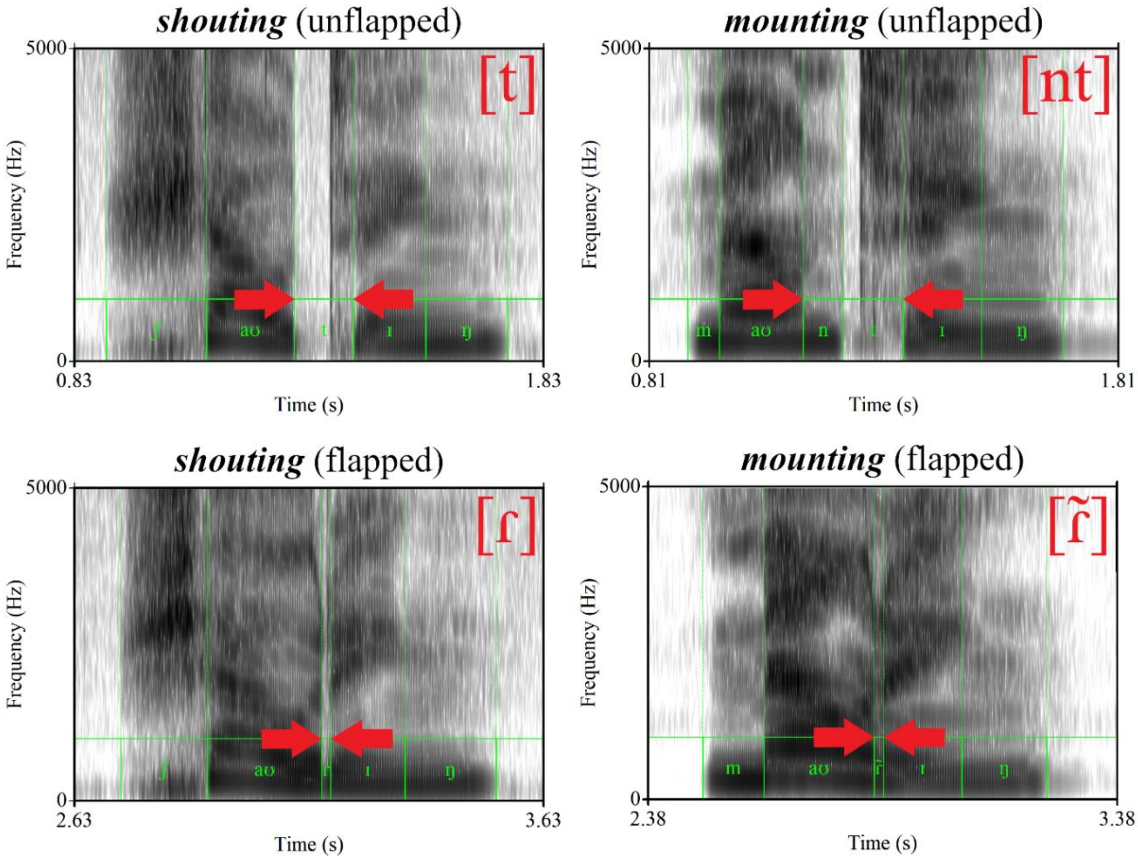
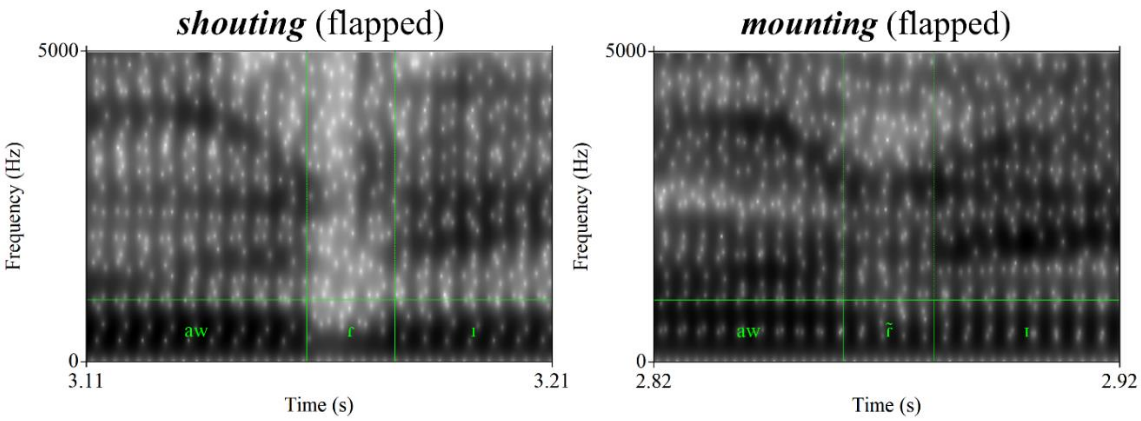


Figure 4. Spectrograms of flapped /t/ (left) and flapped /nt/ in the words shouting and mounting, respectively. The images are zoomed in versions of the flapped utterances in Figure 3.



2. Methods

2.1. Stimuli

A set of 24 monosyllabic nonsense words was created as the stimuli for the production experiment. Each stimulus word contained one of the eight stressed vowels [ɪ, ɛ, æ, ʌ, ɑ, ej, oj, aw] and one of the word-final consonants /t n nt/, resulting in eight sets of minimal pairs (see Table 1) where the final consonants, if followed by either of the added suffixes '-ing' or '-er', are in an ideal environment for flapping to occur (i.e. between two vowels, the second of which is unstressed). The eight sets of stimuli all began with a plosive+liquid cluster (one of /dr kl pr bl/), and the initial cluster for each minimal pair group was chosen to ensure a lack of orthographic or phonetic resemblance to existing lexical items in all cases (e.g. if all stimulus words had began with /dr/, one of the words would have been *DRAIN*, which is an existing word). The stimulus words were written in capital letters throughout the task. A set of 48 monosyllabic filler items was also created, with the same initial consonant clusters and vowels as the target stimuli. The only difference was the word-final consonant(cluster), which, in each case, was one of /p b m mp k ng/ (e.g., *drup, drib, proim,* etc.). The filler stimuli were also written in capital letters, in the same fashion as the target items.

Table 1. Stimulus words used in the experiment

V́_	**/t/**	**/n/**	**/nt/**
ɪ	drit	drin	drint
ɛ	dret	dren	drent
æ	drat	dran	drant
ʌ	drut	drun	drunt
ɑ	drot	dron	dront
ej	clait	clain	claint
oj	proit	proin	proint
aw	blout	bloun	blount

2.2. Participants

Six North American native speakers of English participated in the experiment, three females and three males, between the ages of 19 and 28 (with a mean age of 24.8 years). One of the female subjects grew up in Ontario, Canada, while the other five subjects were raised in various areas of the United States: one in California, two on the East Coast and two in Texas. All six subjects spent most or all of their lives in a North American, primarily English-speaking environment; two of them were monolingual English speakers, while the other four subjects listed English as their only native language, and each of them spoke at least one other language between intermediate and advanced levels. None of the subjects reported any illnesses or disabilities affecting their vocal tract, speech, or hearing, and none of them were habitual smokers.

2.3. Procedure

Due to the pandemic, the recording of the samples could not take place in a traditional soundproof studio environment. Instead, subjects were sent detailed instructions regarding the experiment and had to record their own samples in a quiet room, using headset microphones with recording capabilities in the 20 Hz to 20 kHz frequency range). They were asked to use *Audacity* (Audacity Team, 2021), and to save the recording as a WAV file. Subjects were also instructed to fill out an online form containing the terms and conditions of the experiment, as well as a questionnaire asking for relevant demographic data (sex, age, birthplace, current primary residence, places lived, languages spoken, education, occupation, prior linguistic knowledge, smoking habits, and any medical conditions affecting the vocal tract).

The production experiment consisted of a wug test (Berko, 1958) where speakers were presented with a sentence containing the stimulus word and had to formulate the '-ing' and '-er' forms of the word, which they had to fill into the blank spaces of two subsequent sentences (see Figure 5). Speakers were asked to read the three sentences in order. The fill-in task was chosen in order to avoid orthographic influence on articulation (by leaving blank spaces instead of spelling out the expected word form) and to minimize the speakers' focus on pronunciation (Braver, 2014). The order of stimuli was randomized, and each stimulus item was presented once to each speaker.

Figure 5. Example of a stimulus item, as presented to speakers during the task

Katie learned how to **DRIT** this week.
She was ______ this whole week.
She is now the best _____ in the world.

After an initial set of examples provided by the author, speakers reliably produced the appropriate forms of the stimuli so that the two blanks in the example above, for instance, would be filled with *'DRIT-ing'* and *'DRIT-er,'* respectively. There was a small number of slip-ups which were not corrected by speakers (e.g. mixing up the two word forms or producing an unexpected preceding vowel), which were carefully marked during the analysis, and were put in the appropriate categories where relevant (e.g. if the speaker articulated the preceding vowel in *'DRANT-ing'* as /ej/ instead of /æ/, it was put into the "high palatal" group). None of the items were discarded from the analysis. Each speaker produced 48 target tokens (two realizations of each of the 24 stimuli, suffixed with '-ing' and '-er'), resulting in 288 sample items overall.

2.4. Acoustic analysis

The samples were manually segmented in *Praat* (Boersma & Weenink, 2022). Subsequently, the target segments (the target consonants of flapping, as well as the vowels immediately preceding them) were analyzed in *VoiceSauce* (Shue et al., 2011) to extract segment duration, as well as the following measurements, with the default 25 ms window size and 1 ms frame shift, and the measured values later averaged over the entire segment (harmonic locations were estimated using the Straight pitch track and 3 as the default number of periods for harmonic estimation).

H1*−H2* was extracted in the case of flapped consonants because it is considered to indicate the spectral tilt of sonorant consonants, which, in turn, signals consonant nasality by appearing steeper (higher value of H1*−H2*) in nasals compared to orals (Fujimura, 1962, Garellek et al., 2016). Figure 6 shows spectral slices of oral [ɾ] and nasal [ɾ̃] measured in 40 ms windows of tokens from speaker S06 of the current experiment. As it can be inferred from the differences between the first and second harmonic of each token, the amplitude subtraction H1*−H2* yields a higher value in terms of the nasal consonant.

A1*−P0 was measured in the case of preceding vowels since it is the value associated with the nasality of vowels: if we subtract the amplitude of the first nasal pole (P0, corresponding to the higher one of H1* and H2*) from that of the first formant (A1*), a lower value signifies a higher degree of nasalization from the adjacent consonant (Styler, 2017). Figure 7 shows spectral slices measured in 50 ms windows of oral and nasal /ʌ/, extracted from tokens produced by speaker S06 of the current experiment. As shown by the indicated values, the amplitude of P0 does not differ considerably between the two tokens; however, A1* is lower in the case of the nasalized vowel. Therefore, A1*−P0 is lower for the nasalized variant of /ʌ/ compared to the oral realization.

Since there is no conclusive acoustic measurement which could serve as the differentiating factor between flapped and unflapped realizations (except, perhaps, for segment duration, which was used as a dependent variable in the analysis and thus would have been a circular, and thus unfavorable value for the purpose of differentiating flapped and unflapped tokens), two phonologists were asked to provide their own judgments regarding flap production in all sample items based on their perception. They had to grade each item on a scale from 1 (“definitely a flap”) to 5 (“definitely not a flap”) and record the perceived underlying consonant in the given utterance. They were asked to give additional feedback on the items they were “unsure” about (a score of 3) and specify the exact reason behind their choice. The tokens were grouped into “flapped” and “unflapped” utterances based on the summarized perception of the phonologists (in cases where their individual scores did not differ by more than 2 points and gave a mean score other than 3, items with a mean score under 3 were classified as flapped, while those with a mean score over 3 were classified as unflapped; in other cases, the two judgments were deemed inconclusive and the author’s own perception served as the deciding vote).

Figure 6. Spectral slices of oral [ɾ] and nasal [ɾ̃] flaps with the amplitudes of the first harmonic (H1) and the second harmonic (H2) indicated

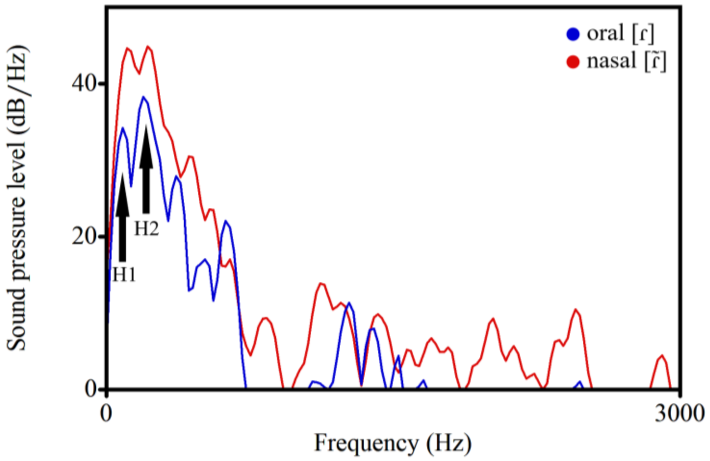
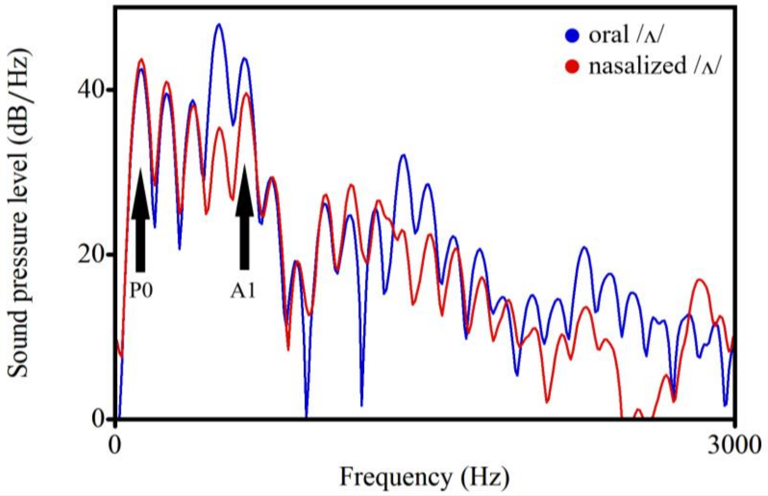


Figure 7. Spectral slices of oral and nasalized /ʌ/ with the amplitudes of the first nasal pole (P0) and the first oral formant (A1) indicated



2.5. Statistical analysis

The analysis of the data was done using *Python* (Python Software Foundation, 2022), with the help of *pandas* (The pandas development team, 2022) and *statsmodels* (Seabold & Perktold, 2010). Several χ^2 tests were conducted to determine the relationship between flap occurrence and potential influencing factors: underlying consonant type, preceding vowel quality, and suffix type. Linear mixed models were used to evaluate the effect of the various factors. Five models were fitted overall. The dependent variables in each of these models were the duration of underlying /n/, the duration of all flapped consonants, mean H1*−H2* of flapped consonants, mean A1*−P0 of nonhigh vowels before flaps, and vowel duration before flaps.

For the /n/ duration model, the fixed effect was flapping status (yes/no). For the other four models concerning flapped tokens only, the type of flap sound (nasal [ɾ̃]/oral [ɾ]) was selected as a fixed effect. Additionally, the quality group of the preceding vowel (high palatal/other) was a fixed effect for the model with flap duration as the dependent variable.

A random intercept for the speaker was added in every model. In the two models concerning consonant duration (both the underlying /n/ durations model and the flapped consonant durations model), the type of the following suffix *(-ing/-er)* was also added as a random intercept in order to account for vowel quality interferences and prosodic differences between the carrier sentences. Finally, for the pre-flap vowel duration and vowel nasality models, vowel quality (ɪ/ɛ/æ/ʌ/ɑ/ej/oj/aw) was added as a random intercept, since different types of vowels are expected to vary in terms of duration, and A1*−P0 differences may vary by vowel quality (Jacewicz et al., 2007; Styler, 2017). Post hoc pairwise comparisons (Tukey's tests) were conducted when the fixed factors had a significant effect on the dependent variable.

3. Results

Various results of the production experiment are detailed in this section. The data and comparisons are presented in an order which roughly follows the structure of the research questions posited at the end of the Introduction. More descriptive results on the likelihood of flapping are introduced first, followed by the acoustics of flap sounds. Then, the various types of effects between flapped consonants and preceding vowels are examined one by one.

3.1. Occurrence of flapping

One of the male subjects, S05 from Texas, had a tendency to hyperarticulate tokens with underlying /t/s, resulting in an almost geminate-like closure duration and a prominent release; his unflapped realizations of /t/ ranged from 108 ms to 268 ms in segment duration. The five other subjects exhibited flapping in tokens with underlying /t/ consistently, except for one unflapped /t/ realization by S03, a female. Of the six subjects, five did not flap /nt/ clusters in the target nonsense words at all. One female subject from Texas, S04, had six tokens (out of sixteen) in which the realization of /nt/ clusters could be classified as flaps. Overall, as Figure 8 shows, it can be said that underlying /t/ was overwhelmingly produced flapped and underlying /nt/ was overwhelmingly produced unflapped by the speakers.

When it came to underlying /n/, however, the likelihood of flapping was not as definitive. Overall, around a third of all /n/ utterances were flapped, as it can be seen in Figure 8. All speakers flapped /n/ in a moderate number of tokens. A χ^2 test was conducted with a significance level of $\alpha = 0.05$, in order to determine whether there was a significant difference between the rate of occurrence of flapping produced by speakers in various underlying consonant groups. The results of the test showed that the likelihood of flapping was significantly different based on the underlying phoneme ($p < 0.01$).

Another χ^2 test was conducted to investigate the relationship between the likelihood of flapping and preceding vowel quality. Tokens were grouped by preceding vowel quality (ɪ/ɛ/æ/ʌ/ɑ/ej/oj/aw) and flapping status (yes/no). The data met the necessary criteria for the test but there was no significant relationship shown between the quality of the preceding vowel and whether or not the consonant was flapped. A final χ^2 test was conducted to see if there was a relationship to consider between the occurrence of flapping and the type of suffix *(-ing/-er)* at the end of each item. The results showed that there was no significant relationship between the two variables. This means that in the current sample, the likelihood of flapping the medial consonant in a token did not depend on either the preceding vowel or the following suffix.

Figure 8. Ratio of flapped and unflapped realizations of each underlying consonant across all speakers.

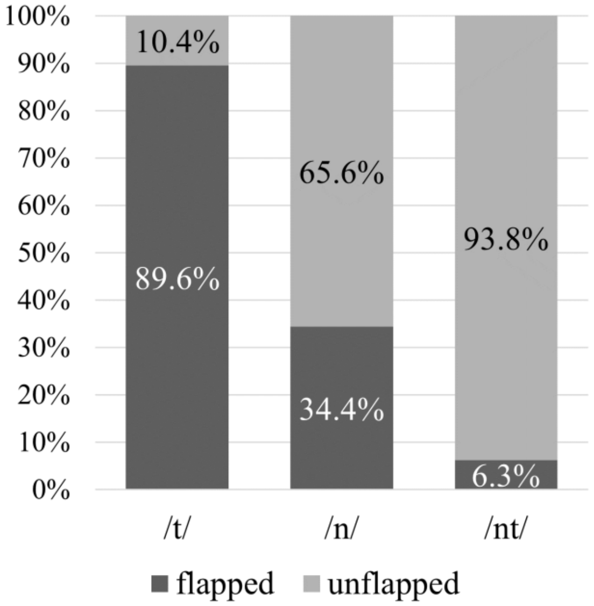
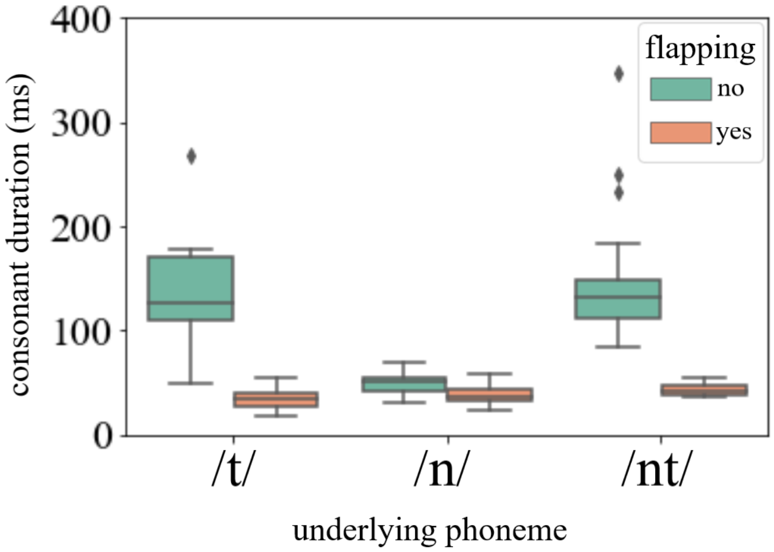


Figure 9. Duration of flapped and unflapped realizations of tokens with each underlying consonant (cluster)

3.2. Consonant duration

As it can be inferred from Figure 9, flapped realizations of each consonant type were considerably shorter than unflapped ones. The small number of unflapped tokens for underlying /t/ and flapped tokens for underlying /nt/ make it impossible to form statistical inferences regarding the flapped and unflapped durations of all 288 realizations; therefore, a linear mixed effects model was used to observe the effect of flapping status on consonant duration only in surface realizations of /n/. The results of the test showed that the occurrence of flapping (yes/no) had a significant effect on the duration of /n/ realizations ($\beta = -10.223 \pm 1.849$, $p < 0.001$), with the post hoc test showing that flapped instances were shorter in duration (38.36 ± 7.87 ms) than unflapped ones (49.49 ± 8.57 ms).



Overall, it can be said that /t/ was overwhelmingly flapped by all speakers, /n/ was flapped in a third of the items on average, and /nt/ was not flapped by most speakers. The only outlier is S04, who flapped six out of sixteen /nt/ utterances and only flapped /n/ once. Because of this distinction, the tokens uttered by S04 were omitted from the statistical analyses. Therefore, the results detailed below concern the flapped utterances of /t/ and /n/ produced by five speakers. Flap sounds produced from underlying /t/ are referred to as oral [ɾ], while those derived from underlying /n/ are referred to as nasal [ɾ̃]. All further comparisons in the current study will rely on the opposition between oral [ɾ] and nasal [ɾ̃] flaps.

3.3. Duration of flap sounds and the effect of vowel quality

The duration of oral and nasal flaps was compared in the samples of all speakers, since it has been observed previously that flaps produced from underlying oral /t/ are significantly shorter than those produced from underlying nasal /n/ and /nt/ (Garai, 2019). Based on the mixed effects model, flap type (oral/nasal) had no significant effect on the duration of the flapped segment, and the duration measurements in oral and nasal flaps don’t differ considerably, as seen in Figure 10. Therefore, it can be inferred that the current sample does not follow the previously observed trend of longer nasal [ɾ̃] and shorter oral [ɾ].

The duration of oral and nasal flaps were analyzed based on the quality of the preceding vowel because a previous study by Zue and Laferrière (1979) found that flap sounds after high palatal vowels are longer in duration than those after other types of vowels. The results of the mixed effects model showed that there was no significant effect of preceding vowel quality on consonant duration, therefore the subjects of the current experiment do not seem to exhibit this phenomenon. The comparison of mean durations of oral [ɾ] and nasal [ɾ̃] flaps in relation to vowel quality is shown in Figure 11.

Figure 10. Duration of flapped consonants, grouped by the type of flap (oral/nasal)

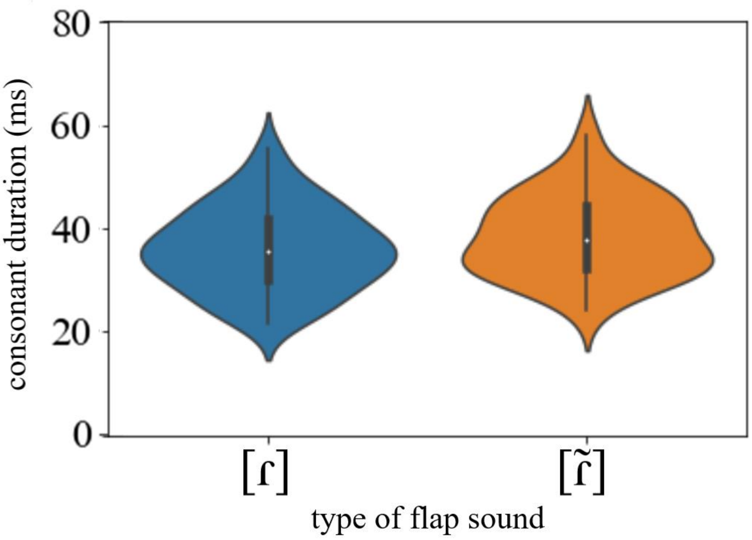
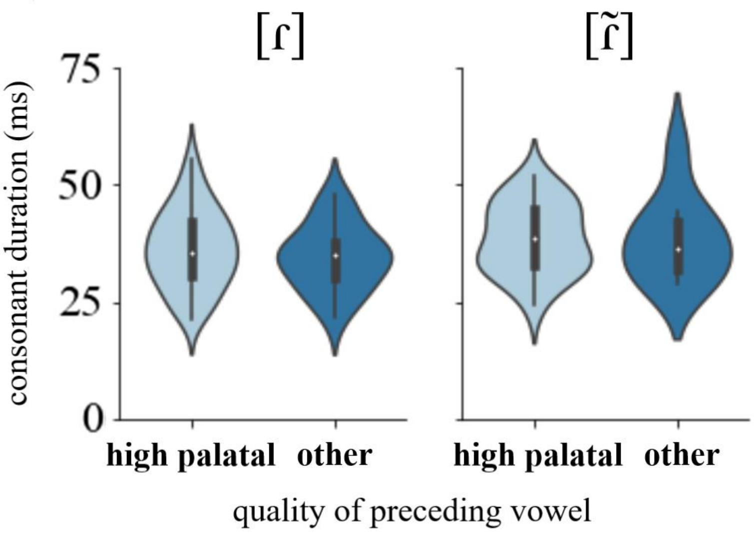


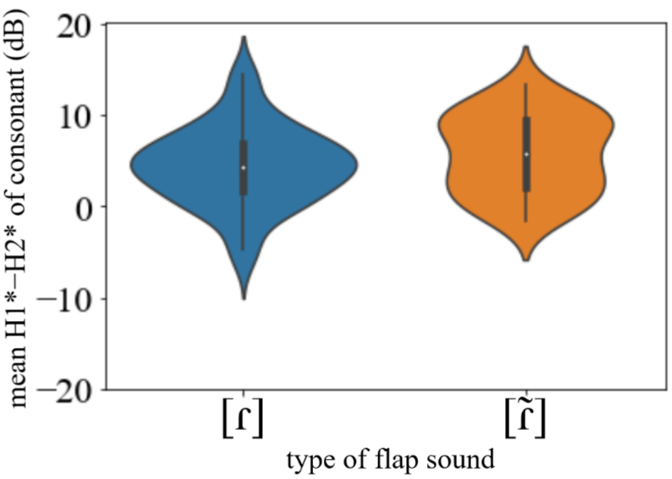
Figure 11. Duration of oral [ɾ] and nasal [ɾ̃] flaps across all speakers, grouped by preceding vowel quality.



3.4. Spectral tilt (nasality) of flap sounds

The relationship between flap type (oral/nasal) and spectral tilt (mean H1*−H2*) was inspected in flapped samples to determine if nasal [ɾ̃] segments retained their nasality after flapping and if oral [ɾ] flaps showed the same level of nasality. The mixed effects model showed that the type of the flap sound (oral/nasal) did not have a significant effect on spectral tilt when only flapped samples were inspected, which indicates that the difference between the mean H1*−H2* values (signaling consonant nasality) of oral [ɾ] versus nasal [ɾ̃] is not present after flapping. The comparison of the data is visualized in Figure 12.

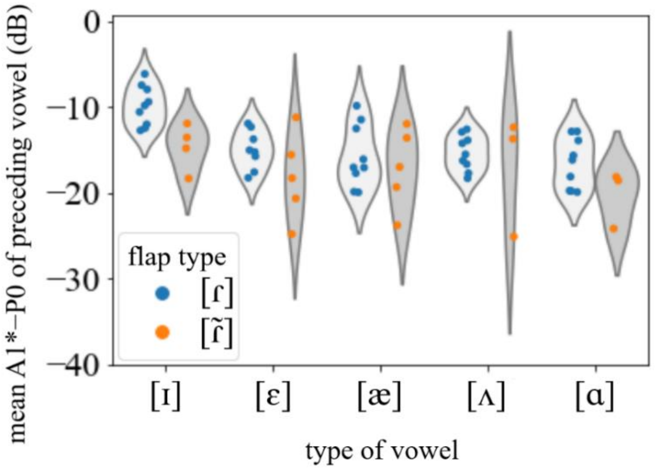
Figure 12. Mean H1*−H2* measurements of flap sounds, grouped by the type of flap (oral/nasal)



3.5. Nasality of preceding vowels

The influence of the flap type on the preceding vowel’s nasality (mean A1*−P0) was examined in the case of nonhigh vowels to show if vowels were nasalized before either oral [ɾ] or nasal [ɾ̃] flaps. The mixed effects model’s results suggest a significant effect of flap type on vowel nasality ($\beta = -3.054 \pm 0.764$, $p < 0.001$; $R^2_m = 0.105$, $R^2_c = 0.615$), with lower values of mean A1*−P0 (suggesting more nasalized vowels) preceding nasal [ɾ̃] flaps (−18.54 ± 5.46 dB) compared to those before oral [ɾ] flaps (−14.19 ± 3.87 dB); this means that vowels before nasal flaps [ɾ̃] were more nasalized than those before oral flaps [ɾ]. The difference in preceding vowel nasality is visualized in Figure 13.

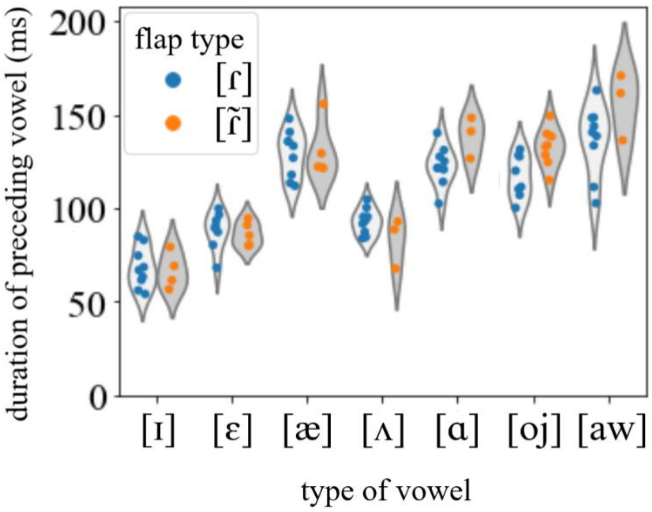
Figure 13. Mean A1*−P0 measurements of preceding vowels, grouped by vowel quality and the type of flap (oral/nasal)



3.6. Duration of preceding vowels

The influence of the type of flap sound on the duration of the preceding vowel was also examined, due to the frequently cited persistent effect of pre-fortis clipping after flapping (Braver, 2014; Scheer, 2017). Based on the mixed effects model, the effect of consonant type on preceding vowel duration was not significant in the current sample. Mean vowel durations before flaps, grouped by the type of flap sound (oral [ɾ] produced from underlying /t/ or nasal [ɾ̃] produced from underlying /n/) and the type of vowel uttered, are shown in Figure 14. The figure only displays the comparison between vowel durations in the cases where the distribution of values was comparable between pre-oral and pre-nasal vowels of the same quality (thus, /ej/ was excluded because there was only one occurrence of /ej/ before a flapped nasal [ɾ̃]). It can be observed that although there are differences between the durations of various types of vowels, there is no considerable difference based on whether the vowel is followed by an oral or a nasal flap, that is, whether the following consonant is underlyingly fortis or not.

Figure 14. Duration of vowels before flaps, grouped by the type of flap (oral [ɾ] or nasal [ɾ̃]) and the type of vowel produced



4. Conclusions

The findings of the production experiment show that the frequency of flapping of /nt/ clusters in nonsense words is considerably lower compared to that of both /t/ and /n/ flapping: the majority of subjects did not flap the cluster in any of their utterances. This substantiates the claim that lexical bias has a considerable effect on lenition processes such as flapping (Connine, 2004). Since the speakers had not had any previous exposure to the stimulus words, a flapped realization was not readily available in the lexicon for them to consider, thus they followed their inherent tendencies and articulated the cluster with a clearly distinct nasal and a hold+release sequence for the plosive. The only consistent lenition tendency observed in the nonsense words of the experiment was the flapping of /t/s, which is a well-established phenomenon in most dialects of American English. In the case of underlying /n/, not only did the frequency of flapping vary by speaker, but the perception judgments were also unsure about the occurrence of flapping in /n/ tokens more often than the other two underlying consonant types. The sample size and distribution of the current study did not allow for the comparison of flapped and unflapped realizations in all consonant types. The analysis of existing words would likely lead to a more balanced set of flapped and unflapped tokens, at least in the case of underlying /nt/; however, in this case, differences in word frequency would have to be accounted for in the analysis.

When it comes to the acoustic properties of nasal and oral flap sounds, the current sample did not exhibit any significant differences. Nasal [ɾ̃] was not realized with a longer duration than oral [ɾ], thus calling into question the previous findings regarding oral v. nasal flap duration (Garai, 2019). Regarding the spectral tilt (and derived nasality) of flaps, there was no significant difference observed between flap sounds produced from different underlying segments. Since H1*−H2* has not been confirmed as a reliable measure of nasality when it comes to flap sounds (only unflapped nasals and oral sonorants, see Garellek et al., 2016), there may be other measurable intensity cues which yield a clearer distinction between nasal and oral flaps. A closer inspection of spectral properties is necessitated in future research to identify these possible cues

Therefore, either [ɾ̃] is equally nasal as [ɾ], or H1*−H2* is not the most advantageous intensity measurement for the comparison of flap nasality. However, the nasality of [ɾ̃] is distinctly shown in the preceding vowel: vowels before nasal flaps were found to have lower values of A1*−P0 than before oral flaps, signaling a high level of nasality before [ɾ]. This follows from the spreading of nasality onto neighboring segments which is already present before flapping.

On the other hand, the results suggest that the effect of pre-fortis clipping cannot be observed on the current sample, as speakers did not articulate the target vowels with different durations based on following underlying consonant fortisness. However, the sample size, especially when broken down by vowel type is not nearly sufficient to draw definite conclusions on the matter. Further investigation is needed to determine how nasal consonants (and sonorants, in general) factor into the process of pre-fortis clipping, especially when it comes to lenited allophones such as flaps. It is important to note that the inclusion of sonorants in the process of pre-fortis clipping is not universally agreed upon, specifically because a sound pair such as /t/ and /n/ differ in various aspects other than fortisness, and there may be numerous factors (linguistic and otherwise) which may impact their duration. Furthermore, the flap-lengthening influence of high palatal vowels (as reported by Zue & Laferrière, 1979) was not observed in the current sample.

It is possible that the participants simply did not exhibit these phenomena, but it is more likely that a larger sample size would lead to more conclusive results. The effect of pre-fortis clipping before flaps, in particular, has been consistently shown between items with underlying /t/ and /n/, therefore it falls onto future research to investigate the possible differences between the duration of vowels before sonorant consonants and fortis/lenis obstruents.

Acknowledgement

I would like to thank Attila Starčević and Péter Őri for their perception judgments. Furthermore, I am grateful to Andrea Deme and Alexandra Markó for their valuable comments during the planning and implementation stages of the research. Supported by the ÚNKP-21-3 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund.

References

Audacity Team (2021). *Audacity(r): Free audio editor and recorder* (version 3.0.0). https://www.audacityteam.org/.

Berko, J. (1958). The child's learning of English morphology. *Word, 14(2),* 150–177.

Boersma, P. & Weenink, D. (2022). *Praat: doing phonetics by computer* (version 6.2.21). http://www.praat.org/. Accessed: 10-01-2022.

Braver, A. (2014). Imperceptible incomplete neutralization: Production, non-identifiability, and non-discriminability in american English flapping. *Lingua, 152,* 24–44.

Bunton, K., Hoit, J. D. & Gallagher, K. (2011). A simple technique for determining velopharyngeal status during speech production. *Seminars in Speech and Language, 32(1),* 69–80.

Connine, C. M. (2004). It's not what you hear but how often you hear it: On the neglected role of phonological variant frequency in auditory word recognition. *Psychonomic Bulletin & Review, 6(11),* 1084–1089.

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

Fujimura, O. & Lindqvist, J. (1971). Sweep-tone measurements of vocal-tract characteristics. The *Journal of the Acoustical Society of America, 49(2),* 541–558.

Garai, Luca (2019). *Flapped nasals, nasalized flaps in American English.* Paper presented at the 34th OTDK, Budapest, 24–25 April, 2019.

Garai, Luca (2021). Influencing factors of nasal flapping in English. In: Gráczi, Tekla Etelka & Ludányi, Zsófia (eds.), *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2021. Budapest: Nyelvtudományi Kutatóközpont.*

Garellek, M., Ritchart, A. & Kuang, J. (2016). Breathy voice during nasality: A cross-linguistic study. *Journal of Phonetics, 59,* 110–121.

Jacewicz, E., Fox, R. A. & Salmons, J. (2007). Vowel duration in three American English dialects. *Am Speech, 82(4),* 367–385.

Jensen, J. T. (1993). *English phonology.* Amsterdam: John Benjamins.

Kahn, D. (1976). *Syllable-Based Generalizations in English Phonology.* Master's thesis, Massachusetts Institute of Technology.

Kenyon, J. S. (1994). *American Pronunciation.* 12th ed. Ann Arbor: George Wahr Publishing Company.

Patterson, D. & Connine, C. M. (2001). Variant frequency in flap production. *Phonetica, 58(4),* 254–275.

Picard, M. (1997). English Flapping and the feature [vibrant]. *English Language and Linguistics, 1(2),* 285–294.

Python Software Foundation (2022). *Python* (version 3.10.7). http://www.python.org. Accessed: 10-01-2022.

Scheer, T. (2017). Voice-induced vowel lengthening. *Papers in Historical Phonology, 2,* 116–151.

Seabold, S. & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. In *Proceedings of the 9th Python in Science Conference*. (92–96).

Shue, Y.-L., Keating, P. A., Vicenik, C. & Yu, K. (2011). Voicesauce: A program for voice analysis. *Proceedings of the international congress of phonetic sciences*. (1846–1849). Hong Kong.

Steriade, D. (2000). Paradigm uniformity and the phonetics– phonology boundary. In Pierrehumbert, J. & Broe, M. (eds.), *Papers in laboratory phonology V: Acquisition and the lexicon,* (313–334). Cambridge: Cambridge University Press.

Styler, W. (2017). On the acoustical features of vowel nasality in English and French. *The Journal of the Acoustical Society of America*, *142(4)*, 2469–2482.

The pandas development team (2022). pandas (version 1.5.0). https://pandas.pydata.org/. Accessed: 10-01-2022.

Vaux, B. (2000). Flapping in English. *Linguistic Society of America,* 1–11.

Zue, V. W. & Laferrière, M. (1979). Acoustic study of medial /t, d/ in American English. *The Journal of the Acoustical Society of America, 66(4),* 1039–1050.